International Journal on Advances in Intelligent Systems

2026 vol. 19 nr. 1&2

The *International Journal on Advances in Intelligent Systems* is Published by IARIA.
ISSN: 1942-2679
journals site: http://www.iariajournals.org
contact: petre@iaria.org

Responsibility for the contents rests upon the authors and not upon IARIA, nor on IARIA volunteers, staff, or contractors.

IARIA is the owner of the publication and of editorial aspects. IARIA reserves the right to update the content for quality improvements.

Abstracting is permitted with credit to the source. Libraries are permitted to photocopy or print, providing the reference is mentioned and that the resulting material is made available at no cost.

Reference should mention:

International Journal on Advances in Intelligent Systems, issn 1942-2679
vol. 19, no. 1 & 2, year 2026, http://www.iariajournals.org/intelligent_systems/

The copyright for each included paper belongs to the authors. Republishing of same material, by authors or persons or organizations, is not allowed. Reprint rights can be granted by IARIA or by the authors, and must include proper reference.

Reference to an article in the journal is as follows:

<Author list>, "<Article title>"
International Journal on Advances in Intelligent Systems, issn 1942-2679
vol. 19, no. 1 & 2, year 2026, <start page>:<end page> , http://www.iariajournals.org/intelligent_systems/

IARIA journals are made available for free, proving the appropriate references are made when their content is used.

Sponsored by IARIA
www.iaria.org

Copyright © 2026 IARIA

International Journal on Advances in Intelligent Systems
Volume 19, Number 1 & 2, 2026

Editors-in-Chief

Hans-Werner Sehring, NORDAKADEMIE - University of Applied Sciences, Germany
Steve Chan, Decision Engineering Analysis Laboratory, USA

Editorial Board

Siby Abraham, NMIMIS Deemed to be University, Mumbai, India
Abdulelah Alwabel, Prince Sattam Bin Abdulaziz University, Saudi Arabia
Razvan Andonie, Central Washington University, USA
Andreas Aßmuth, Technical University of Applied Sciences OTH Amberg-Weiden, Germany
Isabel Azevedo, ISEP-IPP, Portugal
Mark Balas, Texas A&M University, College Station, USA
Leila Ben Ayed, National School of Computer Science | University la Manouba, Tunisia
Freimut Bodendorf, University of Erlangen-Nuremberg, Germany
Karsten Böhm, FH Kufstein Tirol - University of Applied Sciences, Austria
Ozgu Can, Ege University, Turkiye
Massimiliano Caramia, University of Rome Tor Vergata, Italy
Coskun Cetinkaya, Kennesaw State University, USA
Steve Chan, Decision Engineering Analysis Laboratory, USA
Mirela Danubianu, "Stefan cel Mare " University of Suceava, Romania
Leandro Dias da Silva, Universidade Federal de Alagoas, Brazil
Roland Dodd, CQUniversity, Australia
Larbi Esmahi, Athabasca University, Canada
Leila Ghomari, Ecole Nationale Supérieure des Technologies Avancées - ENSTA, Algiers, Algeria
Todorka Glushkova, University of Plovdiv, Bulgaria
Victor Govindaswamy, Concordia University Chicago, USA
Gregor Grambow, Aalen University, Germany
Maki K. Habib, The American University in Egypt, New Cairo, Egypt
Yousif A. Hamad, Siberian Federal University, Russia / Imam Ja'afar Al-Sadiq University, Iraq
Ridewaan Hanslo, University of Pretoria, South Africa
Tzung-Pei Hong, National University of Kaohsiung, Taiwan
Ahmed Kamel, Concordia College, Moorhead, USA
Rajkumar Kannan, Bishop Heber College (Autonomous), India
Leoneed M. Kirilov, Institute of Information and Communication Technologies - Bulgarian Academy of Sciences, Bulgaria
Dalia Kriksciuniene, Vilnius University, Lithuania
Panos Linos, Butler University, USA
Sandra Lovrencic, University of Zagreb, Croatia
Marcio Mendonça, Federal University of Technology - Paraná, Brazil
Harald Milchrahm, Technical University Graz | Institute for Software Technology, Austria
Shahab Mokarizadeh, Royal Institute of Technology (KTH), Stockholm, Sweden
Fernando Moreira, REMIT | Universidade Portucalense, Portugal
Joan Navarro, La Salle Campus Barcelona - Universitat Ramon Llull, Spain
Filippo Neri, University of Naples "Federico II", Italy
Rui Neves Madeira, Instituto Politécnico de Setúbal / Universidade Nova de Lisboa, Portugal
Andrzej Niesler, Wroclaw University of Economics and Business, Poland

Michel Occello, Université Grenoble Alpes, France
Matthias Olzmann, noventum consulting GmbH, Germany
Angel Pasqual del Pobil, Jaume I University, Spain
Agostino Poggi, University of Parma, Italy
Christoph Reich, NEC Laboratories America Inc., USA / TU Darmstadt, Germany
Fátima Rodrigues, Instituto de Engenharia | Politécnico do Porto, Portugal
José Rouillard, University of Lille, France
Minoru Sasaki, Ibaraki University, Japan
Ingo Schwab, Center of Applied Research (CAR) | University of Applied Sciences Karlsruhe, Germany
Hans-Werner Sehring, NORDAKADEMIE - University of Applied Sciences, Germany
Vasco N. G. J. Soares, Polytechnic Institute of Castelo Branco | Instituto de Telecomunicações, Portugal
Donald Sofge, Naval Research Laboratory (NRL), USA
Claudius Stern, FOM University of Applied Sciences for Economics and Management, Kassel, Germany
Sotirios Terzis, University of Strathclyde, UK
Mihaela Vranić, University of Zagreb, Croatia
Alexander Wijesinha, Towson University, USA
Yingcai Xiao, University of Akron, USA

International Journal on Advances in Intelligent Systems
Volume 19, Numbers 1 & 2, 2026

CONTENTS

pages: 1 - 13
Agent-Based Modeling in Energetic Refurbishment
Mart Verhoog, IU International University of Applied Sciences Cologne, Germany

pages: 14 - 30
Extending Process Virtualization Theory for Digital Collaboration: Social Presence, Situation Awareness, and Leadership in Virtual and AI-Augmented Teams
Inga F. Schlömer, IU International University, Germany

pages: 31 - 52
Learner Models: The Gap between Educational Opportunities vs. Data Protection Challenges
Felix Böck, Center for Responsible Artificial Intelligence (CRAI), Coburg University of Applied Sciences and Arts, Germany
Hendrik Link, Public Law, IT Law, and Environmental Law, University of Kassel, Germany
Dieter Landes, Center for Responsible Artificial Intelligence (CRAI), Coburg University of Applied Sciences and Arts, Germany

pages: 53 - 63
Converter-Level Modeling and Coordinated Control of a Wind–Battery Powered Surface Water Heat Pump DC Microgrid: Case Study of Ouistreham, France
Joyce Assaf, Groupe de Recherche en Electrotechnique et Automatique du Havre (GREAH), Université Le Havre Normandie, France
Mamadou-Baïlo Camara, Groupe de Recherche en Electrotechnique et Automatique du Havre (GREAH), Université Le Havre Normandie, France
Damien Guilbert, Groupe de Recherche en Electrotechnique et Automatique du Havre (GREAH), Université Le Havre Normandie, France
Brayima Dakyo, Groupe de Recherche en Electrotechnique et Automatique du Havre (GREAH), Université Le Havre Normandie, France

pages: 64 - 73
Soil Moisture Prediction and Benchmark Reliability: A Review with Multi-Site SMAP–ISMN Analysis
Kamrul Hasan, Grand Valley State University, USA
Cliff Mutegi, Grand Valley State University, USA
Arnold Muiruri, Grand Valley State University, USA
Puneeth Kumar Amudala, Grand Valley State University, USA
Andrew Njiyo, Grand Valley State University, USA

pages: 74 - 83
Improving Object Detection with Semi-Supervised Learning from Temporal Information
Viljar Holm Elvevoll, University of Bergen, Norway
Kim Tallaksen Halvorsen, Insitute of Marine Research, Norway
Ketil Malde, Insitute of Marine Research, Norway

pages: 84 - 93
Socially Responsible Artificial Intelligence in People and Pandemic Analytics: A Resilience-Oriented Framework

for Sustainable Decision-Making

Otabek Khamidov, Thomas Jefferson University, United States
Les Sztandera, Thomas Jefferson University, United States

Agent-Based Modeling in Energetic Refurbishment

Exploring Communication and Stakeholder Cooperation in Energy Transitions

Mart Verhoog
Marketing and Communication Department
IU International University of Applied Sciences
Cologne, Germany
e-mail: mart.verhoog@iu.org

Abstract— **Energetic refurbishment is essential for achieving European climate and energy targets, yet renovation rates remain below expectations. This paper presents an exploratory agent-based model of German homeowners' refurbishment decisions, structured by the Hierarchy of Effects model and grounded in survey data. The model distinguishes between mass communication (e.g., campaigns) and personal communication (e.g., consultancy) and analyzes their combined effects. Results show that mass communication alone increases awareness but has limited impact on intentions, whereas even moderate levels of personal communication strongly accelerate adoption, with diminishing returns at higher cost levels. More importantly, the findings demonstrate that communication effectiveness is regime-dependent: the impact of mass communication, personal communication, and social influence varies systematically with the overall level of communication intensity, giving rise to distinct diffusion dynamics. The study further shows that communication costs are distributed across multiple stakeholders and that effective acceleration requires coordinated action among governments, municipalities, consultants, and homeowners, concluding that refurbishment strategies must balance communication instruments and stakeholder cooperation, as isolated measures fail to generate efficient and scalable adoption.**

Keywords*— *Agent-Based Modeling (ABM); energetic refurbishment; mass and personal communication; stakeholder cooperation; energy transition.

I. INTRODUCTION AND MOTIVATION

A preliminary version of this model was presented in an earlier conference paper [1], where the focus was on identifying an "optimal mix" of communication measures. In contrast, the present paper adopts an explicitly exploratory perspective, extends the model with additional assumptions and sensitivity analysis, and provides a more comprehensive discussion of policy implications.

A. Policy and Practical Context

Energetic refurbishment has become a central pillar in international and European energy policy debates. Buildings are responsible for roughly 40% of total energy consumption in the European Union (EU) and approximately 36% of greenhouse gas emissions [2]. Within this context, improving the energy efficiency of residential buildings is considered a prerequisite for reaching the EU's climate-neutrality target by 2050 and for meeting the interim goals of the "Fit for 55" package. In Germany, the challenge is particularly pronounced: more than two thirds of the housing stock was constructed before the existence of energetic building standards, which results in disproportionately high heating demands and a large ecological footprint [3].

The urgency of accelerating refurbishment has grown since the outbreak of the war in Ukraine. The dependence of Germany and other European countries on Russian natural gas exposed significant vulnerabilities in energy security [1]. Sharp increases in energy prices since 2022 have further underscored the need to reduce demand in the heating sector and to promote energy efficiency measures at scale. Policymakers therefore emphasize not only ecological but also geopolitical and economic benefits of refurbishment.

Despite strong policy goals, refurbishment rates remain far below the required levels. Research and practice have documented several persistent barriers such as high upfront costs, long payback periods, behavioral inertia among homeowners, and concerns about consultants or contractors. [4]. These barriers often translate into hesitation, procrastination, or refusal to undertake necessary investments. Governments have traditionally relied on marketing-like instruments such as mass media campaigns to motivate homeowners. However, evidence suggests that such measures alone have achieved little impact in changing long-term refurbishment behavior [5]. In recent years, governments and local authorities have begun experimenting with sales-like measures — such as subsidized visits of energy consultants, on-site audits, or personalized advice. These measures introduce direct interaction with households and therefore hold potential to increase the effectiveness of policy interventions [6].

Understanding how different combinations of marketing and sales measures influence homeowners' decision-making is crucial. More broadly, it is necessary to study not only which measures are effective but also how their costs and responsibilities are distributed among multiple actors: governments, consultants, and homeowners themselves. This multi-actor dimension is central to the present work. However, existing approaches often focus on individual communication measures in isolation, without systematically consid-

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

ering how their effectiveness depends on the overall level of communication intensity.

B. ABM in the Energy Transition

ABM has emerged as a promising approach to study complex socio-technical transitions. ABM simulates the decisions and interactions of heterogeneous agents — such as households, firms, or policymakers — and allows researchers to observe how individual behaviors aggregate into system-level outcomes. The method is particularly well suited for exploring emergent phenomena, feedback effects, and nonlinear dynamics that are characteristic of the energy transition [7].

Over the past decade, ABM has been widely applied to study the adoption and diffusion of renewable energy technologies. For example, Baranzini et al. [8] analyzed the role of *social contagion* in the adoption of photovoltaic (PV) systems, Rai and Robinson [9] investigated the impact of policy incentives and peer effects on electric vehicle (EV) diffusion, and Derkenbaeva et al. [10] modeled household decisions related to decentralized energy systems in Amsterdam. These studies demonstrate that heterogeneous preferences, social influence, and local contexts can significantly shape technology adoption pathways.

However, the application of ABM to energetic refurbishment remains comparatively scarce. Most existing models focus on investment in new technologies (e.g., solar PV), while the more mundane but equally impactful decision to insulate a home, replace a heating system, or upgrade windows has received less attention [3]. Refurbishment decisions differ fundamentally from renewable adoption: they involve higher complexity, more intrusive interventions in existing buildings, and multiple technical and financial uncertainties. This gap in ABM research motivates the present study.

C. Distinguishing Marketing and Sales in Decision-Making

From a marketing science perspective, communication is not a monolithic concept but can be differentiated into *mass communication* (e.g., advertising campaigns, information brochures, social media outreach) and *personal communication* (e.g., face-to-face advice, consultancy visits, or sales talks). Classical theories such as the Hierarchy of Effects (HOE) model [11] emphasize that these different communication modes affect different stages of the decision-making process. Mass communication is effective in creating awareness and knowledge, whereas personal communication is crucial in later stages such as conviction and intention formation [12].

Research has repeatedly demonstrated that homeowners require different types of information at different stages of the refurbishment journey. Early phases are characterized by information gathering and the reduction of uncertainty. Later phases demand personalized guidance on financial, technical, and constructional aspects [13]. Mass communication alone cannot satisfy these needs, while personal communication is often too costly or resource-intensive to be scaled to entire populations.

Although ABM has been successfully used to simulate consumer behavior in marketing contexts [14], most existing models do not differentiate between marketing and sales. They treat communication as a generic process rather than a layered, staged phenomenon. This constitutes another gap addressed by the present paper: the integration of marketing and sales distinctions into an ABM of refurbishment decisions. In addition, the effectiveness of these communication mechanisms may depend not only on their functional role within the decision process, but also on their combined intensity within the system. This raises the question of how different levels of mass and personal communication jointly shape diffusion dynamics.

D. Stakeholder Cooperation in Refurbishment

A further underexplored dimension is the role of stakeholder cooperation. Energetic refurbishment requires the alignment of efforts across multiple actors:

- **Governments and municipalities** fund and design communication campaigns, create subsidy schemes, and regulate standards.
- **Energy consultants and contractors** provide personalized advice and execute technical measures.
- **Homeowners** make final investment decisions and bear substantial portions of the costs.

The interplay between these actors determines both the effectiveness of refurbishment programs and the distribution of costs and responsibilities. For example, governments may pay for mass media campaigns, while households finance consultancy services either directly or indirectly via subsidies. Understanding these distributions is critical for evaluating the fairness and feasibility of different policy approaches [5; 6; 13; 15].

In practice, lack of cooperation can lead to suboptimal outcomes: governments over-investing in mass campaigns with little impact, consultants facing limited demand, and homeowners feeling unsupported or overwhelmed. Conversely, cooperative approaches that align incentives and responsibilities can accelerate refurbishment. The present study takes this cooperation as a central lens, highlighting how different mixes of marketing and sales measures not only affect decision outcomes but also shape the cooperative dynamics between stakeholders.

E. Contribution of this Paper

This paper builds on a previous conference contribution by the author [1], which presented a first agent-based simulation of refurbishment decisions under different mixes of marketing and sales measures. While the conference paper introduced the basic model and preliminary results, the present journal article expands the analysis substantially in four directions:

- Exploratory framing – The model is presented not as a tool for strict optimization but as an exploration of plausible pathways. This framing allows for a more nuanced discussion of the directions the results may suggest, rather than prescribing single “optimal” solutions.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

- Transparency of assumptions and process – The modeling approach is clarified as a visual exploration of simulation outcomes rather than an algorithmic optimization. An explicit table of assumptions is introduced to enhance transparency.
- Stakeholder cooperation focus – The paper emphasizes that refurbishment acceleration depends on cooperation between governments, consultants, and homeowners, and that costs are ultimately shared among these stakeholders.
- Expanded content – The paper provides an enriched literature review, more detailed results, and enhanced figures and tables compared to the conference version.
- Interaction perspective – The analysis goes beyond isolated parameter effects by examining how different levels of communication intensity jointly shape diffusion dynamics and lead to distinct patterns of adoption.

Taken together, these extensions shift the contribution of the paper from identifying isolated parameter effects toward demonstrating that communication effectiveness is *regime*-dependent. The interaction between mass communication, personal communication, and social influence produces systematically different diffusion dynamics depending on the overall level of intervention intensity. This highlights that communication effects cannot be meaningfully assessed in isolation, but must be understood as part of an interacting system of mechanisms.

The remainder of this article is structured as follows. Section II presents the expanded literature review on ABM, refurbishment decision-making, and stakeholder cooperation. Section III introduces the conceptual model, the assumptions, and the data used for the simulation. Section IV presents the results of the exploratory analysis, with special attention to stakeholder cost distribution. Section V discusses implications for research, practice, and policy. Section VI concludes by summarizing the contributions and outlining future work.

II. LITERATURE REVIEW

This section reviews the state of the art, outlines relevant research streams, and identifies gaps that motivate the present study.

A. Energetic Refurbishment and Policy Context

Energetic refurbishment has become a cornerstone of European climate and energy policy. The European Commission's Renovation Wave strategy explicitly calls for doubling annual renovation rates by 2030 to meet climate neutrality targets [2]. In Germany, more than two thirds of the housing stock predates energetic building standards, contributing to high heating demand and carbon emissions [3].

Multiple barriers explain the sluggish pace of refurbishment. These include high upfront costs, long payback times, and behavioral inertia among homeowners [4]. Financing mechanisms are increasingly seen as a way to overcome these obstacles. Bertoldi et al. [16] provide a comprehensive overview of emerging instruments in the EU, while Nykamp [17] emphasizes the importance of policy mixes to align incentives.

Recent research highlights the role of local governance. Ebrahimigharehbaghi et al. [18] show how municipal approaches in the Netherlands can accelerate retrofitting of owner-occupied homes. Jowkar et al. [19] stress that social factors shape barriers and motivations in Norwegian refurbishment projects. Together, these studies suggest that refurbishment must be understood not only as a technical or economic decision but as a socio-political process requiring cooperation across multiple levels of governance.

B. ABM in Energy Transitions

ABM is increasingly recognized as a powerful tool for studying socio-technical transitions. ABM simulates the interactions of heterogeneous agents and captures emergent dynamics that traditional equilibrium models cannot represent [7].

A large body of ABM research has focused on renewable energy adoption. Du et al. [8] reviewed applications in solar PV and other domains, highlighting the importance of socio-demographic factors and peer effects. More recently, Derkenbaeva et al. [10] developed an ABM of household transitions in Amsterdam, incorporating spatial and social dynamics. Van der Kam et al. [20] constructed an empirical ABM of consumer co-adoption of low-carbon technologies, providing insights into policy mixes.

When it comes to refurbishment, the literature is still limited. Du, Han & de Vries [3] provide a review of renovation adoption and diffusion models, calling for more empirical applications. Chen, Fan & Yao [21] modeled evolutionary dynamics of homeowners' retrofit decisions in complex networks, offering methodological advances that are directly relevant to the present study. Shen et al. [22] combined ABM with performance-based modeling to design integrated retrofit frameworks. These examples illustrate the growing but still fragmented use of ABM in the refurbishment domain.

An emerging strand of work explores methodological innovation. Tian et al. [23] integrate ABM with deep learning to simulate household energy consumption, while Khalil & Fatmi [24] use ABM to examine behavioral changes during the COVID-19 pandemic. Such studies suggest that ABM can be used not only predictively but also exploratorily, to probe the potential outcomes of different interventions under uncertainty. This exploratory use is the perspective adopted in the present paper.

C. Marketing and Sales in Consumer Decision-Making

In marketing science, the HOE model has long been used to describe how communication influences consumer decision-making [11]. Early stages such as awareness and knowledge are typically influenced by mass communication, while later stages such as conviction and intention depend more on personal communication [12; 25; 26].

Simulation studies have begun to operationalize these ideas. Jager [14] suggested how the 4 P's of marketing could be incorporated into social simulations, and Delre et al. [27] developed an ABM of promotional activities. However, most

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

ABM applications in consumer behavior still treat communication as a homogeneous influence, without distinguishing between marketing and sales.

Recent empirical work on sustainable consumption underscores the relevance of this distinction. Chen [28] analyzed how integrated communication tools affect green product purchase intentions. Correia et al. [29] demonstrated the influence of green marketing communication on actual purchase behavior. Krsnik and Erjavec [30] incorporated lifestyle and psychological factors to explain differences in consumer responses - an approach conceptually similar to the Sinus lifestyle segmentation used in the present model. Iliopoulou et al. [31] highlight the role of social media engagement in translating knowledge into sustainable action, underscoring the evolving nature of mass communication channels.

Together, these studies suggest that communication strategies must be tailored not only to different stages of the decision process but also to different consumer segments. This reinforces the rationale for incorporating both mass and personal communication in the ABM presented here.

D. Stakeholder Cooperation in Energy Efficiency

The acceleration of energetic refurbishment depends on the coordinated actions of multiple stakeholders. The governance literature identifies several dimensions of cooperation that are particularly relevant in this context.

One important perspective is multilevel governance. Björklund et al. [32] synthesize more than two decades of EU experience with energy-efficient and zero-carbon buildings and emphasize that successful policy implementation requires alignment across multiple governance levels — European, national, regional, and municipal. Isolated initiatives at one level often remain ineffective without such vertical coordination.

A second dimension concerns the role of municipal governance. Local authorities are often the actors closest to homeowners and can provide tailored support. Ebrahimigharehbaghi et al. [18] demonstrate that municipal programs in the Netherlands have been central to encouraging retrofits of owner-occupied homes, for instance by facilitating personal advice and acting as trusted intermediaries.

Cooperation is not limited to public authorities. Research also points to the role of private actors and hybrid governance arrangements. Galvin [15] shows that private housing companies in Germany increasingly integrate energy efficiency into their strategic agendas, thereby assuming functions traditionally associated with public governance. Such hybrid approaches indicate that responsibilities and resources can be distributed across state and non-state actors.

Finally, cooperation is also shaped by social contexts. Jowkar et al. [19], for example, identifies how social values influence barriers and motivations for renovation in Norway. This suggests that governance models must be sensitive to context rather than assuming uniform homeowner behavior.

Taken together, the literature highlights that energetic refurbishment requires multi-actor alignment, with complementary roles for governments at different levels, municipalities, private stakeholders, and socially embedded motivations. While each of these strands of research has provided valuable insights, there is still limited integration across them, particularly in combination with quantitative modeling approaches.

III. ABM FOR REFURBISHMENT

This section introduces the ABM, including its conceptual structure, empirical basis, assumptions, and simulation design.

A. Model Overview

The study builds on an ABM designed to investigate how different combinations of mass communication and personal communication influence homeowners' decision-making regarding energetic refurbishment. A preliminary version of this model was presented in an earlier conference paper [1], where the focus was on identifying an "optimal mix" of measures. While such a framing provided a concise entry point for the conference format, it implied a degree of precision that does not fully reflect the exploratory character of the model.

In the present journal article, the same model is taken further and deliberately reframed as an exploratory tool. Rather than prescribing an optimal solution, the simulations are used to compare plausible scenarios and to reason about their implications for policy and stakeholder cooperation. This shift reflects both the limitations of the assumptions (which are numerous and simplifying) and the opportunities to use ABM as a way of probing into the dynamics of decision-making under uncertainty.

The model represents 17,000 simulated households, segmented according to the German Sinus lifestyle typology. Within the model, mass communication (MC) and personal communication (PC) denote the corresponding communication variables and their intensity levels, while the terms mass communication and personal communication are also used in the broader conceptual sense to describe the two communication forms. Each agent interacts with its environment and is exposed to varying levels of communication efforts. Their decision-making follows the HOE framework, which distinguishes between stages such as awareness, knowledge, attitude, conviction, and intention. This allows the model to capture not only whether households decide to refurbish, but also how communication influences different stages of the process.

In addition to communication, the model incorporates neighborhood-level social influence, reflecting evidence that refurbishment decisions can be accelerated when neighboring households have already taken positive action. The combined setup thus enables the exploration of how marketing-like and sales-like measures, together with social contagion, may interact to shape refurbishment pathways.

B. Decision-Making Framework

The decision-making logic of the model builds on the HOE model, which conceptualizes consumer behavior as a sequence of stages leading from awareness to final purchase intention. In the context of energetic refurbishment, this framework is adapted to represent how homeowners move

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

through the following stages: awareness, knowledge, attitude, conviction, and intention. Each agent in the simulation progresses through these stages over time, influenced by exposure to communication measures and by social interaction with neighboring agents.

Figure 1 illustrates the adapted HOE model for energetic refurbishment. In this representation, mass communication is assumed to exert its strongest influence in the early stages of awareness and knowledge, but its effect diminishes sharply in the later stages. By contrast, personal communication is relatively weak in the initial stages but becomes decisive in the conviction and intention phases, when households require reassurance, tailored advice, and concrete guidance for action.

To formalize this mechanism, each agent is equipped with a *threshold* variable that accumulates influence from different touchpoints. Once the threshold reaches a defined value, the agent progresses to the next stage of the HOE model. This stepwise updating process is applied iteratively in each simulation period, allowing agents to accumulate communication effects and social influence over time. The strength of each communication touchpoint depends on both its type (mass or personal) and the stage of the decision process. Correction factors are applied to represent the decreasing impact of mass communication and the increasing impact of personal communication across successive stages.

In addition to communication, the model integrates social contagion effects. When one household in a neighborhood decides positively on refurbishment, its neighbors' threshold variables are temporarily increased. This mechanism captures localized peer influence, which has been documented in related energy adoption studies. The resulting setup reflects a multi-channel decision environment, in which homeowners' trajectories are shaped by both institutional communication and social dynamics.

C. Data and Sample

The ABM is empirically grounded in data from the Sinus lifestyle typology, a widely used segmentation of German society that combines sociocultural orientation with social status. The typology distinguishes ten distinct lifestyles, ranging from traditional to progressive orientations and from lower to upper socioeconomic status. Previous research has shown that these lifestyles differ substantially in their attitudes and behaviors toward energetic refurbishment [6].

The empirical basis of the model is a representative survey of n = 4,471 German homeowners, which provides information on the status quo of the refurbishment decision-making process for each lifestyle group. The data capture the proportion of homeowners at different stages of the HOE model (awareness, knowledge, attitude, conviction, and intention). For example, adaptive pragmatics show a relatively high likelihood of forming refurbishment intentions (22.4%), while precarious households demonstrate much lower intention levels (9.3%). These differences highlight the importance of sociocultural segmentation for understanding heterogeneous responses to communication measures.

Figure 1. Hierarchy of effect model for energetic refurbishment [33]

Based on these survey results, the model initializes a population of 17,000 simulated homeowners. The distribution across Sinus lifestyles is scaled to reflect their prevalence in German society. Each agent is assigned an initial position in the decision-making process according to the empirical distributions reported in Table I. This approach ensures that the simulation starts from a realistic baseline, with heterogeneous agents reflecting actual sociocultural patterns in Germany.

By grounding the ABM in representative data, the model captures not only aggregate tendencies but also lifestyle-specific differences in responsiveness to communication. This provides a richer foundation for exploring how different mixes of marketing-like and sales-like measures influence the trajectories of diverse homeowner groups.

D. Assumptions

As with any ABM, the present simulation relies on a set of assumptions that specify how agents behave, how communication affects decision-making, and how costs are calculated. To increase transparency, these assumptions are explicitly listed in Table II. While they may not capture all real-world complexities, they provide a consistent framework for comparing alternative communication scenarios. The number and scope of assumptions also underline that the model should be interpreted as exploratory rather than predictive.

E. Simulation Design

The simulation experiments were designed to explore how different levels of mass communication (MC) and personal communication (PC) influence the development of refurbishment intentions among German homeowners. Both communication variables were systematically varied across five levels (0–5), resulting in a matrix of possible combinations.

TABLE I. EMPIRICAL DATA FROM [33]

German's Sinus lifestyle	Status quo in DM				
	Aware-ness [%]	Know-ledge [%]	Atti-tude [%]	Con-viction [%]	Inten-tion [%]
Homeowners in general	98.7	77.1	48.5	33.5	14.3
1. Established conservat.	99.0	80.1	51.0	33.7	14.9
2. Liberal-intellectual	99.7	82.9	57.9	40.6	13.3
3. High achiever	98.9	79.5	48.4	32.9	18.4
4. Movers and shakers	98.4	76.9	52.3	34.6	16.6
5. Adaptive pragmatics	99.5	69.0	53.9	42.4	22.4
6. Social-ecological	99.7	80.0	52.3	36.2	17.0
7. New middle class	98.5	76.9	44.1	32.6	11.8
8. Traditional	99.0	78.8	44.5	30.8	10.6
9. Precarious	99.6	64.4	36.1	20.6	9.3
10. Escapists	92.8	59.3	33.7	21.9	12.2

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

TABLE II. KEY ASSUMPTIONS FOR MODEL INITIALIZATION AND SIMULATION

Variable	Value / Assumption	Source
Decision model	HOE stages: awareness → knowledge → attitude → conviction → intention	[11; 12; 33]
Threshold for stage change	4	Assumed
Mass comm. impact	Declines by stage (1.0; 0.5; 0.0; 0.0)	[25; 26]
Personal comm. impact	Increases by stage (0.0; 0.0; 0.5; 1.0)	[25; 26]
Social contagion	+30% per neighbor, max. +120%	[8]
Simulation period	2023–2050 or until 95% refurbishment intention	Policy target
Agents	17,000 homeowners (scaled from n = 4,471 survey)	[3; 33]
Mass comm. cost	€ 13.50 CPM	[3]
Personal comm. cost	€ 16.67 per touchpoint	[3]
Consultant capacity	30,000 consultants; max. 3M touchpoints/year	[3]

For each combination, the model was run for a period of 27 years (2023–2050), corresponding to the timeframe of European climate neutrality targets. In addition, a second set of experiments was conducted in which the model was terminated once 95% of households had developed refurbishment intentions. This dual approach makes it possible to compare scenarios with respect to both cumulative outcomes and the speed of diffusion.

To account for stochasticity in agent-based simulation, each scenario was repeated 20 times. The analysis therefore relies on 720 observations (36 combinations × 20 repetitions). For each observation, the model recorded:

- the distribution of homeowners across HOE stages,
- the proportion of agents reaching refurbishment intention,
- the time required to reach 95% refurbishment intention (where applicable), and
- the accumulated economic costs of mass and personal communication.

These costs are calculated at an aggregate (macro) level for each scenario, enabling a direct comparison between effectiveness and cost efficiency. Optimization of outcomes was not conducted algorithmically. Instead, the results were evaluated visually and comparatively across scenarios. This exploratory procedure makes it possible to identify plausible pathways, highlight trade-offs between speed and cost, and reason about how communication measures may be distributed among different stakeholders.

The model was implemented in NetLogo, a widely used platform for ABM. The implementation follows a modular structure, allowing systematic variation of key parameters and reproducible simulation runs. The model implementation is available on GitHub (under https://github.com/mart-verhoog/RefurbishmentModel).

IV. RESULTS

This section presents the simulation results, focusing on the effects of different communication scenarios and key parameter variations.

A. Overall Development of Refurbishment Intentions

The first experiments examined the overall development of refurbishment intentions over the simulation period of 27 years (2023–2050). Figure 2 summarizes the proportion of homeowners that reached the intention stage under different combinations of mass communication (MC) and personal communication (PC). This figure focuses on the final level of refurbishment intention reached after the simulation period.

The results show that mass communication alone is insufficient to bring about substantial change. Starting from a baseline of 14.3% refurbishment intention (Table I), sustained exposure to maximum levels of mass communication (MC = 5, PC = 0) increases this share only marginally, reaching 18.6% after 27 years. This finding echoes observations in practice that large-scale campaigns, while important for raising awareness, have limited impact on actual decision-making.

In contrast, the inclusion of personal communication strongly accelerates progression through the decision-making process. Even modest levels of personal communication (PC = 1) combined with high mass communication (MC = 5) lead to near saturation, with over 99% of households reaching refurbishment intention by 2050.

Scenarios with higher levels of personal communication (PC = 2–3) achieve similar results but at higher costs, as will be discussed in Section IV.C.

The results also confirm that personal communication cannot fully substitute for mass communication. Scenarios with high personal communication but low or absent mass communication (e.g., MC = 0, PC = 3) produce higher overall intentions than MC-only cases, but they fail to achieve saturation within the simulated timeframe. This pattern indicates that mass communication plays a necessary role in creating initial awareness and knowledge, which then enables personal communication to be effective in the later stages of decision-making.

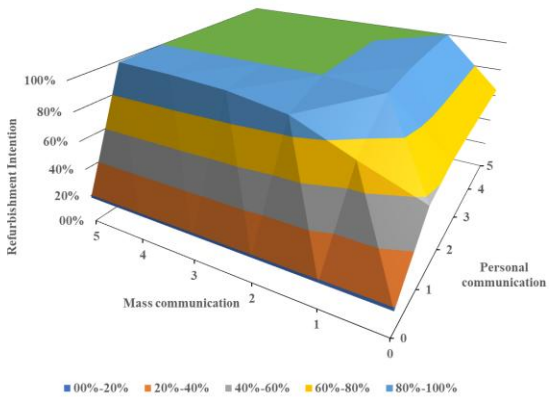


Figure 2. Final refurbishment intention (%) after 27 years across communication scenarios

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Overall, these findings suggest that the interaction of mass and personal communication is critical. Mass communication alone is inefficient, and personal communication alone is insufficient, but the combination of the two produces the strongest outcomes. This dynamic provides the basis for further analysis of the speed and cost trade-offs in subsequent sections.

B. *Speed of Reaching 95% Refurbishment Intention*

In addition to analyzing outcomes after a fixed simulation period of 27 years, the model was also used to examine how quickly different communication strategies enable the system to approach saturation. For this purpose, the simulation was terminated once 95% of households had reached the refurbishment intention stage. This indicator captures the speed of diffusion and allows a more direct comparison between alternative mixes of mass and personal communication.

The results reveal a clear trade-off between the level of personal communication and the speed of adoption. Scenarios that rely primarily on mass communication fail to reach the 95% threshold within the simulated timeframe, regardless of intensity. By contrast, the addition of even modest levels of personal communication substantially accelerates diffusion. For example, a combination of MC = 5 and PC = 1 achieves the 95% threshold in less than 20 years, while MC = 5 with PC = 2 shortens this period to around 15 years. Increasing personal communication further reduces the time to saturation, but the marginal gains diminish as PC levels rise beyond 3.

These dynamics are illustrated in Figure 3, which plots the average number of years required to reach 95% refurbishment intention under different communication mixes. In contrast to the previous results, this figure focuses on the speed of diffusion, measured as the number of years required to reach a 95% intention level. The pattern highlights that mass communication alone cannot achieve saturation, whereas the incremental addition of personal communication quickly reduces the time required. At the same time, higher PC levels imply higher costs, raising important questions about efficiency and burden-sharing among stakeholders.

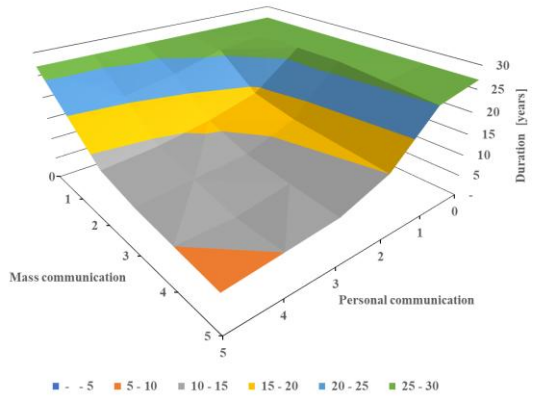


Figure 3. Time required (in years) to reach 95% refurbishment intention across communication scenarios

To quantify these implications, Table III presents the additional communication costs associated with faster adoption pathways. The results suggest that a moderate investment in personal communication offers the most favorable balance between speed and cost, while excessive reliance on personal contact delivers limited time savings at disproportionately higher expense.

C. *Economic Costs of Communication*

The acceleration of refurbishment intentions through communication measures generates additional costs, which differ by type and scale. Their distribution is critical for evaluating the feasibility of alternative strategies.

As shown in Table III (Section IV.B), scenarios with higher levels of personal communication result in disproportionately higher expenditures compared to those relying mainly on mass communication. While the addition of modest levels of PC (e.g., PC = 1) provides a substantial acceleration of adoption, it also raises costs by several hundred million euros over the simulation horizon. Increasing PC further produces strongly diminishing returns, with limited time savings at much higher expense.

Mass communication, by contrast, is relatively inexpensive per exposure but has limited effect on intention formation without personal communication. This implies that mass communication is necessary but not sufficient, whereas personal communication is effective but costly.

An important consideration is that the expenditures are not borne by a single actor. In practice, governments and municipalities typically finance mass campaigns, while homeowners or public subsidy schemes cover parts of the costs of consultancy. Energy consultants themselves also face capacity costs, often supported by professional associations or specific subsidy instruments. To illustrate these responsibilities, Table IV summarizes the typical allocation of cost responsibilities among stakeholders.

TABLE III. ADDITIONAL COSTS FOR FASTER REFURBISHMENT

Combinations of MC and PC		Characteristics of combination		
MC	PC	Duration [years]	Accumulated costs [mi. €]	Additional costs vs. MC = 5 / PC = 1[mi. €]
5	5	9	460.4	181.7
4	4	10	409.3	130.6
5	4	10	411.6	132.9
4	5	10	509.3	230.6
5	3	11	342.7	64.0
3	5	11	557.4	278.7
4	3	12	371.1	92.4
3	4	12	488.4	209.7
3	3	13	399.0	120.3
2	5	13	656.1	377.4
2	4	14	566.0	287.3
5	2	15	317.3	38.6
4	2	16	334.8	56.1
2	3	16	487.4	208.7
3	2	17	351.8	73.1
2	2	19	388.8	110.1
1	5	20	1,004.8	726.1
1	4	21	845.0	566.3
1	3	23	696.4	417.7
1	2	26	526.7	248.0

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

It is important to emphasize that Table IV does not assign exact numerical shares. The model calculates only aggregate costs, while the table provides a qualitative illustration of how these costs are typically distributed in practice. This perspective underscores the role of stakeholder cooperation in financing communication strategies for refurbishment acceleration.

D. Segment-Specific Patterns

While aggregate outcomes provide an overall picture of refurbishment dynamics, the segmentation of households according to Sinus lifestyles reveals important differences in responsiveness to communication measures. Table I in Section III showed the heterogeneous baseline distribution of homeowners across the HOE stages. Building on this initialization, the simulation results indicate that communication strategies affect lifestyle groups in uneven ways.

Households from higher-status, future-oriented lifestyles such as adaptive pragmatics and high achievers respond more quickly to both mass and personal communication. In these groups, modest levels of personal communication in combination with high mass communication are sufficient to achieve near saturation well before 2050. Their baseline intention levels are already above average, and communication measures therefore build on an existing disposition to act.

By contrast, groups such as the precarious and traditional lifestyles demonstrate slower progression through the HOE stages. Even with strong personal communication efforts, saturation is not achieved within the simulated period. These households often remain stuck in the conviction stage, reflecting structural barriers such as financial constraints. The results suggest that for these groups, communication alone may not be sufficient; additional policy instruments, e.g., financial support, could be required.

Intermediate groups such as the established conservatives and liberal intellectual fall between these extremes. They reach high intention levels under combinations of mass and personal communication, but the speed of adoption is slower than in more progressive groups. This indicates that while they are not fundamentally resistant, they may require sustained exposure over time.

Taken together, the segment-specific patterns highlight the importance of tailoring communication mixes to different homeowner groups. A uniform strategy may produce strong aggregate outcomes but mask pockets of low responsiveness. Recognizing heterogeneity is crucial for designing efficient and equitable refurbishment policies. A detailed numerical breakdown by lifestyle segments is beyond the scope of this paper.

TABLE IV. TYPICAL DISTRIBUTION OF COMMUNICATION COSTS ACROSS STAKEHOLDERS

Communication type	Typical stakeholder(s) bearing costs
Mass communication campaigns	National government, municipalities
Personal communication (consultant visits)	Homeowners (direct payment), subsidy schemes
Consultant capacity and infrastructure	Consultants, professional associations, partly public support

E. Exploratory Insights

The results of the simulation are not intended to deliver precise forecasts or to prescribe an optimal communication strategy. Instead, they provide a structured exploration of plausible pathways, highlighting trade-offs and dependencies that are relevant for both policy and practice.

A first key insight is that mass communication and personal communication are complementary rather than substitutable. Mass campaigns are necessary to create broad awareness and knowledge, but they remain ineffective without the reinforcing role of personal advice in the later stages of decision-making. Conversely, personal communication is powerful in driving conviction and intention but cannot generate sufficient reach on its own.

A second insight is that speed and cost are inherently linked. Faster adoption requires more personal communication, which is costly and resource-intensive. Moderate levels of PC offer the best balance between speed and cost, while very high levels of PC accelerate adoption only marginally but at disproportionately higher expense.

A third insight concerns stakeholder cooperation. The model makes clear that communication costs are substantial and cannot realistically be borne by a single actor. Governments are best positioned to finance mass campaigns, while homeowners and subsidy schemes contribute to the costs of consultancy, and consultants themselves provide the necessary capacity. Effective acceleration therefore depends on shared responsibility and coordination among stakeholders.

Finally, the analysis underscores the exploratory value of ABM. By making assumptions explicit and systematically varying key parameters, the approach allows researchers to reason about possible outcomes, tensions, and leverage points. Rather than predicting an optimal solution, the model highlights the conditions under which refurbishment pathways may accelerate or stall, and it provides a transparent basis for policy reflection.

F. Sensitivity Analysis

To assess the robustness of the results presented above, a targeted sensitivity analysis was conducted. The core research question of this study inherently adopts a sensitivity perspective by examining how varying levels of mass and personal communication influence decision outcomes. Building on this, two key parameters were systematically varied: the decision threshold and the strength of social contagion.

The analysis was performed under three communication scenarios representing different intervention intensities: low (MC = 2, PC = 0), medium (MC = 3, PC = 1), and high (MC = 4, PC = 3). For each scenario, parameters were varied *ceteris paribus*. The threshold parameter was tested at three levels (3, 4, 5) while keeping social contagion constant at 30%, and vice versa (15%, 30%, 45%) while keeping the threshold fixed at 4. Each configuration was simulated 20 times, and mean values were reported for the final share of homeowners reaching refurbishment intention and the number of years required to reach 95% intention. Table V presents the results of the threshold sensitivity analysis, while Table VI summarizes the results for social contagion.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

TABLE V. SENSITIVITY OF MODEL OUTCOMES TO VARIATION IN THE THRESHOLD PARAMETER

Scenario	MC	PC	Threshold	Final Intention (%)	Years to 95%
A	2	0	3	100.0	6
A	2	0	4	94.9	10
A	2	0	5	59.8	NA
B	3	1	3	100.0	5
B	3	1	4	100.0	6
B	3	1	5	100.0	8
C	4	3	3	100.0	4
C	4	3	4	100.0	4
C	4	3	5	100.0	5

TABLE VI. SENSITIVITY OF MODEL OUTCOMES TO VARIATION IN SOCIAL CONTAGION STRENGTH

Scenario	MC	PC	Contagion	Final Intention (%)	Years to 95%
A	2	0	0.15	65.2	NA
A	2	0	0.30	60.8	NA
A	2	0	0.45	57.8	NA
B	3	1	0.15	100.0	7
B	3	1	0.30	100.0	6
B	3	1	0.45	100.0	6
C	4	3	0.15	100.0	5
C	4	3	0.30	100.0	4
C	4	3	0.45	100.0	4

Note: "NA" indicates that 95% adoption was not reached by the end of the simulation period.

The results show that the effect of parameter variation depends strongly on the level of communication intensity. Under low communication intensity (MC = 2, PC = 0), the threshold parameter has a substantial impact on model outcomes, determining whether diffusion occurs at all. Under moderate intensity (MC = 3, PC = 1), all configurations reach full adoption, but the threshold influences the speed of diffusion. Under high communication intensity (MC = 4, PC = 3), the model becomes largely insensitive to threshold variation, with rapid convergence in all cases.

In contrast, variation in social contagion shows only minor effects across all scenarios. While small differences can be observed under low communication intensity, the overall influence of this parameter remains limited compared to the threshold.

Overall, the results indicate that the model exhibits context-dependent robustness. While outcomes are sensitive to parameter variation under low communication intensity, the main patterns remain stable under moderate and high intensity. This suggests that the observed dynamics reflect structural properties of the model rather than being driven solely by specific parameter choices. At the same time, the analysis highlights that sensitivity is not uniformly distributed across parameters and regimes, but follows a consistent and interpretable pattern. In particular, individual resistance—captured by the threshold parameter—becomes critical under low-intensity conditions, while diminishing in relevance as communication pressure increases. Taken together, this combination of structured sensitivity and stability supports the validity of the model in exploring communication-driven diffusion processes.

V. DISCUSSION

This section discusses the findings in relation to existing literature, derives implications for policy and practice, and reflects on the limitations of the model.

A. Integration with Existing Literature

The simulation findings resonate with and extend several strands of existing research in marketing, energy transitions, and ABM.

First, the model's differentiation between mass communication and personal communication aligns closely with classical marketing theory. The HOE model [11; 12] and subsequent marketing research emphasize that mass communication is primarily effective in the early stages of awareness and knowledge, while personal communication is crucial in the conviction and intention stages [25; 26]. The results of this study empirically support these theoretical expectations in the specific domain of energetic refurbishment.

Second, the findings connect to the growing body of research on the role of energy consultancy and advice. Studies such as [5; 6; 13] have shown that personalized advice is often decisive for overcoming barriers to renovation. Our model results mirror this insight, demonstrating that even modest levels of personal communication trigger sharp increases in refurbishment intentions, whereas mass campaigns alone achieve only marginal effects.

Third, the results contribute to the literature on ABM in energy transitions. Previous applications have largely focused on technology adoption, such as photovoltaic systems [8]. Renovation has been underexplored, despite its equal or greater importance for climate targets. By applying ABM to energetic refurbishment and explicitly modeling communication dynamics, this study extends the methodological repertoire of ABM applications in the energy domain.

Finally, the exploratory framing of the model reflects a broader methodological trend. As noted by [24], ABM is increasingly recognized not only as a predictive tool but also as a means to support exploratory analysis and policy learning under uncertainty. The present study adopts this perspective explicitly, using the model to reason about plausible pathways, trade-offs, and stakeholder responsibilities rather than to prescribe a single optimal solution.

Taken together, the results extend existing literature on communication-driven diffusion processes by suggesting that the effects of communication mechanisms are not uniform, but depend on the overall level of intervention intensity. While prior research emphasizes the role of mass communication, interpersonal influence, and social contagion, the present findings indicate that their relative importance varies systematically across different regimes. In particular, the results suggest that parameter effects should not be interpreted in isolation, but in relation to the broader communication context in which they operate.

B. Implications for Policy and Practice

The simulation results provide several implications for policy design and practical implementation. They show that communication measures are necessary for accelerating refurbishment but that their effectiveness depends critically on

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

cooperation among stakeholders and coordination across governance levels.

First, the findings highlight the importance of multilevel governance. Björklund et al. [32] emphasize that energy renovation policies in the EU succeed only when European, national, and local actions are aligned. In the simulation, national-level mass communication raises awareness but fails to generate substantial refurbishment intentions on its own. Only when combined with personal communication—often facilitated at the municipal level—does the system accelerate toward saturation. This indicates that higher-level campaigns must be complemented by localized, trusted interventions.

Second, the results underscore the leverage of municipal governance. Ebrahimigharehbaghi et al. [18] demonstrate how municipalities in the Netherlands mobilize homeowners by offering targeted programs and advice. The simulation confirms this mechanism: even modest amounts of personal communication (PC = 1) are sufficient to dramatically speed up adoption. This suggests that small-scale municipal initiatives, such as subsidized consultation visits or neighborhood information events, may provide high returns relative to their cost.

Third, the model points to the role of private actors and hybrid governance. Galvin [15] argues that private housing companies increasingly integrate energy efficiency goals into their corporate strategies, taking on responsibilities traditionally reserved for public authorities. The cost structures in the simulation reflect this: while governments typically finance mass campaigns, the expenses of consultancy are partly borne by homeowners and subsidy schemes. Consultants themselves face capacity costs. These findings indicate that hybrid arrangements—where public and private actors share the financial and organizational burden—are not only realistic but necessary for large-scale acceleration.

Finally, the results suggest the importance of social contexts. Jowkar et al. [19] find that social values influence barriers and motivations for renovation in Norway. In the model, lifestyle groups such as precarious households lag behind others despite strong communication measures, reflecting deeper structural and social barriers. This implies that communication alone cannot reach all groups equally and must be embedded in broader strategies that also address financial constraints and trust.

Taken together, these insights point to a policy approach that avoids reliance on a single "optimal" communication instrument. Instead, policymakers should pursue policy mixes that combine broad awareness-building through mass campaigns with personalized, trust-based interactions at the local level, and that mobilize both public and private actors in financing and implementation. The simulation shows that such cooperation is not optional but essential for achieving large-scale acceleration in a cost-effective and socially inclusive manner.

The results further suggest that policy interventions should be aligned with the prevailing communication intensity within the system. Under low-intensity conditions, efforts to reduce individual barriers to adoption—such as financial constraints or informational deficits—are particularly critical, as these factors can determine whether diffusion occurs at all. In contrast, under moderate communication intensity, increasing the level or coordination of communication activities can significantly accelerate diffusion processes. Finally, under high-intensity conditions, additional communication efforts may yield diminishing returns, as adoption levels are already close to saturation. This highlights the importance of tailoring policy strategies to the specific diffusion regime rather than applying uniform intervention approaches.

C. Limitations of the Model

While the simulation provides valuable exploratory insights, several limitations must be acknowledged. These limitations relate to the scope of the assumptions, the operationalization of decision-making, and the boundaries of the model.

First, the model is based on a substantial number of simplifying assumptions. Table II made these explicit, covering thresholds, communication impacts, social contagion, and cost structures. Such assumptions are necessary to make the model tractable but inevitably reduce its realism. For example, the influence of communication measures is represented by fixed correction factors, which may not fully capture the diversity of responses across households. Similarly, the cost parameters are treated as stable averages, whereas in reality they vary across regions, time, and communication channels. In addition, the threshold parameter represents a stylized proxy for individual resistance to behavioral change and is not directly calibrated to empirical data, which limits its interpretability in real-world terms.

More fundamentally, the model deliberately abstracts from structural barriers such as high renovation costs (typically €50,000–€150,000) and demographic constraints, e.g., older homeowners facing long payback periods. These factors are not explicitly modeled, as the analytical focus is on isolating communication-driven decision dynamics. Holding such constraints constant allows for a clearer interpretation of how mass and personal communication influence the formation of refurbishment intentions. At the same time, this abstraction limits the model's ability to assess real-world feasibility and financial constraints, and should be considered when interpreting the results.

The empirical basis of the model builds on earlier data and conceptual assumptions that predate the recent energy market disruptions triggered by the war in Ukraine. However, the model is designed to capture structural behavioral mechanisms rather than short-term market conditions. Its purpose is therefore not to reproduce a specific market situation, but to explore general patterns of decision-making under different communication regimes. While this limits direct empirical applicability to current market conditions, the conceptual insights into communication effects and decision dynamics remain valid.

The model is partially calibrated to empirical data, as the 17,000 simulated agents reflect both the distribution of Sinus lifestyles and the empirically observed status quo across the stages of the HOE model. However, not all parameters are calibrated in the same way. In particular, the threshold parameter is modeled in a stylized manner and represents a

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

proxy for individual resistance to behavioral change. Similarly, the social contagion mechanism is not empirically calibrated but implemented as a simplified representation of neighborhood-level peer effects. Accordingly, the model combines empirically grounded initialization with exploratory behavioral assumptions. Its value lies less in exact prediction than in examining how different communication regimes interact with heterogeneous starting conditions and decision dynamics. To account for this, a targeted sensitivity analysis was conducted for these non-empirically calibrated parameters, allowing the robustness of the model results to be assessed under systematic ceteris paribus variation.

Second, the decision-making process is modeled using a static adaptation of the HOE framework. While this provides a clear and structured logic, it assumes that all households progress sequentially through awareness, knowledge, attitude, conviction, and intention. In practice, decision-making may be more iterative, with households skipping or revisiting stages, or being influenced by external shocks such as sudden policy changes or energy price fluctuations. These dynamics are not fully captured in the current model.

Third, beyond structural constraints, the model focuses exclusively on communication measures. Other policy instruments, such as financial incentives, regulatory standards, or market mechanisms, are not included. This narrow focus is deliberate, as it isolates the specific role of communication, but it limits the model's capacity to assess the combined effect of policy mixes. Given the importance of subsidies and regulations in real-world renovation programs, this omission must be borne in mind when interpreting the results.

Fourth, the simulation treats stakeholder cooperation qualitatively rather than quantitatively. While Table IV outlined typical cost responsibilities across actors, the model does not calculate exact financial allocations to governments, consultants, or homeowners. This makes the stakeholder dimension illustrative rather than predictive. Future work could strengthen this aspect by coupling ABM results with economic or institutional modeling to better represent the distribution of costs and responsibilities.

In addition, the sensitivity analysis indicates that the role of social contagion within the current model structure is comparatively limited. Variations in contagion strength lead to only minor changes in diffusion outcomes across most scenarios. This may be a result of the specific model design and parameterization, and suggests that indirect social influence is dominated by direct communication effects under the given assumptions.

Finally, the model is designed as an exploratory tool rather than a predictive engine. The results should therefore not be interpreted as forecasts of refurbishment adoption in Germany, but as illustrations of plausible trajectories under different communication mixes. At the same time, the sensitivity analysis shows that model behavior is not uniformly stable, but depends on the overall communication context, which further supports its use as a tool for exploring conditional dynamics rather than generating precise predictions. This limitation is also a strength: it highlights the value of ABM for reasoning under uncertainty, while avoiding the false impression of precision.

D. Future Research Directions

The limitations of the present study also open up several promising directions for future research. These concern methodological refinements, integration with additional policy instruments, and deeper consideration of social factors.

A first avenue lies in broadening the policy scope. The current model isolates communication measures to make their effects explicit. However, real-world renovation programs combine communication with financial incentives, regulatory standards, and technical support. Extending the model to simulate policy mixes would make it possible to analyze synergies and trade-offs between different instruments. For example, combining subsidies with communication could test whether financial incentives amplify the effectiveness of personal advice among households with financial constraints.

A second direction involves enhancing the temporal dynamics of the simulation. In this study, outcomes were evaluated after a fixed period of 27 years or at the 95% threshold of adoption. While informative, this design could be enriched with continuous time series analyses that show how different communication mixes shape trajectories year by year. A new figure comparing diffusion curves across selected scenarios (e.g., MC = 5/PC = 0 vs. MC = 5/PC = 1 vs. MC = 0/PC = 3) would illustrate dynamic differences more vividly. This type of visual enrichment would strengthen the exploratory potential of the model and facilitate policy dialogue.

A third research direction concerns heterogeneity and social diversity. The present model integrates lifestyle segmentation through Sinus milieus, but it treats social factors only indirectly. Empirical studies [19] show that social values strongly influence renovation decisions. Future models could incorporate social or regional variations in responsiveness, willingness to cooperate, thus improving the model's ability to represent diverse contexts. In addition, such extensions could help to better capture indirect social influence mechanisms, which appear only in a limited form in the current model.

A fourth opportunity lies in refining agent behavior. The current HOE-based decision logic assumes sequential progression through stages. Future models could include more adaptive or iterative pathways, allowing agents to skip, revisit, or regress between stages. They could also simulate learning effects, where repeated exposure to communication or peer influence gradually lowers thresholds. This would bring the model closer to real-world decision-making dynamics. It would also allow for a more realistic representation and empirical calibration of behavioral thresholds, which are currently modeled in a stylized manner.

Finally, further work could focus on quantifying stakeholder cooperation. The present study illustrates typical distributions of costs qualitatively. Future research could couple the ABM with economic models or institutional analyses to assign financial burdens more precisely to governments, consultants, and homeowners. Such an approach would support more detailed assessments of cost-sharing fairness and efficiency.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Taken together, these directions point to a rich research agenda. By expanding scope, improving temporal dynamics, integrating social diversity, refining agent behavior, and quantifying stakeholder roles, future work can strengthen the ability of ABM to inform effective and equitable refurbishment policies. In particular, further research could build on the regime-dependent dynamics observed in this study to better understand how different intervention intensities shape diffusion processes under varying conditions.

VI. Conclusion

This paper has presented an exploratory agent-based simulation of communication-driven energetic refurbishment decisions among German homeowners. Building on earlier work, the analysis expanded the model with explicit assumptions, lifestyle segmentation, and a stakeholder perspective. The results demonstrate that communication is a necessary but not sufficient driver of refurbishment acceleration: its effectiveness depends on the interplay of mass communication and personal communication, the balance between speed and cost, and the cooperation of multiple actors in financing and delivering these measures. In addition, the results indicate that these effects depend strongly on the overall level of communication intensity, rather than on individual parameters in isolation.

Three key contributions emerge from this study. First, the model confirms and operationalizes a central insight from marketing theory: mass communication is essential for generating awareness and knowledge but weak in creating conviction and intention, whereas personal communication is decisive in later decision stages. Their combination, rather than either instrument alone, produces the strongest effects.

Second, the results indicate that diffusion dynamics can be characterized by distinct regimes, depending on the overall level of communication intensity. Under low-intensity conditions, adoption is highly sensitive to individual resistance, while under moderate intensity, diffusion outcomes become stable but vary in speed. Under high-intensity conditions, diffusion approaches saturation, and parameter variations have only limited effects. This regime-dependent behavior highlights that the effectiveness of communication strategies cannot be assessed in isolation, but must be understood in relation to the broader communication context.

Third, the paper situates refurbishment acceleration within a governance perspective. By aligning the simulation results with insights from the literature on multilevel governance, municipal engagement, hybrid public–private arrangements, and social influences, the study highlights that effective communication strategies require coordination across actors and contexts. Acceleration is unlikely to be achieved through isolated measures; it depends on cooperation among governments, municipalities, consultants, homeowners, and, in some cases, private housing companies. This perspective is consistent with the regime-dependent dynamics observed in the model, which emphasize the need for coordinated interventions across different actors and intensity levels.

The study also underscores the value of ABM as an exploratory tool. By making assumptions explicit and systematically varying key parameters, including targeted sensitivity analyses, the model provides a structured way to reason about plausible pathways, tensions, and leverage points in refurbishment policy. While not predictive, the approach supports policy learning under uncertainty and offers a transparent platform for dialogue among policymakers, practitioners, and researchers.

Future work should extend the model to include policy mixes beyond communication, explore time series dynamics, incorporate social diversity more explicitly, and quantify stakeholder cost allocations. These extensions would further strengthen the relevance of ABM for the design of effective and equitable renovation strategies.

In conclusion, the findings highlight that communication-driven refurbishment processes are governed not only by the presence of communication measures, but by their interaction within different levels of communication intensity. In particular, the results suggest that the relevance of individual resistance, communication strategies, and social influence varies systematically across different diffusion contexts. While the model is exploratory in nature, it provides a structured perspective on how communication strategies shape adoption dynamics and supports a more differentiated understanding of policy interventions. Rather than offering precise predictions, the study contributes to identifying the conditions under which communication measures are most likely to be effective.

References

[1] M. Verhoog, “How marketing and sales can push and accelerate residential refurbishment,” in Proc. SIMUL 2023, IARIA, 2023, pp. 35–37.

[2] European Commission, “A renovation wave for Europe – greening our buildings, creating jobs, improving lives,” Brussels, 2020.

[3] M. Verhoog, Steuerung von akteuren und entscheidungen in baunetzwerken. Wiesbaden: Springer Gabler, 2018.

[4] H. Du, Q. Han, and B. de Vries, “Modelling energy-efficient renovation adoption and diffusion process for households: a review and a way forward,” Sustainable Cities and Society, vol. 77, Art. 103560, 2022, doi: 10.1016/j.scs.2021.103560.

[5] R. Schüle, A. Bierwirth, and T. Madry, Zukunft der energieberatung in deutschland. Ludwigsburg: Wüstenrot Stiftung, 2011.

[6] E. Dunkelberg and I. Stieß, Energieberatung für eigenheimbesitzer/innen. Berlin: IÖW, 2011.

[7] S. Railsback and V. Grimm, Agent-based and individual-based modeling: a practical introduction. Princeton: Princeton Univ. Press, 2012.

[8] A. Baranzini, S. Carattini, and M. Péclat, “What drives social contagion in the adoption of solar photovoltaic technology?,” GRI Working Paper 270, 2017.

[9] V. Rai and S. A. Robinson, “Agent-based modeling of energy technology adoption: empirical integration of social, behavioral, economic, and environmental factors,” Environ. Model. Softw., vol. 70, pp. 163–177, 2015, doi: 10.1016/j.envsoft.2015.04.014.

[10] E. Derkenbaeva, G. J. Hofstede, and E. van Leeuwen, “Simulating households’ energy transition in Amsterdam: an agent-based modeling approach,” Energy Res. Soc. Sci., vol. 99, p. 103087, 2023, doi: 10.1016/j.erss.2023.103087.

[11] T. Barry, “The development of the hierarchy of effects: an historical perspective,” Curr. Issues Res. Advertising, vol. 10, no. 2, pp. 251–295, 1987.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

[12] R. Lavidge and G. Steiner, "A model for predictive measurements of advertising effectiveness," J. Marketing, vol. 25, pp. 59–62, 1961.

[13] C. Maby, E. Koirala, J. Parag, and B. K. Sovacool, "Energy advice in Europe 2007: a review of current practice in advice on sustainable energy in the countries of the European Union," European Commission, 2007.

[14] W. Jager, "The four P's in social simulation: a perspective on how marketing could benefit from the use of social simulation," J. Business Res., vol. 60, no. 8, pp. 868–875, 2007, doi: 10.1016/j.jbusres.2007.02.003.

[15] R. Galvin, "An under-developed dimension in upgrading energy-inefficient German rental buildings: corporate social responsibility as a hybrid form of governance," Energy Res. Soc. Sci., 2023.

[16] P. Bertoldi, M. Economidou, and V. Palermo, "How to finance energy renovation of residential buildings: review of current and emerging financing instruments in the EU," WIREs Energy Environ., 2020.

[17] H. Nykamp, "Policy mix for a transition to sustainability: green buildings in Norway," Sustainability, vol. 12, no. 2, 2020.

[18] S. Ebrahimigharehbaghi, Q. K. Qian, G. de Vries, and H. J. Visscher, "Municipal governance and energy retrofitting of owner-occupied homes in the Netherlands," Energy & Buildings, vol. 274, 2022, doi: 10.1016/j.enbuild.2022.112423.

[19] M. Jowkar, A. Temeljotov-Salaj, C. M. Lindkvist, and M. Støre-Valen, "Sustainable building renovation in residential buildings: barriers and potential motivations in Norwegian culture," Construction Management and Economics, vol. 40, no. 3, pp. 161–172, 2022, doi: 10.1080/01446193.2022.2027485.

[20] M. van der Kam, M. Lagomarsino, and E. Azar, "An empirical agent-based model of consumer co-adoption of low-carbon technologies to inform energy policy," Cell Reports Sustainability, vol. 1, no. 12, 2024, doi: 10.1016/j.crsus.2024.100268.

[21] F. Chen, R. Fan, and Q. Yao, "Evolutionary dynamics of homeowners' energy-efficiency retrofit decision-making in complex networks," J. Environ. Manag., 2023, doi: 10.1016/j.jenvman.2022.116849.

[22] Y. Shen et al., "Integrated framework for space- and energy-efficient retrofitting in multifunctional buildings: a synergy of agent-based modeling and performance-based modeling," Build. Simul., vol. 17, pp. 1579–1600, 2024, doi: 10.1007/s12273-024-1148-z.

[23] S. Tian et al., "An agent-based modeling approach combined with deep learning method in simulating household energy consumption," J. Building Eng., 2021, doi: 10.1016/j.jobe.2021.103210.

[24] M. A. Khalil and M. R. Fatmi, "How residential energy consumption has changed due to COVID-19 pandemic? an agent-based model," Sustainable Cities and Society, vol. 81, 2022, doi: 10.1016/j.scs.2022.103832.

[25] P. Kotler and K. Keller, Marketing management, 13th ed. Pearson, 2009.

[26] P. De Pelsmacker, M. Geuens, and J. Van den Bergh, Marketing communications, 6th ed. Pearson, 2017.

[27] S. A. Delre, W. Jager, T. H. A. Bijmolt, and M. A. Janssen, "Targeting and timing promotional activities: an agent-based model for the takeoff of new products," J. Business Res., vol. 60, no. 8, pp. 826–835, 2007, doi: 10.1016/j.jbusres.2007.02.002.

[28] C. S. Chen, C. C. Yu, and K. Y. Tu, "Exploring the impact of integrated marketing communication tools on green product purchase intentions among diverse green consumer segments," Sustainability, vol. 15, no. 24, 2023, doi: 10.3390/su152416841.

[29] E. Correia, S. Sousa, C. Viseu, and M. Larguinho, "Analysing the influence of green marketing communication in consumers' green purchase behaviour," Int. J. Environ. Res. Public Health, vol. 20, no. 2, 2023, doi: 10.3390/ijerph20021356.

[30] S. Krsnik and K. Erjavec, "Comprehensive study on the determinants of green behaviour of Slovenian consumers: the role of marketing communication, lifestyle, psychological, and social determinants," Sustainability, vol. 16, no. 17, 2024, doi: 10.3390/su16177555.

[31] E. Iliopoulou, E. Koronaki, A. Vlachvei, and O. Notta, "The power of green communication and social media engagement in sustainable food consumption," Sustainability, vol. 16, 2024, doi: 10.3390/su16219202.

[32] M. Björklund, F. von Malmborg, and J. Nordensvärd, "Lessons learnt from 20+ years of research on multilevel governance of energy-efficient buildings in the EU," Energy Efficiency, 2023, doi: 10.1007/s12053-023-10178-6.

[33] M. Verhoog, "A target group-specific communication approach to push refurbishment," in Energy for Sustainability Int. Conf., Funchal, 2017.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Extending Process Virtualization Theory for Digital Collaboration:

Social Presence, Situation Awareness, and Leadership in Virtual and AI-Augmented Teams

Inga F. Schlömer
IT and Engineering Department
IU International University
Erfurt, Germany
e-mail: inga.schloemer@iu.org

Abstract—**Virtual collaboration has become a central mode of organizing knowledge work, yet sustaining social and informational dynamics in digitally mediated teams remains a persistent challenge. Building on an earlier conference publication presented at the DIGITAL 2025 Conference on Advances on Societal Digital Transformation, this article extends a systematic literature review on digital collaboration tools by broadening its theoretical foundation, integrating leadership as a contextual factor, and incorporating a complementary exploratory review on artificial intelligence (AI) in virtual collaboration. Anchored in Process Virtualization Theory, and enriched by insights from Media Richness Theory, Media Synchronicity Theory, and Social Exchange Theory, the study examines how digital collaboration tools contribute to social presence and situation awareness in virtual teams. Using qualitative content analysis, the review synthesizes findings from peer-reviewed literature to identify recurring patterns in how collaboration technologies support interpersonal connection, shared understanding, and trust in remote work settings. The results show that social presence is primarily fostered through media-rich and synchronous interaction formats, informal communication opportunities, and practices that strengthen belonging, while situation awareness is supported by persistent information structures, transparent communication routines, and ongoing knowledge exchange. Trust emerges as a cross-cutting condition that links these dimensions and mediates the effectiveness of digital collaboration. Leadership practices play a pivotal role in shaping how technological affordances are used, influencing whether tools enhance cohesion and awareness or contribute to overload and control perceptions. The complementary AI-focused review highlights an emerging ambivalence: while AI-enabled functionalities may enhance efficiency, transparency, and coordination, they also pose risks to trust and social presence when algorithmic processes become opaque or intrusive. By synthesizing these insights, the article advances a socio-technical and socio-algorithmic extension of Process Virtualization Theory and offers implications for research, leadership practice, and the design and governance of digital collaboration tools.**

Keywords—digital transformation; virtual collaboration; process virtualization theory; collaboration tools.

I. INTRODUCTION

This paper builds upon an earlier study presented at the DIGITAL 2025 Conference on Advances on Societal Digital Transformation [1], in which a systematic literature review examined how digital collaboration tools enhance social presence and situation awareness in virtual collaboration. The present article extends this analysis by broadening the theoretical foundation through the inclusion of Media Richness Theory (MRT), Media Synchronicity Theory (MST), and Social Exchange Theory (SET), thereby allowing for a more differentiated understanding of collaboration processes in virtual settings. In addition, leadership is being analyzed as a critical moderating component in this context. Furthermore, a complementary literature review is conducted to explicitly examine the role of AI in digital tools used for virtual collaboration. This additional review systematically synthesizes findings from the broader body of academic research on AI-enabled functionalities, such as automated facilitation, adaptive feedback mechanisms, and intelligent support, to evaluate how the integration of AI shapes interaction quality, efficiency, and relational dynamics within virtual teams.

Virtual collaboration has become an integral part of organizational operations, driven by advancements in digital collaboration technologies and further accelerated by the COVID-19 pandemic. Virtual collaboration is commonly used as an umbrella term for digitally mediated work practices, although it often encompasses analytically distinct processes of coordination, communication, and collaboration. In a narrower sense, collaboration refers to the joint production of outcomes based on shared goals, while coordination concerns the alignment of interdependent activities and communication the exchange and interpretation of information [2].

Virtual teams have evolved from a niche practice to a widespread organizational norm [3]. The transition from traditional office environments to remote and hybrid models has necessitated the virtualization of work processes, as defined by the Process Virtualization Theory (PVT) [4]. PVT assesses how physical processes translate to virtual settings.

The pandemic-induced shift to remote work challenged existing assumptions about the limitations of virtual processes, demonstrating that even knowledge work processes with high relational demands could be effectively virtualized [5]. This evolution from a temporary adjustment

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

to a new post-pandemic digital era has opened opportunities to re-evaluate how digital tools can support not only operational efficiency but also the social and relational aspects of teamwork. Central to this discussion are the Information Technology (IT) characteristics of social presence and situation awareness, which contribute to creating a sense of interpersonal connection and maintaining contextual understanding within remote teams.

Through an extended systematic literature review, this paper advances current understanding of how digital collaboration tools shape social presence and situation awareness in virtual teams and how these effects are influenced by leadership practices and emerging technological capabilities. Building on insights from MRT, MST, and SET, the analysis integrates complementary theoretical lenses to explain not only how communication technologies enable interaction, but also why certain configurations foster trust, cohesion, and effective coordination in remote work settings. Furthermore, by incorporating an additional review of AI-augmented collaboration tools, the article examines how intelligent functionalities increasingly mediate virtual teams and influence social presence and situation awareness. In doing so, this research contributes to digital transformation scholarship by offering a more nuanced, socio-technical perspective on virtual collaboration that accounts for leadership, theory-driven interpretation, and the growing role of AI in shaping contemporary remote work practices.

Following this introduction, Section II establishes the theoretical foundation of the study by outlining PVT and synthesizing prior research on social presence and situation awareness, complemented by additional perspectives from MRT, MST, and SET. Section III details the methodological approach, including the systematic review protocol, data analysis procedures, and the extension of the original review conducted for the earlier conference publication, as well as limitations. Section IV presents the findings of the qualitative content analysis, structured around social presence, situation awareness, trust, and leadership as key analytical categories. To further contextualize these findings, Section V illustrates concrete examples of collaboration tool use drawn from the reviewed literature. Section VI reports the results of the complementary exploratory review on AI-enabled collaboration tools and discusses their opportunities and limitations within virtual teamwork. Section VII synthesizes the conceptual contributions and implications of the study, extending PVT and outlining practical and research-related implications. Finally, Section VIII concludes the article by summarizing the extended perspective and situating the findings within the broader discourse on digital and AI-mediated collaboration.

II. Theoretical Foundation of Virtual Collaboration

The increasing prevalence of remote and hybrid work models has fundamentally reshaped contemporary office settings, leading to a growing reliance on virtual modes of collaboration [3]. This shift is particularly pronounced in the domain of knowledge work, which is characterized by activities centered on the creation, dissemination, and application of knowledge rather than on manual or physical labor. Within this context, virtual teams have become an important and widely established organizational form. They are commonly understood as groups of individuals working toward a shared objective while being geographically dispersed and depending primarily on digital technologies to structure their collaboration and bridge spatial, temporal, and organizational divides [6]. The COVID-19 pandemic, coupled with mandatory remote work policies, further accelerated the adoption of both remote and hybrid team structures [7]. As a result, most organizations today exhibit some degree of virtuality, which increasingly blurs the boundaries between virtual and non-virtual teams and highlights the need for frameworks capable of explaining this transformation [8]. In this regard, PVT offers a valuable perspective by providing a conceptual approach to assessing, which organizational processes can be effectively transferred to virtual environments [4]. The next section will explore this theory in more detail. It will outline the theory's core concepts, review its empirical applications, and discuss its current extensions. Building on this foundation, the subsequent section introduces additional theoretical lenses. MRT, MST, and SET refine the analysis of technology-mediated collaboration and illuminate complementary mechanisms that PVT alone does not fully capture.

A. Process Virtualization Theory

The virtualization of work processes is a fundamental prerequisite for enabling remote collaboration. PVT provides a theoretical framework for analyzing the digital transformation of business processes by examining the extent to which physical processes can be effectively replicated in virtual environments [4]. At the core of PVT lies the concept of process virtualizability, which determines how well a process can be executed without reliance on physical interactions between individuals or individuals and objects. Process virtualizability is specified as the dependent construct and is commonly operationalized either by the quality of outcomes produced by the virtualized process or by indicators of its actual usage. This concept plays a crucial role in guiding organizations through digital transformation by identifying, which processes are most suited for virtualization [9]. Four constructs determine whether a process can be effectively virtualized: *sensory*, *relationship*, *synchronization*, as well as *identification and control requirements* [4], [9]. These factors negatively affect virtualizability, meaning that as their importance increases, the feasibility of virtualization decreases [4].

Sensory requirements denote the extent to which a process demands the reproduction of rich, embodied experience for its participants, encompassing not only the classic modalities of sight, sound, touch, taste, and smell, but also the felt, affective qualities that accompany participation, such as excitement or vulnerability [4]. When successful execution requires direct sensory access to other people or to physical artifacts, the process becomes dependent on a material context that current virtual environments typically cannot fully replicate. Because virtualization removes or

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

weakens these forms of physical co-presence and object interaction, high sensory requirements systematically constrain the feasibility of migrating such processes into virtual settings.

Relationship requirements refer to the extent to which a process relies on sustained interpersonal engagement in social or professional contexts, through which participants exchange knowledge, build trust, and sometimes form friendships [4]. Conceptually, these requirements are closely aligned with sensory requirements, as both concern facets of embodied, interpersonal experience, that are difficult to reproduce in digital settings [10]. Empirical work shows that meaningful relationships can develop via virtual media when participants have sufficient time to interact and gain familiarity with the tools [4]. However, high relationship requirements typically necessitate compensatory arrangements, such as occasional in-person encounters, structured opportunities for rapport-building, or the deliberate cultivation of a shared identity, which raise the effort needed to virtualize the process and increase resistance to virtualization.

Synchronization requirements capture the extent to which a process depends on tightly coordinated timing, whether in real time or in carefully ordered sequences, so that individual activities must occur with minimal delay [4]. Physical settings typically satisfy these requirements by default, since co-located participants can interact with one another and with relevant artifacts almost instantaneously, provided the number of actors and objects remains manageable. By contrast, virtual environments do not inherently guarantee low-latency coordination. Achieving near–real-time interaction often entails additional technological and organizational effort. Notably, asynchronous arrangements can also be advantageous, enabling participants to contribute at personally convenient times and to engage in more deliberate, reflective responses. Overall, while synchrony tends to be "built in" to collocated processes, it must be actively engineered in virtual ones, making high synchronization requirements a potential obstacle to virtualization.

Identification and control requirements describe the extent to which a process depends on reliably establishing the unique identity of participants and on exercising effective oversight or influence over their behavior [4]. Such requirements are generally associated with lower process virtualizability, because virtual settings heighten risks of identity fraud and reduce the transparency needed to attribute actions and enforce norms. When the actors cannot be physically verified and their conduct is less observable, detecting who performs which activity and shaping their behavior becomes more difficult, thereby increasing the effort required to virtualize the process. At the same time, anonymity can be beneficial in processes that do not depend on knowing the identity of others. For instance, anonymous group brainstorming may stimulate creativity and reduce excessive reluctance to speak up in front of high-ranking participants.

Process virtualization can be achieved with or without digital technologies [4]. Any arrangement that enables execution without immediate physical co-presence, whether technological or not, constitutes virtualization. Books are historical examples that demonstrate virtualization predates contemporary computing by enabling information exchange without face-to-face interaction. What distinguishes the present era is the dramatic expansion of virtual processes driven by increasingly capable and widely accessible IT, particularly the internet. To clarify how technology shapes this development, the core PVT model specifies three IT characteristics, that positively moderate the relationship between the theory's main constructs and process virtualizability: representation, reach, and monitoring capability.

Representation refers to IT's ability to present process-relevant information in a way that makes actors, objects, their properties, and their interactions understandable to those involved [4]. IT enables multimodal portrayals through audio and video and, increasingly, through haptic and olfactory interfaces, which help to meet higher sensory requirements. Sight and sound are well established, but smell and touch remain more challenging. However, they continue to advance. Non-IT media also provide representational functions. For example, a physical team Kanban board depicts tasks, status, and interdependencies for co-located collaborators. What differentiates IT-based representation is its dynamism: digital renderings can be updated in real time as conditions change, whereas non-IT representations are typically static. In addition, IT-enabled profiling and matching can support relationship-intensive processes by aligning participants not only on complementary interests but also on similar competencies or overlapping functional domains in organizational settings.

Reach refers to IT's capacity to decouple process participation from shared time and place, enabling engagement across both temporal and geographic boundaries [4]. Temporally, reach allows team processes to continue outside of conventional working hours. For example, distributed development teams use issue trackers, code repositories, and document comment threads to progress work, so that one site can hand off at the end of the day and another can pick up later. Spatially, reach lets dispersed colleagues participate in the same workflow by co-authoring in real time, conducting virtual stand-ups, or pairing remotely on code and models via shared sessions. By connecting individuals who might not otherwise interact and helping them discover peers with related competencies or overlapping domains through enterprise profiles and communities of practice, reach makes high relationship requirements easier to satisfy. It also meets the demand for highly synchronized collaboration by enabling live, two-way collaboration in incident "war rooms," shared whiteboards, and remote pair-programming environments where participants can respond and seek clarification in real time. In tandem with representation, which provides richer interaction contexts, reach helps teams find one another and sustain timely, effective engagement, thereby strengthening process virtualizability.

Monitoring capability refers to an IT system's ability to authenticate participants and record, attribute, and analyze

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

their actions. This eases processes with strict identification and control requirements [4]. In IT-mediated settings, users typically authenticate via credentials or comparable methods, enabling systematic, fine-grained, and automated tracking of participation and behavior. Although risks such as identity spoofing persist, advances in identification technologies, for example biometrics, help to mitigate these vulnerabilities. In some cases, the oversight resulting from this can exceed what is feasible in face-to-face contexts, providing a greater capacity for identification and control in practice. Non-IT arrangements also support monitoring, for example, punch clocks and timecards provide basic attendance records, but the automated and analytic affordances of IT offer a qualitatively different level of observability.

B. Extensions to PVT

PVT has been applied and extended across a growing range of contexts, underscoring both its explanatory reach and its boundary conditions. In higher education, researchers show that virtualization requirements embedded in the learning process shape undergraduates' satisfaction and their continuance intentions in online settings [11]. In consumer finance, research on online banking shows that the core PVT requirements serve as prerequisites for attitudes toward virtualizability, user resistance, and actual use [12]. Public-sector evidence indicates that perceived virtualizability predicts intentions to adopt government e-services and is intertwined with resistance dynamics among users of an e-pay slip system [13]. Beyond these domains, PVT has been applied to analyze accommodation services [14]. Used in fintech research, it shows that rising synchronism and identification/control readiness can outweigh declining relationship readiness and boost customers' intention to use [15]. Studies on remote work emphasize sensory and synchronization requirements as significant obstacles to effective virtualization [16]. Others expand PVT by incorporating e-commerce-specific attributes [17] and participant-related factors [18].

The COVID-19 pandemic has provided a unique context for re-evaluating PVT through a crisis-driven perspective [5]. The global health crisis necessitated the rapid virtualization of knowledge work processes, many of which had previously been resistant to digital transformation. As a result of this accelerated transformation, virtualizability evolved from a dependent variable to a determining factor in the fulfillment of process requirements. Organizations increasingly relied on IT characteristics to address these virtualization challenges, leveraging digital tools to simulate sensory elements, facilitate relationship-building, support synchronous interactions, and enhance monitoring capabilities. In this revised understanding of PVT, two additional IT characteristics — social presence and situation awareness — were identified, which further contribute to the effective virtualization of knowledge work processes. Social presence refers to the ability of IT to create a sense of interpersonal connection in virtual settings. The Social Presence Theory states that the degree of perceived psychological and emotional connection between individuals in a mediated communication environment influences the effectiveness and quality of interactions [19], [20]. Social presence refers to the psychological experience in technology-mediated interaction whereby communication partners are sensed as tangible, "real" others, despite the exchange occurring through computer-mediated tools and electronic platforms [21]. It emphasizes how communication technologies vary in their ability to convey social cues in virtual settings.

Situation awareness — essentially "knowing what is going on" [22] — centers on perceiving relevant context information, interpreting its meaning, and anticipating its future development. When shared among team members, this becomes team situation awareness [5], a collective view of conditions, implications, and likely trajectories that is vital for the relational demands of remote work.

The digital age has challenged the notion that processes can only be virtualized to a limited extent, as the empirical reality of widespread remote work has shown that IT can indeed support the social aspects of collaboration in virtual teams. Although some critics have pointed out that many recommendations overlooked the involuntary nature of pandemic-related remote work [23], the urgency to virtualize has receded in today's post-pandemic digital era. This transition creates an opportunity to reflect on progress made in remote collaboration and to reevaluate the potential of collaboration tools through the lens of social presence and situation awareness. Both of these IT characteristics present innovative pathways for understanding how virtual environments can foster not only operational efficiency but also the interpersonal and contextual dimensions of teamwork.

C. Complementary Theoretical Perspectives on Digital Collaboration

To enhance the explanatory power of PVT and deepen the understanding of how collaboration tools influence social presence and situation awareness, this study also draws upon three complementary theoretical perspectives: MRT, MST, and SET. These frameworks provide valuable insights into how digital collaboration tools shape relational dynamics, cognitive alignment, and trust development in virtual teams.

MRT posits that communication media vary in their capacity to effectively convey information, especially in ambiguous or unclear situations [24]. Richer media, such as face-to-face interaction or video conferencing, enable immediate feedback, the transmission of multiple cues (e.g., gestures, tone of voice), and personal focus, thereby supporting the exchange of nuanced and emotionally rich messages. Leaner media, for example email or text messaging, have limited bandwidth for relational and contextual cues and are therefore less suitable for interactions requiring emotional engagement or complex interpersonal coordination.

In the context of virtual collaboration, MRT thus provides a theoretical foundation for understanding why certain technologies are more effective in fostering social presence and affective trust. At the same time, the theory also informs decisions about how task-related information should be communicated in order to support situation awareness. Richer media are particularly appropriate when

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

shared understanding of evolving tasks, interdependencies, or unexpected changes are required, whereas leaner media may suffice for the dissemination of routine updates or clearly structured information. From this perspective, media richness influences not only relational outcomes but also the extent to which team members can develop and maintain a coherent understanding of the current work situation.

Compared to PVT, which evaluates the conditions under which entire organizational processes, including both person-to-person and person-to-object interactions, can be effectively migrated into virtual settings, MRT addresses a more focused question: how individuals select communication media for specific interpersonal tasks. Its unit of analysis is the medium's capacity to convey cues relative to task ambiguity. MRT explains media choice in bilateral communication, PVT captures the broader feasibility of virtualizing diverse process types beyond communication alone [4].

MST further refines this perspective by distinguishing between two fundamental communication processes: conveyance, referring to the transmission of new or factual information, and convergence, which involves the development of shared understanding among communication partners [25]. According to MST, synchronous communication tools, such as chat or video conferencing platforms, are particularly well suited for convergence processes, as they enable rapid feedback, mutual adjustment, and the co-construction of meaning. These characteristics support not only the alignment of interpretations and expectations, but also the experience of social presence through ongoing interpersonal engagement.

Asynchronous tools, by contrast, primarily facilitate conveyance by allowing time for reflection, parallel information processing, and the structured dissemination of task-related content. From this perspective, MST helps explain how different temporal configurations of communication media contribute to situation awareness: synchronous interaction supports real-time awareness of team dynamics and emerging issues, while asynchronous communication sustains a shared understanding of tasks, responsibilities, and progress over time. In this way, MST complements PVT by specifying how the temporal organization of communication within virtual processes shapes both social presence and situation awareness in digitally mediated collaboration.

In addition to media theories, SET offers a relational lens to understand how digital interactions cultivate trust and commitment within remote teams. SET argues that social behavior is driven by reciprocal exchanges that generate mutual obligations and affective bonds over time [26]. In virtual environments, where informal interactions and spontaneous reciprocity are constrained, the strategic use of digital tools becomes essential for initiating and maintaining such social exchanges. For instance, routine social check-ins or informal chats can simulate reciprocal behavior and foster perceptions of goodwill and reliability, critical foundations for trust [27]. SET thereby informs how collaboration tools can be intentionally used to reinforce relational stability and long-term cooperation despite geographic and temporal dispersion.

Together, these theoretical lenses enrich the interpretation of how collaboration tools influence virtual team processes and offer a broader conceptual grounding for analyzing the findings of this study.

III. Research Design and Methodological Approach

This paper builds on a systematic review of the existing body of literature to gain deeper insights and a more comprehensive understanding of the contribution of digital collaboration tools to the social aspects of virtual collaboration. By analyzing prior research, the study seeks to clarify how digital collaboration tools influence key social mechanisms in virtual teams. Specifically, this systematic literature review analyses how digital collaboration tools shape social presence and situation awareness in distributed work settings. Against this background, the review is guided by the following research questions:

RQ1: In what ways do digital collaboration tools foster social presence in remote settings?

RQ2: In what ways do digital collaboration tools enhance situation awareness in remote teams?

A. Systematic Literature Review Design and Search Procedure

The review employed a structured approach to address the research questions [28]. The review scope was defined as a summary of previous research activities and their underlying theories on remote work, collaboration tools, and virtual team collaboration [29]. The research covered several relevant databases (ACM, EBSCO, Emerald, IEEE, Science Direct, Springer). The systematic search combined the keywords "collaboration tool," "new work," "remote work," "hybrid work," "team cohesion," "social bonding," "group cohesion," "sense of community," and "virtual team dynamics" with Boolean AND/OR operators to ensure comprehensive, relevant results. Figure 1 summarizes the four-iteration review process and resulting article counts as proposed by [30].

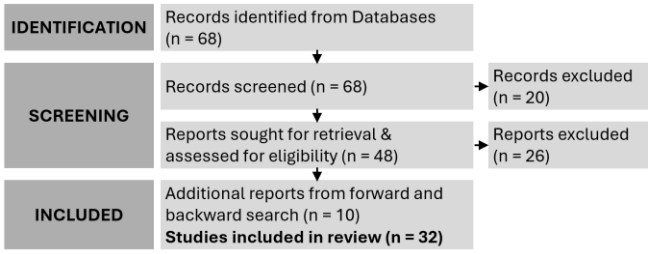


Figure 1. First Review Process.

A database search using the predefined terms was followed by abstract screening based on the criteria in Table I, full-text review using the same exclusions, and forward and backward citation tracking, adding ten articles and yielding 32 eligible studies.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

TABLE I. EXCLUSION CRITERIA.

Exclusion Criteria	Published before 2010. No use of collaboration tools. Focus on teams without a degree of virtuality. Articles in any other language as German and English

B. Qualitative Content Analysis and Coding Scheme

The collected data was analyzed using qualitative content analysis [31]. Recurring themes and patterns were identified that shed light on key factors in the use and effectiveness of collaboration tools for supporting social presence and situation awareness.

Social presence is understood as the degree to which digital collaboration tools facilitate awareness of the other person and interpersonal relationships during interaction [19]. Accordingly, the category of *interpersonal relationship/team cohesion* was included, complemented by the concepts of *collaboration* and *community/belonging*. These categories capture the extent to which digital platforms can replicate or substitute the relational dynamics of face-to-face settings, which are essential for long-term engagement and collective performance in virtual teams.

Situation awareness, in contrast, highlights the informational dimension of virtual collaboration. It refers to the collective understanding of work dynamics within a team, encompassing awareness of colleagues' tasks, progress, and interdependencies. In remote settings, this awareness cannot emerge spontaneously, as in co-located offices, but must be actively supported through deliberate communication practices and digital tools. For this reason, *knowledge exchange* was defined as a key category. Closely related is *informal communication*, which plays a vital role in filling contextual gaps, enabling quick clarifications, and creating shared mental models of ongoing work. Without such informal channels, situation awareness can easily break down, resulting in delays, misunderstandings, and weakened trust.

Finally, the analysis revealed *trust* as a cross-cutting concept that underpins both social presence and situation awareness. Trust determines whether employees are willing to disclose problems, share incomplete information, and rely on one another in the absence of direct oversight. It also mediates the effectiveness of digital communication, as technological affordances alone cannot ensure openness or psychological safety. Consequently, trust was integrated into the framework as an essential condition for virtual collaboration.

In addition to the analytical categories originally derived from the systematic review reported in the earlier conference publication [1], the review process was extended in the context of developing the present journal article. This extended review introduced *leadership* as an additional analytical category. The extension was motivated by recurring evidence across the literature indicating that leadership practices play a decisive role in shaping social presence and situation awareness in virtual teams. Accordingly, relevant text segments were newly coded, enabling the analysis to capture how leadership-related behaviors, such as structuring communication processes, providing feedback, or demonstrating empathy, interact with digital collaboration tools to influence the social and informational dynamics of remote collaboration. The appendix presents the resulting concept matrix, which presents the analysis of the literature, additionally classified by research design and tool type.

C. Exploratory Review of AI-Enabled Collaboration Tools

To broaden the conceptual scope of this study and critically reflect on emerging trends in digital collaboration, an additional exploratory literature analysis was conducted on the role of AI in virtual teamwork. While the original systematic review focused on collaboration tools in general, this supplementary analysis aimed to examine how AI-enabled functionalities, such as meeting summarizers, smart assistants, emotion detection, and intelligent task coordination, may influence social presence and situation awareness.

The rationale for this extension stems from the increasing integration of AI into mainstream collaboration platforms and the potential shift in how interpersonal and contextual dynamics are mediated. As AI technologies increasingly shape how information is exchanged and interpreted in virtual teams, they raise new questions about the boundaries and assumptions of PVT.

To identify relevant literature, a structured keyword-based search was conducted in July 2025 across the following academic databases: ACM Digital Library, EBSCO, IEEE, Science Direct and Springer. The search combined the terms "artificial intelligence" AND "collaboration tools" AND "remote work". Only peer-reviewed journal articles and conference proceedings published in English were considered. Figure 2 illustrates the review process in accordance with the PRISMA protocol [30].

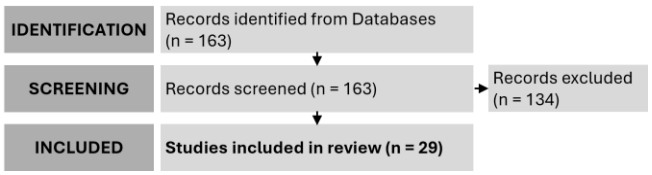


Figure 2. Second Review Process.

The search yielded 163 documents. After screening them based on their relevance to the research scope, 29 were included. These sources were analyzed using qualitative content analysis, focusing on AI functions, positive as well as negative effects of AI use in virtual collaboration and future uses of AI [31].

This supplemental dataset does not feed directly into the original concept matrix, as its thematic focus differs from the initial review sample. The results inform a separate critical

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

discussion on the opportunities and limitations of AI-augmented collaboration within the framework of process virtualization (see Section VI).

D. Methodological Limitations

Although the reviews followed a structured protocol across multiple databases, the reliance on a single researcher remains a significant limitation. Potential selection bias might have affected the breadth and inclusivity of identified studies, as different reviewers could have interpreted inclusion and exclusion criteria differently or identified additional relevant publications. Moreover, focusing on specific search terms may have overlooked research addressing remote collaboration from adjacent perspectives (e.g., organizational psychology, workplace design). Together, these factors could limit the generalizability of the findings and should be considered when interpreting the results.

IV. Findings: Social and Informational Dimensions of Virtual Collaboration

This section presents the results of the qualitative content analysis and synthesizes recurring patterns identified across the reviewed literature. The findings illustrate how digital collaboration tools are discussed in relation to social presence, situation awareness, and trust in remote team settings, while also highlighting leadership as a key factor shaping these dynamics. The following subsections report the findings in accordance with the analytical categories and summarize dominant themes identified across the analyzed studies.

A. Social Presence in Virtual Teams

Social presence is a critical factor in virtual collaboration. Many studies emphasize that digital platforms can be designed to foster the feeling of togetherness typically experienced in face-to-face settings. Synchronous communication tools, such as video conferencing and instant messaging, have been found to be particularly effective at cultivating social presence by enabling immediate, reciprocal interaction and conveying rich social cues (e.g., facial expressions and gestures), all of which are vital for establishing and maintaining interpersonal bonds [32], [33].

According to MST, convergence-oriented tasks such as relationship-building benefit from high synchronicity media, which facilitate real-time feedback and emotional resonance [25]. Immersive digital environments, such as virtual workspaces and avatars, can further strengthen this sense of connection by simulating physical co-presence, thereby enhancing engagement and team cohesion [34], [35]. For instance, the platform Gather lets remote colleagues initiate spontaneous interactions by virtually approaching one another [34] (further elaborated in Section V). This ability for rapid, informal encounters can mitigate the isolation that often arises in remote settings [32], [36], [37].

These findings align with MRT, which suggests that media capable of transmitting multiple cues and supporting immediate feedback are more suitable for conveying relational and affective content than leaner, text-based alternatives [24].

One study draws on the Embodied Social Presence Theory (ESPT), which suggests that virtual workspaces should replicate significant aspects of physical interaction to enhance communication and collaboration. This enables participants to interpret nonverbal cues in digital environments [35]. Through the use of synchronous video tools and virtual reality platforms, ESPT's focus on embodied co-presence reduces ambiguity in remote interactions and reinforces interpersonal relationships.

Some authors emphasize that organizational culture and shared values can be reinforced through collaboration tools when they are used for team-building events, recognition programs, and personal updates [38]. Similarly, planned informal interactions can improve social presence by compensating for the lack of casual conversations in the corridor [39]. Teams that allocate time for social check-ins tend to feel more aligned and engaged. Studies further highlight the importance of intentional, well-structured digital communication to replicate the camaraderie of co-located offices [40], [41]. Frequent touchpoints, such as virtual coffee chats and informal communication channels, help remote employees preserve a communal sense of identity [36], [39]. Meanwhile, personalized introductions — including Asynchronous Semi-Guided Professional Introductions (ASGPIs) — facilitate the exchange of personal and professional information, thereby strengthening interpersonal connections [32].

It is also recognized that even asynchronous tools can support social presence by enabling indirect reciprocity and sustained connections over time, yet their overall impact is constrained by the absence of immediate interaction and emotional engagement [35]. According to MST, these asynchronous channels are better suited for conveyance tasks and fall short in facilitating convergence processes that rely on shared understanding and mutual responsiveness [25].

However, studies indicate that an overreliance on asynchronous communication can lead to feelings of social isolation and a reduced sense of belonging, particularly in teams that lack prior face-to-face interactions [38]. The literature stresses that choosing the right mix of synchronous and asynchronous channels is vital for sustaining high-quality interactions [37], [42]. Platforms, such as Slack, Microsoft Teams, Gather, or Zoom, allow for both project-focused conversations and casual "break room" chats [43]. This multimodal approach ensures that teams can adapt their communication methods to both the complexity and urgency of tasks [33]. Well-timed, media-rich interactions, such as video calls at the beginning of a team's lifecycle, foster personal relationships. Text-based channels and shared documents, on the other hand, are sufficient for routine status updates as collaboration progresses [42], [44]. This provides a direct link and a transition to situation awareness.

B. Situation Awareness in Distributed Collaboration

While social presence fosters emotional connection, situation awareness centers on the ability to stay informed about colleagues' work progress, needs, and challenges in

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

real time, essential for coordinating team efforts and ensuring alignment. While traditional offices often rely on spontaneous observations and overheard conversations [45], remote work environments lack these spontaneous exchanges and must adopt purposeful strategies to maintain comparable levels of context [46]. The inability to observe colleagues' work activities or engage in impromptu discussions often leads to knowledge silos and reduced collaboration efficiency [45]. In addition, researchers found that virtual team members often hesitate to seek help because problems are more difficult to explain and responses are slower [47]. To counteract these effects, many teams have adapted their communication practices by integrating asynchronous chat tools, such as Slack, which facilitate real-time knowledge exchange despite physical separation [39] (further elaborated in Section V). Tools, such as email, shared documents, and discussion forums, contribute to collaboration by supporting structured information exchange and enabling parallel collaboration [33], [48]. Also from the perspective of MST, asynchronous channels, like shared documents or message boards, may be preferable for information distribution (conveyance) [25].

It is underlined that knowledge sharing keeps remote teams aligned and efficient [49]. Similarly, researchers argue that social support, which includes knowledge-sharing practices, strengthens connectivity and reduces workplace stress, fostering a more cohesive understanding of team dynamics [37]. Furthermore, a study indicates that mutual exchange of knowledge builds trust and reinforces collective competence [48]. The authors highlight the significance of informal knowledge exchange in virtual environments, where psychological safety plays a crucial role in facilitating open discussions and information flow. For example, short, informal encounters ("huddle interactions") or informal side chats during official meetings allow members to ask quick questions, seek clarifications, or exchange immediate feedback, much as they might in the physical workspace [39], [45]. Additionally, these ongoing knowledge exchanges, whether over group messaging or video calls, counteract the risk of isolation and misalignment [39], [44].

Well-structured check-ins and routine status updates via shared platforms help team members anticipate when others might need assistance [42]. This transparency not only aids productivity but can also reduce stress and foster mutual trust, as teams learn to predict and fill potential gaps [40], [42], [50]. This more flexible backup behavior can also be enabled by "total information awareness", meaning relevant communications are circulated to all team members, including those not directly involved [51]. Such communication practices are linked to outcomes like improved team performance [48]. Consistent rules and norms can further reinforce positive team functioning in virtual contexts [41], [52], [53]. However, it is emphasized that virtual teams must deliberately invest in relationship-building [47]. For example, proactively introducing new colleagues through informal sessions can foster familiarity even faster than in co-located teams.

C. *Trust as a Cross-Cutting Enabling Condition*

Trust can be defined as a dynamic, shared state that emerges within teams and reflects the belief that colleagues are reliable and aligned in terms of goals and values [27]. In virtual teams, trust is closely intertwined with social presence and situation awareness, as both the relational and informational dimensions of collaboration shape how team members perceive and rely on one another. Research highlights that even brief face-to-face encounters, such as an initial in-person kickoff, can accelerate the development of early trust, which subsequently carries over into virtual collaboration. When in-person interaction is not feasible, structured opportunities for informal exchange, for example virtual teambuilding events or open communication channels, can partially substitute for missing physical closeness and foster interpersonal bonding [47], [54]. In these contexts, digital communication technologies must support both task-related coordination and relational interaction to facilitate the development and maintenance of trust [55]. Trust and team cohesion are highly interdependent: when early trust is established, cohesive collaboration develops more easily, whereas low initial trust complicates maintaining unity across geographic and temporal divides [27].

Similarly, a study highlights that without deliberate efforts to foster socio-emotional bonds, relationships between subordinates and managers can deteriorate in purely virtual settings, further undermining trust [50].

In the context of social presence, shared emotional cues, authentic self-disclosure, and informal interactions foster affective trust. High-synchronicity media convey nonverbal cues that help team members assess intentions, while excessive asynchronous text communication can undermine trust due to misinterpretation [33]. A well-balanced media strategy that supports both structured collaboration and spontaneous social contact strengthens trust at the group level [43], [48].

This finding resonates with SET, which emphasizes that trust develops through reciprocal exchanges that signal commitment and mutual support over time [26]. In virtual contexts, well-designed communication rituals, for example structured check-ins or digital recognition, can function as symbolic gestures that sustain perceived reliability [27].

From a situation awareness perspective, low trust hampers information flow, as individuals hesitate to disclose difficulties or seek help in perceived hostile environments. Weak trust discourages transparent data sharing due to fears of appearing incompetent or misusing sensitive information [51]. Conversely, high-trust teams proactively coordinate, offer support, and monitor emerging challenges, enhancing performance in remote settings [27], [42]. At the leadership level, trust mitigates the "invisibility problem" in remote work, enabling fair treatment, nonintrusive oversight, and accountability without micromanagement [23], [50], [56].

Social presence creates the emotional conditions under which trust thrives. Situation awareness ensures group members have enough insight into one another's actions and challenges to maintain a trustworthy, dependable workflow.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Tying these two elements together, trust becomes the linchpin that channels technological affordances into truly cohesive remote teamwork.

D. Leadership as a Contextual Shaping Factor

Leadership plays a central role in shaping how digital collaboration tools foster social presence and situation awareness in remote work. Studies show that leaders must actively structure digital communication to strengthen interpersonal connection and team cohesion. Inclusive leadership practices, such as fair recognition and open participation, help employees feel part of the group and thus increase social presence [23], [34]. At the same time, compassionate leadership that emphasizes empathy and individualized support reduces the sense of isolation in virtual settings and helps maintain engagement when face-to-face cues are absent [37]. Conversely, insufficient contact or lack of feedback from managers working remotely can weaken trust and reduce perceived presence [51].

With respect to situation awareness, leaders are responsible for ensuring transparency and visibility of work progress. Digital routines such as structured check-ins, clear goal-setting, and regular updates provide team members with the contextual information needed to anticipate each other's needs [40], [42], [57]. Shared leadership models can also strengthen situation awareness when responsibilities are distributed fairly and task interdependencies are made explicit [41]. However, it is cautioned that not all leadership functions can be shared without risking confusion. Some decision-making authority remains essential with the formal leader to maintain clarity.

Digital collaboration tools are both an enabler and a challenge in this context. On the one hand, platforms, messaging tools, and video conferencing enhance awareness of team activities and make collaboration visible [44], [58]. On the other hand, excessive monitoring or rigid control mechanisms may erode trust and reduce employees' willingness to share information, thereby undermining situation awareness [38], [52]. Training initiatives are therefore critical. Leaders who are skilled at using digital tools strategically can align digital practices with organizational goals, fostering social presence and contextual awareness among their teams [36], [49], [54], [55].

V. Illustrative Case Examples from Literature

To further contextualize the findings of this review, this section presents two illustrative examples from the literature that demonstrate how specific collaboration tools establish social presence and situation awareness as central IT characteristics of virtual collaboration. The examples illustrate how these socially oriented characteristics are embedded in concrete tool affordances and everyday communication practices.

A. Gather: Virtual Office Environments Supporting Social Presence and Situation Awareness

The use of Gather as a virtual office environment in remote software engineering teams provides a particularly illustrative example of how social presence and situation awareness can be jointly supported through digital technology [34]. Gather simulates key spatial features of co-located work environments by allowing team members to navigate shared virtual spaces with avatars and to initiate interactions based on proximity. This design enables spontaneous, low-threshold encounters that closely resemble informal office interactions, thereby fostering a strong sense of social presence grounded in perceived co-presence and interpersonal accessibility. At the same time, these spatial and visual affordances support situation awareness by making social and task-related activity continuously observable. Team members can see who is present, who is engaged in conversations, and how different areas of the virtual workspace are used, which contributes to an ongoing shared understanding of availability, interaction patterns, and work context. Accordingly, Gather demonstrates how social presence and situation awareness serve as enabling IT characteristics that compensate for the absence of physical co-location, enabling meaningful collaboration in virtual settings.

MRT further helps explain why Gather is particularly effective for interactions involving ambiguity or negotiation, as it combines multiple cues, instantaneous communication, and personal focus. In addition, MST highlights how the platform supports convergence processes through real-time interaction, allowing teams to align interpretations and coordinate action in situ. Leadership practices play a mediating role by encouraging participation in shared spaces and legitimizing informal interaction, thereby reinforcing the reciprocal social exchanges through which trust and cohesion emerge over time.

B. Slack: Asynchronous Communication and Coordination in Remote Work

Another example is offered by a study investigating Slack adoption during the abrupt transition to remote work at the onset of the COVID-19 pandemic [39]. In this context, professional service teams initially tried to recreate in-office collaboration by holding frequent synchronous video meetings. However, this approach soon led to fatigue and inefficiencies. As a response, many teams shifted to Slack as their primary coordination channel. Through persistent channels and threaded discussions, Slack supported situation awareness by making task updates, process decisions, and contextual information visible and accessible across time. Team members were able to reconstruct the state of ongoing work, monitor progress, and understand dependencies without requiring continuous real-time interaction. Simultaneously, Slack sustained a baseline level of social presence through informal messages, quick questions, and lightweight interpersonal exchanges that approximated the ease of brief, in-person interactions. This example illustrates how social presence and situation awareness operate as core IT characteristics that enable distributed coordination and mutual orientation in asynchronous collaboration. MST helps clarify how Slack primarily supports conveyance processes, allowing parallel information processing and reflection, while still enabling convergence through focused discussions when needed. Leadership practices were central in shaping

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

these dynamics by establishing norms around responsiveness, channel use, and availability, thereby balancing efficiency concerns with the maintenance of relational connection and inclusiveness. From a SET perspective, these practices support sustained reciprocity by signaling commitment and accessibility without imposing constant immediacy.

These examples demonstrate how different categories of collaboration tools operationalize social presence and situation awareness as key IT characteristics of virtual collaboration. Immersive environments such as Gather emphasize embodied co-presence and real-time mutual awareness, whereas asynchronous platforms like Slack foreground persistence, transparency, and temporal flexibility. Both cases highlight the central role of tool affordances in fostering social presence and situation awareness and demonstrate how leadership practices mediate their effective integration into remote workflows.

VI. AI-Augmented Collaboration in Virtual Teams

Recent advances in AI are increasingly reshaping digital collaboration tools used in remote work settings. While collaboration technologies have traditionally been conceptualized as infrastructures that mediate interaction and support shared understanding across spatial distance, recent research highlights a shift toward AI-driven functionalities that actively structure communicative processes. Across a wide range of organizational and technological contexts, studies describe how AI is embedded in collaboration environments, where it supports coordination, filters information, and augments communicative exchanges among geographically distributed team members [59], [60], [61]. Through these mechanisms, AI is involved in establishing social presence by influencing perceptions of availability, responsiveness, and emotional expression. At the same time, AI affects situation awareness by organizing access to relevant information and collective states.

Moreover, the literature consistently emphasizes the ambivalent character of AI-mediated collaboration. Several contributions suggest that AI-enabled features can strengthen social presence in remote teams by reducing interactional friction and supporting more continuous and context-sensitive communication, for instance through intelligent interface design, personalization mechanisms, or enhanced virtual meeting environments [62], [63], [64].

Other studies warn that increased algorithmic mediation may weaken social presence when communication becomes standardized, opaque, or overly optimized, thereby constraining opportunities for informal interaction and mutual sensemaking [65], [66]. Comparable tensions are evident with regard to situation awareness, which may be enhanced through AI-supported analytics while simultaneously becoming detached from socially grounded cues.

Studies further illustrate that AI-mediated functionalities are no longer peripheral but increasingly embedded in routine collaboration and managerial practices. Current uses of AI in collaboration tools include automated scheduling, intelligent task allocation, real-time transcription, sentiment analysis, and workflow automation, often in conjunction with cloud-based infrastructures and enterprise systems [67], [68], [69]. In software development and engineering contexts, AI-supported meeting tools and analytics are described as particularly relevant for maintaining shared situation awareness in distributed teams [53], [70]. These applications are frequently framed as responses to the growing complexity of remote collaboration, where maintaining awareness of tasks, roles, and progress requires continuous coordination across temporal and spatial boundaries.

Across the reviewed literature, positive effects of AI-augmented collaboration tools are frequently associated with enhanced transparency, efficiency, and informed decision-making [53], [67], [71], [72]. Several studies report that AI-based analytics increase visibility into participation patterns, workload distribution, and performance indicators, thereby supporting situation awareness in complex organizational environments [73], [74]. In crisis-related contexts, AI-enabled tools are described as helping organizations sustain coordination and awareness despite disruption and uncertainty [75]. From this perspective, AI is portrayed as reinforcing shared understanding and continuity in distributed collaboration, particularly where direct interaction is limited.

However, negative effects and unintended consequences are equally prominent in the literature. Multiple contributions raise concerns about surveillance, loss of autonomy, and the normalization of pervasive monitoring through AI-driven systems, which may undermine trust and weaken social presence [66], [71]. Other studies highlight risks related to algorithmic bias, security vulnerabilities, and misuse of AI-enabled infrastructures, which may distort awareness structures and marginalize certain actors or interaction styles [76], [77], [78]. These findings point to a tension between efficiency gains and the erosion of socially grounded cues that are central to effective collaboration.

Looking ahead, the literature anticipates a continued expansion of AI capabilities within collaboration tools and organizational systems. Future developments are expected to include more adaptive, predictive, and autonomous systems that integrate collaboration data with broader human capital and labor market analytics [79], [80]. At the team level, AI-driven feedback and support systems are discussed as potential mechanisms for sustaining well-being, performance, and awareness in remote settings [81]. At the same time, several contributions emphasize that the implications of these developments depend on governance frameworks, ethical considerations, and organizational choices [82], [83]. Rather than converging on a single trajectory, the literature portrays AI-augmented collaboration as an evolving and contested domain, in which opportunities for enhanced social presence and situation awareness coexist with significant risks of fragmentation, depersonalization, and loss of agency.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

VII. Conceptual Contributions and Practical Implications

Based on the results presented, this study makes several conceptual contributions to the field of virtual collaboration research. It extends the PVT model and refines its applicability to contemporary, digitally mediated knowledge work. In contrast to traditional interpretations of PVT, which focus on assessing the virtualization potential of organizational processes based on sensory, relational, synchronous, and identification and control requirements [4], [9], this review builds on and systematizes more recent extensions of the theory by explicitly foregrounding social presence and situation awareness as IT characteristics of digital collaboration [5]. The contribution of this study lies in synthesizing and structuring existing literature to demonstrate how these socially oriented IT characteristics enable the virtualization of relationally demanding collaboration processes. By integrating these characteristics into a broader PVT-informed analysis, the review clarifies their role in shaping interaction quality, shared understanding, and coordination in contemporary virtual work settings.

Building on crisis-driven extensions of PVT [5], the findings indicate that these characteristics are not merely supportive conditions but constitute core mechanisms through which virtual collaboration becomes sustainable beyond temporary or emergency contexts. Social presence captures the relational dimension of process virtualization by enabling interpersonal connection, belonging, and cohesion in the absence of physical co-location, whereas situation awareness reflects the informational dimension by supporting shared understanding of tasks, roles, and interdependencies [19], [22].

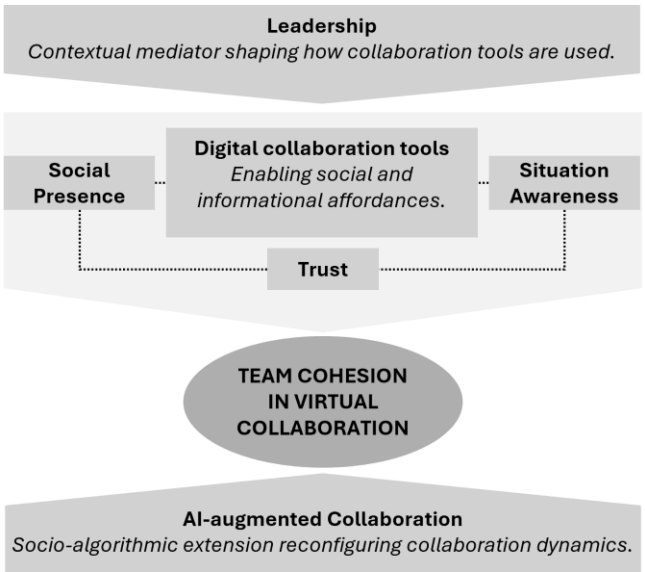


Figure 3. Conceptual synthesis of how digital collaboration tools support team cohesion in virtual collaboration.

Figure 3 synthesizes the core findings of this review by illustrating how digital collaboration tools contribute to team cohesion in virtual collaboration through distinct but interrelated social mechanisms. The analysis shows that collaboration tools primarily foster social presence and situation awareness by enabling social and informational affordances. Social presence supports relational closeness and a sense of interpersonal connection, while situation awareness facilitates shared understanding of tasks, activities, and coordination processes. Trust emerges as a central, relational outcome of the dynamic interplay between these two mechanisms rather than as an isolated characteristic, reinforcing both relational continuity and collaborative stability over time. Leadership is conceptualized as a contextual mediator that shapes how collaboration tools and their affordances are used in practice, thereby influencing the extent to which social presence, situation awareness, and trust can develop. Finally, AI-augmented collaboration is positioned as a socio-algorithmic extension that increasingly reconfigures collaboration dynamics by reshaping coordination, visibility, and sensemaking processes, without replacing the underlying social mechanisms on which team cohesion in virtual collaboration ultimately depends.

A key conceptual contribution of this article lies in its explicit integration of AI as an emerging layer of mediation in virtual collaboration. Whereas PVT traditionally conceptualizes information technology as an enabling infrastructure, the reviewed literature on AI-enabled collaboration tools points to a qualitative shift in how collaboration is shaped. AI has evolved from being a passive support technology to playing a more active role in group interactions [59], [62]. AI-supported functionalities increasingly structure communication, prioritize information, coordinate tasks, and interpret collective states, thereby actively influencing both social presence and situation awareness. This socio-algorithmic dimension challenges implicit assumptions of PVT regarding the passive role of IT and suggests the need to conceptualize algorithmic agency as an integral component of process virtualization in contemporary collaboration settings.

The study further refines virtualization research by integrating complementary theoretical lenses that address limitations of PVT when applied to socially complex work processes. MRT explains why media-rich communication formats remain crucial for sustaining social presence and affective trust in ambiguous contexts [24]. MST clarifies how communication structures support either convergence or conveyance processes that underpin situation awareness [25]. SET adds a relational perspective by emphasizing that trust and commitment emerge through reciprocal interactions over time, even in digitally mediated environments [26]. Together, these perspectives support a socio-technical and socio-algorithmic extension of PVT that foregrounds communication dynamics, relational expectations, and trust as central elements of virtual collaboration.

Another important contribution is the explicit incorporation of leadership as a contextual shaping force. The reviewed literature demonstrates that leadership

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

practices influence how digital and AI-enabled collaboration tools are used, interpreted, and legitimized in everyday work. By structuring communication routines, setting norms for transparency, and regulating visibility and monitoring features, leaders shape whether collaboration technologies foster trust, social presence, and shared awareness or instead contribute to overload, disengagement, and perceptions of control [23], [41], [52]. Conceptually, framing leadership as a contextual factor shaping how digital collaboration tools are used in practice extends PVT toward a more explicitly organizational and practice-oriented perspective.

From a practical standpoint, the findings imply that organizations should move beyond technology-centric approaches to virtual collaboration. Effective digital collaboration requires communication environments that combine synchronous and asynchronous interaction, support informal as well as task-oriented exchanges, and provide persistent yet adjustable visibility into work processes [39], [47]. Leadership development initiatives should therefore emphasize relational, communicative, and ethical competencies alongside technical skills, particularly in increasingly AI-mediated collaboration settings. For designers of collaboration tools, the results highlight the importance of configurable systems that allow teams to calibrate interaction richness, synchronicity, and algorithmic intervention in line with evolving task and relational demands [56], [84].

VIII. Conclusion and Future Work

This article extends an earlier conference publication [1] by broadening the theoretical and conceptual understanding of virtual collaboration and by integrating additional analytical perspectives on leadership and AI. Drawing on systematic literature reviews, the study demonstrates that virtual collaboration can replicate many relational and informational aspects of co-located work when technological affordances are aligned with communication practices, leadership behaviors, and organizational norms. Anchored in PVT [4], [9] and complemented by MRT [24], MST [25], and SET [26], the analysis highlights social presence and situation awareness as central mechanisms through which trust, cohesion, and shared understanding can be sustained in digitally mediated collaboration [19], [22].

Across the reviewed literature, social presence emerges as a relational outcome that is particularly supported by media-rich and synchronous interaction formats, informal communication opportunities, and practices that foster belonging and interpersonal connection [32], [35], [39]. Situation awareness, in contrast, is primarily enabled by persistent information structures, transparent communication routines, and ongoing knowledge exchange that make tasks, roles, and interdependencies visible to distributed team members [47], [48], [57]. Trust operates as a cross-cutting condition that links these two dimensions by shaping whether individuals are willing to share information, disclose uncertainty, and rely on one another in the absence of physical co-location [27], [51]. Leadership plays a decisive role in translating technological affordances into meaningful social and informational outcomes by structuring communication, legitimizing tool use, and balancing transparency with autonomy [23], [41], [52].

At the same time, the findings underscore that technology alone cannot ensure effective virtual collaboration. The complementary analysis of AI-enabled collaboration tools reveals an ambivalent dynamic: while AI-supported functionalities may enhance efficiency, coordination, and transparency [73], [75], they also raise concerns related to agency, surveillance, and the erosion of socially grounded interaction cues when algorithmic processes become opaque or intrusive [66], [71]. These tensions reinforce the conclusion that digital and AI-enabled tools must be embedded in appropriate governance structures and leadership practices to avoid undermining trust and social presence [84].

Beyond summarizing these contributions, this study points to several avenues for future research. Conceptually, future work would benefit from a more consistent analytical differentiation between coordination, communication, and collaboration in studies of virtual work. Rather than subsuming these dimensions under the broad label of virtual collaboration, distinguishing them as related but distinct processes can support more precise theorizing and empirical operationalization of digital collaboration processes [2].

Methodologically, longitudinal and experimental research designs are needed to move beyond descriptive studies and examine causal relationships between specific digital affordances, leadership behaviors, and outcomes such as trust, cohesion, and team performance [42]. Comparative studies across organizational contexts, industries, and cultural settings could further illuminate boundary conditions of PVT and clarify how virtual collaboration evolves over time.

In addition, the growing integration of AI into collaboration tools points to a distinct and increasingly urgent research agenda. While existing studies predominantly focus on short-term efficiency gains and coordination benefits, comparatively little is known about the long-term effects of AI-mediated collaboration on social presence, situation awareness, and trust [59], [60]. As AI systems increasingly structure communication flows, prioritize information, and generate interpretations of collective states, they assume a more active role in shaping collaboration processes rather than merely supporting them. Future research should therefore examine how different degrees of algorithmic mediation affect perceptions of agency, accountability, and reciprocity over time, particularly in contexts where AI substitutes for human judgment or interaction [67], [68], [69]. Longitudinal studies could shed light on whether AI-enhanced transparency and efficiency stabilize collaboration or gradually erode relational quality and trust when algorithmic processes remain opaque or are perceived as intrusive [66]. Moreover, further research is needed to identify governance mechanisms and design principles that allow organizations to calibrate AI support in ways that preserve socially grounded cues, maintain psychological safety, and align algorithmic intervention with task complexity and relational demands [82], [83].

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

REFERENCES

[1] I. F. Schlömer, "Revisiting Process Virtualization: A Systematic Review of How Collaboration Tools Support Social Presence and Situation Awareness," in *DIGITAL 2025: Advances on Societal Digital Transformation: July 6th-10th, 2025, Venice, Italy*, 2025, pp. 50–56. Accessed: Aug. 26, 2025. [Online]. Available: https://www.thinkmind.org/library/DIGITAL/DIGITAL_2025/digital_2025_2_70_10043.html

[2] L. M. Camarihna-Matos, H. Afsarmanesh, and H. Afsarmanesh, "Concept of Collaboration." Accessed: Jan. 07, 2025. [Online]. Available: https://www.igi-global.com/gateway/chapter/www.igi-global.com/gateway/chapter/17627

[3] J. M. Barrero, N. Bloom, and S. J. Davis, "The Evolution of Work from Home," *Journal of Economic Perspectives*, vol. 37, no. 4, pp. 23–49, Nov. 2023, doi: 10.1257/jep.37.4.23.

[4] E. Overby, "Process Virtualization Theory and the Impact of Information Technology," *Organization Science*, vol. 19, no. 2, pp. 277–291, Mar. 2008, doi: 10.1287/orsc.1070.0316.

[5] A. Zeuge, C. Schaefer, A. Weigel, A. Eckhardt, and B. Niehaves, "Crisis-driven digital transformation as a trigger for process virtualization: Fulfilling knowledge work process requirements for remote work," *International Journal of Information Management*, vol. 70, p. 102636, Jun. 2023, doi: 10.1016/j.ijinfomgt.2023.102636.

[6] J. Lipnack and J. Stamps, "Virtual teams: The new way to work," *Strategy & Leadership*, vol. 27, no. 1, pp. 14–19, Jan. 1999, doi: 10.1108/eb054625.

[7] M. Sako, "From remote work to working from anywhere," *Commun. ACM*, vol. 64, no. 4, pp. 20–22, Apr. 2021, doi: 10.1145/3451223.

[8] S. Zheng, M. Yan, Y. Liang, Y. Chen, Q. Wei, and S. Li, "Understanding the positive and negative effects of team virtuality: A theoretical review and research agenda," *Human Resource Management Review*, vol. 34, no. 2, p. 101013, Jun. 2024, doi: 10.1016/j.hrmr.2024.101013.

[9] E. Overby, "Migrating Processes from Physical to Virtual Environments: Process Virtualization Theory," in *Information Systems Theory*, vol. 28, Y. K. Dwivedi, M. R. Wade, and S. L. Schneberger, Eds., in Integrated Series in Information Systems, vol. 28. , New York, NY: Springer New York, 2012, pp. 107–124. doi: 10.1007/978-1-4419-6108-2_6.

[10] M. Chaudhary, V. Munjal, and V. Kumar, "Investigating Student Resistance and Intentions to Use E-learning Systems in Higher Education: An Integrated Application of PVT and UTAUT2 Model," *CAIS*, vol. 56, pp. 167–193, 2025, doi: 10.17705/1CAIS.05607.

[11] A. Alarabiat, O. Hujran, D. Soares, and A. Tarhini, "Examining students' continuous use of online learning in the post-COVID-19 era: an application of the process virtualization theory," *ITP*, vol. 36, no. 1, pp. 21–47, Jan. 2023, doi: 10.1108/ITP-02-2021-0142.

[12] B. Balci, D. Grgecic, and C. Rosenkranz, "Why People Reject or Use Virtual Processes: A Test of Process Virtualization Theory," *AMCIS 2013 Proceedings*, May 2013, [Online]. Available: https://aisel.aisnet.org/amcis2013/AdoptionOfIT/GeneralPresentations/4

[13] P. S. Ackom, K. Owusu Kwateng, F. K. Tetteh, and M. Wiesche, "Understanding e-government services: integration of process virtualization theory and user resistance," *DPRG*, vol. 24, no. 5, pp. 418–434, Nov. 2022, doi: 10.1108/DPRG-11-2021-0153.

[14] I. Barbeitos and T. Oliveira, "Continuance in Portuguese Peer-To-Peer Accommodation Services Through the Lens of Process Virtualization Theory," *Information Systems Management*, vol. 42, no. 1, pp. 68–88, Dec. 2024, doi: 10.1080/10580530.2024.2332197.

[15] A. Verbovetska, "The Impact of Financial Technology on Customer Intention to Use Financial Services through the Lenses of Process Virtualization Theory," *ICIS 2019 Proceedings*, 2019.

[16] Y. Feng, J. Park, and M. Feng, "What is holding back business process virtualization in the post-COVID-19 era? Based on process virtualization theory (PVT)," *Front. Psychol.*, vol. 14, Feb. 2023, doi: 10.3389/fpsyg.2023.1084180.

[17] M. Barth and D. Veit, "Which Processes Do Users Not Want Online? Extending Process Virtualization Theory," *ICIS 2011 Proceedings*, Dec. 2011.

[18] B. Balci and C. Rosenkranz, "'Virtual or material, What do You Prefer?' a Study of Process Virtualization Theory," *ECIS 2014 Proceedings*, Jun. 2014.

[19] J. Fulk, J. Schmitz, and C. W. Steinfeld, "A Social Influence Model of Technology Use," in *Organizations and Communication Technology*, Newbury Park, London, New Delhi: Sage Publications, 1990.

[20] J. Short, E. Williams, and B. Christie, "The social psychology of telecommunications." John Wiley, 1976.

[21] K. Kreijns, K. Xu, and J. Weidlich, "Social Presence: Conceptualization and Measurement," *Educ Psychol Rev*, vol. 34, no. 1, pp. 139–170, Mar. 2022, doi: 10.1007/s10648-021-09623-8.

[22] M. R. Endsley, "Toward a Theory of Situation Awareness in Dynamic Systems," *Hum Factors*, vol. 37, no. 1, pp. 32–64, Mar. 1995, doi: 10.1518/001872095779049543.

[23] S. Torres and M. A. Orhan, "How it started, how it's going: Why past research does not encompass pandemic-induced remote work realities and what leaders can do for more inclusive remote work practices," *Psychology of Leaders and Leadership*, vol. 26, no. 1, pp. 1–21, 2023, doi: 10.1037/mgr0000135.

[24] R. L. Daft and R. H. Lengel, "Organizational Information Requirements, Media Richness and Structural Design," *Management Science*, vol. 32, no. 5, pp. 554–571, May 1986, doi: 10.1287/mnsc.32.5.554.

[25] A. R. Dennis, R. M. Fuller, and J. S. Valacich, "Media, Task, and Communication Processes: A Theory of Media Synchronicity.," *MIS Quarterly*, vol. 32, no. 3, pp. 575–600, Sep. 2008, doi: 10.2307/25148857.

[26] P. Blau, *Exchange and Power in Social Life*, 2nd ed. New York: Routledge, 2017. doi: 10.4324/9780203792643.

[27] J. V. Dinh, D. L. Reyes, L. Kayga, C. Lindgren, J. Feitosa, and E. Salas, "Developing team trust: Leader insights for virtual settings," *Organizational Dynamics*, vol. 50, no. 1, 2021, doi: 10.1016/j.orgdyn.2021.100846.

[28] J. vom Brocke *et al.*, "Reconstructing the giant: On the importance of rigour in documenting the literature search process," in *17th European Conference on Information Systems, ECIS 2009*, 2009, pp. 2206–2217.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

[29] H. M. Cooper, "Organizing knowledge syntheses: A taxonomy of literature reviews," *Knowledge in society*, vol. 1, no. 1, pp. 104–126, Jan. 1988.

[30] M. J. Page *et al.*, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, p. n71, Mar. 2021, doi: 10.1136/bmj.n71.

[31] U. Kuckartz, *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*, 4., Überarbeitete Aufl. in Grundlagentexte Methoden. Weinheim: Beltz, 2018.

[32] G. C. Muresan, S. Mititelu, J. Andres, and M. C. Schraefel, "'Should I Introduce myself?': Asynchronous semi-guided professional introductions for enhanced perceived team effectiveness in new virtual dyadic teams," *International Journal of Human-Computer Studies*, vol. 188, p. 103279, Aug. 2024, doi: 10.1016/j.ijhcs.2024.103279.

[33] O. Torro, H. Pirkkalainen, and H. Li, "Media synchronicity in organizational social exchange," *ITP*, vol. 35, no. 8, pp. 162–180, Dec. 2022, doi: 10.1108/ITP-06-2020-0384.

[34] L. Okpara, C. Werner, A. Murray, and D. Damian, "The role of informal communication in building shared understanding of non-functional requirements in remote continuous software engineering," *Requirements Eng*, vol. 28, no. 4, pp. 595–617, Dec. 2023, doi: 10.1007/s00766-023-00404-z.

[35] J. Yang and H. Fang, "Integrating Embodied Social Presence Theory and Process Virtualization Theory to assess business process virtualizability: The mediating role of embodied co-presence," *PLoS ONE*, vol. 19, no. 6, p. e0305423, Jun. 2024, doi: 10.1371/journal.pone.0305423.

[36] S. Bilderback and M. D. Kilpatrick, "Global perspectives on redefining workplace presence: the impact of remote work on organizational culture," *JEET*, vol. 4, no. 1, pp. 62–72, Oct. 2024, doi: 10.1108/JEET-08-2024-0023.

[37] K.-L. Tan, A. K. S. Sim, T. Y. Lew, T.-H. Cham, and I. S. H. Hii, "Beyond laptops and tables: unveiling Singapore's success in hybrid work through a two-wave gender multigroup analysis of compassionate leadership," *ER*, vol. 47, no. 1, pp. 148–172, Jan. 2025, doi: 10.1108/ER-02-2024-0108.

[38] A. Asatiani and L. Norström, "Information systems for sustainable remote workplaces," *The Journal of Strategic Information Systems*, vol. 32, no. 3, p. 101789, Sep. 2023, doi: 10.1016/j.jsis.2023.101789.

[39] A. Whillans, L. Perlow, and A. Turek, "Experimenting during the shift to virtual team work: Learnings from how teams adapted their activities during the COVID-19 pandemic," *Information and Organization*, vol. 31, no. 1, p. 100343, Mar. 2021, doi: 10.1016/j.infoandorg.2021.100343.

[40] E. De Bruyne and D. Gerritse, "Exploring the future workplace: results of the futures forum study," *JCRE*, vol. 20, no. 3, pp. 196–213, Nov. 2018, doi: 10.1108/JCRE-09-2017-0030.

[41] C. Mayer, T. Sivatheerthan, S. Mütze-Niewöhner, and V. Nitsch, "Sharing leadership behaviors in virtual teams: effects of shared leadership behaviors on team member satisfaction and productivity," *TPM*, vol. 29, no. 1/2, pp. 90–112, Feb. 2023, doi: 10.1108/TPM-07-2022-0054.

[42] L. L. Gilson, P. Costa, T. A. O'Neill, and M. T. Maynard, "Putting the 'TEAM' back into virtual teams," *Organizational Dynamics*, vol. 50, no. 1, p. 100847, Jan. 2021, doi: 10.1016/j.orgdyn.2021.100847.

[43] C. Miller, P. Rodeghero, M.-A. Storey, D. Ford, and T. Zimmermann, "'How Was Your Weekend?' Software Development Teams Working From Home During COVID-19," in *2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE)*, May 2021, pp. 624–636. doi: 10.1109/ICSE43902.2021.00064.

[44] A. Alaiad, Y. Alnsour, and M. Alsharo, "Virtual Teams: Thematic Taxonomy, Constructs Model, and Future Research Directions," *IEEE Transactions on Professional Communication*, vol. 62, no. 3, pp. 211–238, Sep. 2019, doi: 10.1109/TPC.2019.2929370.

[45] T. Losev, S. Storteboom, S. Carpendale, and S. Knudsen, "Distributed Synchronous Visualization Design: Challenges and Strategies," in *2020 IEEE Workshop on Evaluation and Beyond - Methodological Approaches to Visualization (BELIV)*, Salt Lake City, UT, USA: IEEE, Oct. 2020, pp. 1–10. doi: 10.1109/BELIV51497.2020.00008.

[46] L. Liu, H. Van Essen, and B. Eggen, "An exploratory study of how to design interventions to support informal communication in remote work," in *Nordic Human-Computer Interaction Conference*, Aarhus Denmark: ACM, Oct. 2022, pp. 1–10. doi: 10.1145/3546155.3546673.

[47] B. Viererbl, N. Denner, and T. Koch, "'You don't meet anybody when walking from the living room to the kitchen': informal communication during remote work," *JCOM*, vol. 26, no. 3, pp. 331–348, Aug. 2022, doi: 10.1108/JCOM-10-2021-0117.

[48] A. Lechner and J. Tobias Mortlock, "How to create psychological safety in virtual teams," *Organizational Dynamics*, vol. 51, no. 2, p. 100849, Apr. 2022, doi: 10.1016/j.orgdyn.2021.100849.

[49] M. Mahtta, R. Pillai, A. Gunasekaran, B. Sivathanu, and N. Kaushik, "A pathway to virtual team performance in the New Normal paradigm," *JOEPP*, vol. 9, no. 4, pp. 549–571, Sep. 2022, doi: 10.1108/JOEPP-11-2020-0218.

[50] I. Wahl, D. Wolfgruber, and S. Einwiller, "Mitigating teleworkers' perceived technological complexity and work strains through supportive team communication," *CCIJ*, vol. 29, no. 3, pp. 329–345, Apr. 2024, doi: 10.1108/CCIJ-05-2023-0061.

[51] T. D. Golden and A. Fromen, "Does it matter where your manager works? Comparing managerial work mode (traditional, telework, virtual) across subordinate work experiences and outcomes," *Human Relations*, vol. 64, pp. 1451–1475, Nov. 2011, doi: 10.1177/0018726711418387.

[52] L. Chernyak-Hai and E. Rabenu, "The New Era Workplace Relationships: Is Social Exchange Theory Still Relevant?," *Ind. Organ. Psychol.*, vol. 11, no. 3, pp. 456–481, Sep. 2018, doi: 10.1017/iop.2018.5.

[53] A. S. L. De Andrade, V. Jackson, R. Prikladnicki, and A. Van Der Hoek, "On meetings involving remote software teams: A systematic literature review," *Information and Software Technology*, vol. 175, p. 107541, Nov. 2024, doi: 10.1016/j.infsof.2024.107541.

[54] E. Raguseo, L. Gastaldi, and P. Neirotti, "Smart work: Supporting employees' flexibility through ICT, HR practices and office layout," *EBHRM*, vol. 4, no. 3, pp. 240–256, Dec. 2016, doi: 10.1108/EBHRM-01-2016-0004.

[55] K. Tworek, "E-leadership shaped by IT adaptability through Employees' Dynamic Capabilities," *Procedia Computer Science*, vol. 225, pp. 357–365, 2023, doi: 10.1016/j.procs.2023.10.020.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

[56] R. Gong and M. Hua, "Designing Multimodal User Interfaces for Hybrid Collaboration: A User-Centered Approach," in *HCI International 2023 – Late Breaking Papers*, vol. 14054, M. Kurosu, A. Hashizume, A. Marcus, E. Rosenzweig, M. M. Soares, D. Harris, W.-C. Li, D. D. Schmorrow, C. M. Fidopiastis, and P.-L. P. Rau, Eds., in Lecture Notes in Computer Science, vol. 14054. , Cham: Springer Nature Switzerland, 2023, pp. 67–82. doi: 10.1007/978-3-031-48038-6_5.

[57] R. Paul, J. R. Drake, and H. Liang, "Global Virtual Team Performance: The Effect of Coordination Effectiveness, Trust, and Team Cohesion," *IEEE Transactions on Professional Communication*, vol. 59, no. 3, pp. 186–202, Sep. 2016, doi: 10.1109/TPC.2016.2583319.

[58] S. A. Newman and R. C. Ford, "Five Steps to Leading Your Team in the Virtual COVID-19 Workplace," *Organ Dyn*, vol. 50, no. 1, p. 100802, 2021, doi: 10.1016/j.orgdyn.2020.100802.

[59] M. Lazarova, P. Caligiuri, D. G. Collings, and H. De Cieri, "Global work in a rapidly changing world: Implications for MNEs and individuals," *Journal of World Business*, vol. 58, no. 1, p. 101365, Jan. 2023, doi: 10.1016/j.jwb.2022.101365.

[60] A. Russell and E. Frachtenberg, "Worlds Apart: Technology, Remote Work, and Equity," *Computer*, vol. 54, no. 7, pp. 46–56, Jul. 2021, doi: 10.1109/MC.2021.3074121.

[61] C. Sarath Kumar and M. Krithika, "Creating Effective Communication Solutions for Hybrid Work Environments using Cisco Webex," in *2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, India: IEEE, Nov. 2024, pp. 296–302. doi: 10.1109/ICECA63461.2024.10800821.

[62] M. Darban, "Navigating virtual teams in generative AI-led learning: The moderation of team perceived virtuality," *Educ Inf Technol*, vol. 29, no. 17, pp. 23225–23248, Dec. 2024, doi: 10.1007/s10639-024-12681-4.

[63] P. Ghadekar, Y. Shaligram, V. Shimpi, O. Shingade, S. Singh, and S. Thakur, "AI-Enhanced Virtual Meeting Platform: Integrating Spatial Audio and Personalization Features along with Smart Meeting Summarization," in *2025 4th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0*, Raigarh, India: IEEE, Apr. 2025, pp. 1–6. doi: 10.1109/OTCON65728.2025.11070561.

[64] K. M. Kuhn, "The constant mirror: Self-view and attitudes to virtual meetings," *Computers in Human Behavior*, vol. 128, p. 107110, Mar. 2022, doi: 10.1016/j.chb.2021.107110.

[65] M. Abid, O. Ben-Salha, K. Gasmi, and N. H. A. Alnor, "Modelling for disability: How does artificial intelligence affect unemployment among people with disability? An empirical analysis of linear and nonlinear effects," *Research in Developmental Disabilities*, vol. 149, p. 104732, Jun. 2024, doi: 10.1016/j.ridd.2024.104732.

[66] M. Constantinides and D. Quercia, "Good Intentions, Bad Inventions: How Employees Judge Pervasive Technologies in the Workplace," *IEEE Pervasive Comput.*, vol. 22, no. 1, pp. 69–76, Jan. 2023, doi: 10.1109/MPRV.2022.3217408.

[67] F. Liu and J. Li, "Application Exploration of Robot Process Automation in Digital Labor Time Management System," *Procedia Computer Science*, vol. 228, pp. 89–97, 2023, doi: 10.1016/j.procs.2023.11.012.

[68] N. K. Sravan, V. A. Vuyyyru, P. Gottapu, A. Theeda, and L. Peddineni, "Agile Management Tools: Technological Evaluations and Future Archetype," in *2024 4th International Conference on Pervasive Computing and Social Networking (ICPCSN)*, Salem, India: IEEE, May 2024, pp. 914–918. doi: 10.1109/ICPCSN62568.2024.00154.

[69] R. Younis, M. Iqbal, K. Munir, M. A. Javed, M. Haris, and S. Alahmari, "A Comprehensive Analysis of Cloud Service Models: IaaS, PaaS, and SaaS in the Context of Emerging Technologies and Trend," in *2024 International Conference on Electrical, Communication and Computer Engineering (ICECCE)*, Kuala Lumpur, Malaysia: IEEE, Oct. 2024, pp. 1–6. doi: 10.1109/ICECCE63537.2024.10823401.

[70] B. M. Moreno-Cabezali, "Strategic Human Capital Management in Engineering Workplaces: A Fuzzy Logic-Based Decision Framework," *IEEE Access*, vol. 13, pp. 125458–125472, 2025, doi: 10.1109/ACCESS.2025.3589207.

[71] D. R. Lombardi, J. C. Sipior, and D. A. Appelbaum, "Audit partners' and associates' insights and experiences of remote work in response to a global crisis Event: Implications for the profession and future directions," *The British Accounting Review*, p. 101701, Jul. 2025, doi: 10.1016/j.bar.2025.101701.

[72] K. Al-Suwaidi, I. Alsyouf, Z. Tahboub, A. Alsaad, and A. Alsyouf, "Integrating fourth industrial revolution technologies in human resource practices: Mediating roles of work flexibility, quality of work life, and distributive fairness in enhancing employee productivity," *Sustainable Futures*, vol. 10, p. 101150, Dec. 2025, doi: 10.1016/j.sftr.2025.101150.

[73] H. Huang and L. Zhao, "Does digital innovation help firms navigate the COVID-19 pandemic? Evidence from China," *Emerging Markets Review*, vol. 62, p. 101171, Sep. 2024, doi: 10.1016/j.ememar.2024.101171.

[74] S. Patil, P. Upadhye, S. R. Khandekar, P. Darekar, A. Ghewari, and U. Bhattacharya, "Artificial Intelligence and Data Science Approaches to Analyzing Human Resource Management in Hybrid and Remote Work Models," in *2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, Chennai, India: IEEE, Dec. 2024, pp. 1–6. doi: 10.1109/ICSES63760.2024.10910810.

[75] G. Narayanamurthy and G. Tortorella, "Impact of COVID-19 outbreak on employee performance – Moderating role of industry 4.0 base technologies," *International Journal of Production Economics*, vol. 234, p. 108075, Apr. 2021, doi: 10.1016/j.ijpe.2021.108075.

[76] R. Mokkapati and V. L. Dasari, "An Artificial Intelligence Enabled Self Replication System Against Cyber Attacks," in *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India: IEEE, Jan. 2023, pp. 698–703. doi: 10.1109/ICSSIT55814.2023.10061089.

[77] B. V. Raghav, N. S. Sree, S. Parnitha, E. S. S. Meghanaa, and R. Challa, "A Comprehensive Analysis on Online Masquerade Attacks," in *2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*,

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Salem, India: IEEE, Jun. 2024, pp. 1375–1380. doi: 10.1109/ICAAIC60222.2024.10575147.

[78] M. Zaoui, B. Yousra, S. Yassine, M. Yassine, and O. Karim, "A Comprehensive Taxonomy of Social Engineering Attacks and Defense Mechanisms: Toward Effective Mitigation Strategies," *IEEE Access*, vol. 12, pp. 72224–72241, 2024, doi: 10.1109/ACCESS.2024.3403197.

[79] G. Manoharan, P. Sharma, V. Chaudhary, P. Chatterjee Biswas, M. K. Sharma, and M. Lourens, "The Future of Work: Examining the Impact of AI/ML on Job roles, Organizational Structures, and Talent Management Practices," in *2024 International Conference on Trends in Quantum Computing and Emerging Business Technologies*, Pune, India: IEEE, Mar. 2024, pp. 1–6. doi: 10.1109/TQCEBT59414.2024.10545125.

[80] H. Zhou, L. Wang, Y. Cao, and J. Li, "The impact of artificial intelligence on labor market: A study based on bibliometric analysis," *Journal of Asian Economics*, vol. 98, p. 101926, Jun. 2025, doi: 10.1016/j.asieco.2025.101926.

[81] R. Kuhlmann and I. Klingenberg, "The impact of leader and team PsyCap on employees' individual PsyCap: An experimental analysis of transmission effects in virtual and non-virtual settings," *Computers in Human Behavior Reports*, vol. 18, p. 100622, May 2025, doi: 10.1016/j.chbr.2025.100622.

[82] U. J. Butt, W. Richardson, A. Nouman, H.-M. Agbo, C. Eghan, and F. Hashmi, "Cloud and Its Security Impacts on Managing a Workforce Remotely: A Reflection to Cover Remote Working Challenges," in *Cybersecurity, Privacy and Freedom Protection in the Connected World*, H. Jahankhani, A. Jamal, and S. Lawson, Eds., in Advanced Sciences and Technologies for Security Applications. , Cham: Springer International Publishing, 2021, pp. 285–311. doi: 10.1007/978-3-030-68534-8_18.

[83] M. Qin, X. Shao, Y. Zhu, and C.-T. Lin, "Harnessing artificial intelligence for environmental protection: Smart air quality management under oil price fluctuations," *Energy Economics*, vol. 151, p. 108892, Nov. 2025, doi: 10.1016/j.eneco.2025.108892.

[84] M. Landowski, J.-S. Effert, F. Günther, M. Tebart, and F. Moerike, "Beyond Hallway Chats ? Negotiating Content Awareness in Hybrid Work Practices," in *Proceedings of Mensch und Computer 2024*, Karlsruhe Germany: ACM, Sep. 2024, pp. 508–513. doi: 10.1145/3670653.3677523.

[85] L. S. Müller, S. Reiners, J. Becker, and G. Hertel, "Long-term effects of COVID-19 on work routines and organizational culture – A case study within higher education's administration," *Journal of Business Research*, vol. 163, p. 113927, Aug. 2023, doi: 10.1016/j.jbusres.2023.113927.

[86] M. Marinho, L. Amorim, R. Camara, B. R. Oliveira, M. Sobral, and S. Sampaio, "Happier and further by going together: The importance of software team behaviour during the COVID-19 pandemic," *Technology in Society*, vol. 67, p. 101799, Nov. 2021, doi: 10.1016/j.techsoc.2021.101799.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

I. APPENDIX

EXPANDED CONCEPT MATRIX

	Classification		Concepts						
Article	Research Design	Tool-Types mentioned	Social Presence			Situation awareness		Trust	Leadership
			Interpersonal Relationship/ Team Cohesion	*Collaboration*	*Community/ belonging*	*Knowledge exchange*	*Informal Communication*		
[23]	REV	Vid, Pl	x		x		x	x	x
[27]	REV + Q	Vid, Em	x	x	x		x	x	
[32]	CON	Vid, Pl	x	x	x	x	x	x	
[33]	REV + CON	VR,Vid,Ch	x	x	x			x	
[34]	Q	VR, Vid, Ch, Pl	x	x	x		x		x
[35]	CON	Pl	x		x				
[36]	CON	Pl	x		x		x		x
[37]	QU	Vid, Ch	x		x	x		x	x
[38]	REV	Pl, Em	x		x	x	x	x	x
[39]	Q	Vid, Ch, Em	x				x		
[40]	Q	Pl	x	x	x		x		x
[41]	QA	ND	x	x	x			x	x
[42]	REV	Vid, Ch, Pl, Em					x	x	x
[43]	QM	Ch, NT	x				x		
[44]	REV	VR, Vid, Ch, Pl, Em	x	x		x		x	x
[45]	Q	Vid, Ch, Pl, Ph	x	x				x	
[46]	Q + CON	Vid, Ch					x		
[47]	Q	Vid, Ch, Em, Ph	x		x		x		
[48]	Q	VR, Vid, Ch, Em	x			x	x	x	
[49]	Q	Vid, PL	x	x		x		x	x
[50]	QU	Vid, Ch, Pl, Em, Ph				x		x	
[51]	QU	Em, Ph, ND	x				x	x	x
[52]	CON	Pl						x	x
[53]	RE	VR, Vid, Ch, Pl, Em	x	x	x		x	x	
[54]	QM	ND		x	x				
[55]	QU	Vid, Ch, Pl		x	x	x			x
[56]	CON	Vid, Ch, Pl, Em, Ph	x	x				x	
[57]	REV + QU	Vid, Ch, Pl, Em	x					x	x
[58]	CON	Vid, Ch, Pl, Em, Ph	x	x	x		x	x	x
[84]	Q	VD	x				x		
[85]	Q	Vid, Ch, Pl, Em, Ph	x	x		x		x	
[86]	QU	Vid, Ch, Pl,	x	x			x		
∑ *(Research Design/ Tool-Types/ Segments with concept)*	Q = 10 CON = 8 REV = 7 QU = 6 **QM = 2** **EXP = 0**	Vid = 22 Pl = 18 Em = 13 Ch= 15 Ph = 7 **VR = 5** ND = 3	85	28	36	16	106	53	110

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Learner Models:

The Gap between Educational Opportunities vs. Data Protection Challenges

Felix Böck[1] [0000-0001-7382-8333], Hendrik Link[2] [0009-0006-0596-2120] and Dieter Landes[1] [0000-0002-0741-3540]

[1] Center for Responsible Artificial Intelligence (CRAI),
Coburg University of Applied Sciences and Arts, 96450 Coburg, Germany
{felix.boeck, dieter.landes}@hs-coburg.de

[2] Public Law, IT Law, and Environmental Law,
University of Kassel, 34109 Kassel, Germany
hendrik.link@uni-kassel.de

***Abstract* — Learners differ vastly in various aspects of what they need for successful learning. Artificial Intelligence (AI) provides a basis for digital learning environments that automatically adapt to learners' needs. To do so, these systems presuppose knowledge of the individual learner. Learner models are digital representations of learner characteristics that aim to enable personalised and adaptive learning experiences, addressing key areas such as transparency, fairness, data protection, modularity, and sustainability. Such learner models form the core of AI-based adaptive learning environments, as they store data about individual learners. This paper collects and discusses requirements, legal issues, and challenges associated with developing and using learner models, particularly in the context of European regulations. By reviewing existing standards, scientific publications, and practical use cases, we identify gaps in standardisation and propose foundational requirements for designing interoperable, legally compliant learner models. Our findings lay the groundwork for developing a reference architecture that facilitates the scalable and ethical integration of learner models in digital learning environments.**

Keywords* — *learner model; learner modelling; requirements; legal issues; compliance; ethical principles; higher education; learning analytics; general data protection regulation; GDPR; lawfulness of processing; data minimisation; learning analytics regulation.

I. INTRODUCTION

Learning is an essential process through which people acquire and continuously develop their knowledge and skills. This process takes place in both formal and informal contexts and is therefore part of our daily lives. With rapidly advancing technology, interactive content enables us to promote learning through individualised learning paths, regardless of time or place. This opens up additional opportunities for flexible and self-directed learning.

This work builds on our preliminary published conference publication [1]. It expands on this by providing a comprehensive view of digital learning environments, not only from the learner's perspective (micro level) but also from the perspectives of lecturers (*meso level*) and educational institution administrators (*macro level*). In addition, illustrative examples make the requirements more tangible and concrete, facilitating a quick understanding of this highly complex topic. Recommended strategies for dealing with the topic are then derived from this, and the appendix contains detailed working aids to simplify the introduction to collecting and processing personal data in the context of digital learning environments.

Numerous studies have analysed the factors that contribute to successful learning. All of these studies have one thing in common: the so-called flow experience [2], [3]. The flow experience describes the channel between underchallenging and overchallenging the learner. To make the learning process as successful and sustainable as possible, care should be taken to ensure that learning occurs within this channel. This channel is individual to each learner and must be considered separately.

However, it is not only in terms of the flow experience that learners differ; as a group, they are also becoming increasingly heterogeneous, as they differ in terms of individual levels of knowledge and competencies, learning styles, individual preferences for (digital) media, and various other factors [4], [5]. A potential solution to accommodate this growing heterogeneity might be digital learning environments that support users in specific situations. In digital learning environments, there are two main approaches to adjusting preferences or learning content: personalisation and adaptation. With personalisation, learners can manually adjust the system to their preferences [6], while with adaptation, the system automatically adapts to individual learners based on recognised individual needs and changing conditions [7]. The latter is exciting from the perspective of digital adaptive learning environments. This automatic support is possible on various levels. In our previous work [1], we focused exclusively on learning environments that support learners at the micro level by adapting to their specific needs in self-directed learning. This paper expands the perspective to include the meso and macro levels, so that the requirements for data collection and its digital (automatic) processing are considered not only from a small-scale student perspective but also more generally. Yet, adaptation presupposes some knowledge of the individual learner, which is usually stored in a learner model (aka user model). The latter constitutes a collection of user characteristics relevant to individual learning support [8],

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

[9]. Although a vast amount of literature exists on learner models, there seems to be no consensus concerning the content and purpose of a learner model [9] and the legal constraints that restrict the use of the information embedded in the learner model. This contribution identifies requirements that a learner model should fulfil to provide individual learning support. Requirements are derived from a comprehensive analysis of scientific publications [9] and standards on learner models [8]. In addition, we address legal issues related to the development and operation of digital learning environments that rely on learners' personal data. To do so, we take a European perspective. With this contribution, we not only want to help identify common requirements for learner models, which could be implemented in a subsequent step through a reference architecture that accounts for legal constraints, but also to lay the foundation for the development and use of digital adaptive learning environments for educational institutions.

In the remainder of this article, Section II first explains the terms used, Section III describes typical application scenarios for learning models from different perspectives (learner, lecturer, and educational institution), and Section IV discusses related work. Section V presents non-functional requirements that are mandatory, desirable or otherwise relevant for learning models and their use in digital adaptive learning environments, before Section VI compares these with legal considerations based on European laws and regulations. Section VII establishes the relationship between the technical and legal requirements and links them, on which Section VIII builds, setting out recommended guidelines for proceeding with this topic, with reference to the working aids in the appendix. Section IX summarises the article and provides an outlook for future research.

II. Definition of Terms

The literature on the science of digital learning presents different views on the same terms or their synonyms. We will therefore briefly review the most important terms to establish a common starting point. As mentioned in the introduction, there are various ways to adjust settings when learning in digital learning environments. Adaptation, often referred to as individualisation in the literature, represents the highest level of differentiation. An earlier stage is personalisation, which allows users to manually adjust the system to their preferences. In a learning context, adaptation refers to the provision of automatic, targeted support throughout the learning process, i.e., tailoring instruction to an individual's performance level and personal interests. For this automatic adaptation to succeed, the learning environment requires a digital representation of the learner, known as the learner model (*LM*).

This paper views a learner model solely as "an explicit representation of the individual characteristics and interaction data of a single learner in the form of a machine internal representation, and comprises many different aspects that ultimately depend on the purpose and goal of the application, [sort of] the digitally processed representation of the learner's characteristics" [10]. Inference mechanisms and reasoning may refer to a learner model, but are separate components rather than part of the model. Furthermore, learner models need to be distinguished from learning analytics. The term' learning analytics' (LA) was coined in an academic context by Siemens and Long in their call for papers for the 1st International Conference on Learning Analytics and Knowledge (LAK11) as "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs" [11]. Learning Analytics focuses exclusively on the learning process (including the analysis of relationships among learners, content, institutions, and teachers) to clarify the added value that big data and analytics can bring to higher education. The intended level of analysis is primarily at the course level (discourse analysis, "intelligent curriculum") and departmental level (predictive modelling, patterns of success/failure), to report – e.g., via dashboards through aggregation and descriptive statistics of the collected data – and the subsequent opportunity for optimisation [12], [13].

While the latter involves analysing group behaviour to predict individual outcomes, a learner model serves as a basis for tailoring the learning process to individual learners. Both approaches examine behavioural data and derive new insights which, once interpreted, enable the implementation of new measures. In short, learning analytics analyses learning activities, whilst learner models digitally represent learners' current status.

Open learner models constitute a notable development, as they make model content accessible to learners and other parties involved in the learning process, such as teachers or parents. Open learner models are visual representations of machine-readable formats of components of learner models [14]. By offering tools for self-reflection in different formats, learners should be supported in various ways [15]. Learning Analytics Dashboards (*LAD*) share a similar objective [16]. Unlike learner models, LADs rely on a static representation of behavioural metrics derived from interaction data, whereas learner models focus on modelling the learner's knowledge and other individual characteristics [16], [17]. Learner models and learning analytics both require personalised learner data, which is subject to the same legal principles. Although goals and ways of working differ, the requirements and legal aspects can be pretty similar. Therefore, we also consider learning analytics and examine if specific aspects also apply to learner models, possibly with adaptations.

Learning environments are (digital) platforms where learners may use different content via various learning elements to educate themselves independently. Put simply, typical digital learning environments usually consist of various components – even when subsequently combined into a monolithic whole. Standard components include, for example, the learner model, which enables individualisation in the first place, something like a domain and competence module, where the content to be learned is structured thematically and competence-oriented (comparable to a module handbook for degree programmes), and the platform with the correspondingly tagged learning elements and the possibility of adaptive learning paths. However, the

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

components can also be divided into smaller parts and cut according to function [18]. Learning environments are not defined further here, as they are sufficient for the meta-level context. It does not matter whether the learning environment is, e.g., a mobile application, an institutional continuing education platform, or even innovative learning settings with XR. Of course, digital learning environments can also be characterised and specified in greater detail, and, in some cases, even classified. For this publication, however, it is only important to note that the term "learning environments" refers to learning platforms that collect implicit and explicit data about learners to support adaptive learning recommendations and adjustments to learning elements.

III. USAGE SCENARIOS

The learning process is a highly multifaceted and complex phenomenon that educators and educational theorists have studied for many centuries. It serves as the basis for effective digital implementation in a digital learning environment.

A. Learning Process

Individual adaptation of the learning process is only viable if information and data about the learner can be used to respond to their needs and provide them with individualised support. The concept just mentioned will not scale to, e.g., entire study cohorts of degree programmes without further aid. The abundance of learner data, its pre-processing and intelligent aggregation make manual evaluation by the instructor for each individual learner almost impossible, which is why computer-aided methods are usually used. Digital recommendation systems automate the provision of individualised support to users, helping them make better decisions [19], often by building on learner models.

The typical sequence of a possible learning process in digital adaptive learning environments, in which the learner receives support through recommendations for learning elements and adjustments to their individual learning path, is defined by the *LEARN* model (**L**earning **E**nvironment for **A**daptive **R**ecommendations and **N**eeds) [18]. The learner selects a subject area in the learning environment to explore or learn about. For this subject area, the system recommends a learning element that best matches the learner's preferences, current knowledge level, and prior knowledge, based on data stored in their personal learning model. Further alternatives are presented in a ranking list from which the learner can choose if they do not like the recommendation. Once the learner has decided on a learning element, they work through it and complete it with a knowledge check – for example, a quiz or an assignment. Throughout their time in the learning environment, new interaction data is collected and fed back into the learner model, along with the learner's feedback and knowledge-check data. Based on the new information obtained, the next appropriate learning element is then recommended, leaving it up to the learner to continue or take a break for the time being [18].

Figure 1 visualises this reference workflow of the learning process within a digital learning platform. The starting point is the completion of learning and teaching activities within the digital learning environment. During the learning process, many different types of behavioural data are collected about learners and their activities. This data is pre-processed and then made persistent to enable reproducible analyses subsequently. In addition to behavioural data, other structured knowledge about a learner (such as previous education, learning type, etc.) is also stored and analysed, resulting in individual recommendations for the next learning activities and feedback on previously visited learning elements.

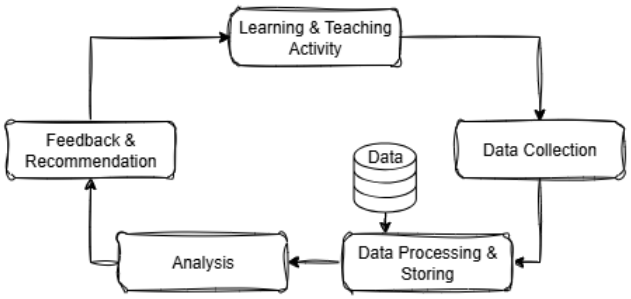


Figure 1. Workflow Reference Learning Process [1]

The basis for adaptive teaching and learning settings is the learner model, which stores all relevant data persistently and makes it available as required. Yet, the learner model should not be considered in isolation but within the overall context of the digital learning environment and its associated relationships, as the added value of learner models only comes into play when they are used in the adaptive learning platform.

B. Possible Reference Architecture

There are many approaches aimed at providing learners with personalised learning materials and pathways based on various learner characteristics, such as knowledge level or learning preferences. These approaches often differ in their underlying architectures, the techniques they use, for example, in learner modelling, but also in other design decisions and, finally, in the technologies used for implementation. However, existing research often lacks a holistic view of the learning environment, particularly regarding the system's use in real-world scenarios. The systems are evaluated and tested in a defined environment, disregarding the fact that they must meet further requirements in daily use. Simulations in protected test environments often overlook data use and the need for proper student consent for data collection. However, data protection in general and the protection of personal data in particular have become significant legal issues, particularly in Europe. Furthermore, integration into existing software systems and infrastructure – such as organisational software in a university context – is not considered. A few publications mention the components used in their systems or frameworks, with the focus usually on the recommendation component and the algorithms and techniques employed [18].

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Böck et al. have already applied the potential learning process described above to a possible reference architecture in another publication [18]. The infrastructure of the reference architecture for digital learning environments can be broadly divided into five sub-components: adaptive learning elements (managed by the learning element registry), the core, the recommender engine, and the two models: the learner model and the domain and competence model. For example, the management of learning courses or learner authentication is handled by the core component. The domain and competence model contains a machine-readable representation of the domains that can be learned using learning elements on the platform, along with their difficulty levels and associated competence levels. These cannot be viewed in isolation but are always linked to a specific domain. Together with the learner model, the domain and competence model form the basis for the recommendation engine to generate the most suitable recommendations for the next learning elements. The architecture consists of loosely coupled, highly coherent modules that enable flexible communication, ensuring efficient processing and high user acceptance. This structure facilitates the easy integration of new systems and the addition of interfaces. As the system is scalable, affected components can handle increased user volumes as required, and updates to individual modules can be carried out independently without affecting the overall system. The architecture is generalised, so that a wide variety of domain and competence models are supported in compatible formats.

C. *Affected Levels*

Learner models and adaptation do matter at different levels: from support in planning a degree program [macro level] (which modules are interesting for me?), to learning patterns of student cohorts during their studies [meso level] (which topics prove difficult within certain student cohorts?), to which concepts within a module [micro level]. We will focus mainly on the micro level. However, we will also consider the meso and macro levels to a large extent if they have similar or even identical requirements and differ only in their objectives.

However, Figure 1 can be applied not only to the learner perspective (micro level) already mentioned, but also to the meso and macro levels. At the meso level, individual learners can be grouped into a cohort, providing the lecturer with information about a specific learning element or topic. After many learners have completed learning elements, aggregation is added in the 'Data Processing & Storing' step to summarise data from multiple learners in the same learning area. In the analysis step, cohort analyses are carried out, for example, to identify interaction patterns and challenges or learning difficulties for the entire group of learners in specific sub-areas or learning elements, and to counteract these with measures. This information is reflected to the lecturer in the 'Feedback & Recommendation' step, including tips and possible measures to address these challenges, for example, by adapting the content or providing additional learning elements. It is not a single learner that is of interest here, but the cohort as a whole in pseudonymised or even partially anonymised data for the lecturer.

At the macro level, it is viewed from a more distant meta-level. This involves activities across courses and degree programmes, so platform-wide use is relevant here. In the 'data collection' phase, aggregated course data is collected, including completion and dropout rates, as well as accessibility and usage metrics. In the next phase, this collected data is stored long-term and linked to institutional data, such as administrative data (e.g., applications, enrolments, or other data). The analyses aim to simplify strategic decisions. Therefore, comparative analyses of different courses or study programmes, as well as longitudinal trend analyses, are carried out to identify systemic patterns and derive measures. This then happens in the 'Feedback & Recommendation' phase, in which strategic institutional recommendations for the management of the educational institution or institute are derived from the analysis results, for example, to optimise curricula and degree programmes, as well as to make improvements at the platform level.

The following section presents examples of typical scenarios from digital teaching to illustrate the specific application purpose. The necessary measures and requirements can then be derived subsequently. It is normal for several actors to be involved in a learning process. In most cases, however, the selection in adult education can be reduced to three relevant actors: the student (referred to as a learner from now on), the teacher, and the educational institution that provides the formal setting for the education. These three actors also form the three levels of adaptive learning described above. For the sake of completeness and better understanding, all three are briefly outlined here, although the focus remains on the micro level and thus on the learner. However, the resulting synergies for the other two actors and related levels in the learning process are also highlighted and considered.

D. *Purpose of Use*

Several publications, e.g., ISO/IEC JTC1/SC36 [19], [21], [22], already present typical scenarios of digital learning at different levels. Teaching and learning activities in digital learning environments are the starting point for analyses of digital learning. Learning analytics are used to personalise and thus improve the learning environment adaptively, and to summarise methods such as learner models and learning analytics. Based on the results of the learning analysis, the learning environment can recommend individualised learning paths alongside distinct learning content. Three main actors are involved in typical learning scenarios: the learner, the instructor, and the educational institution.

Learner (*micro level*)**.** In the past, log entries from the learning environment were difficult for non-technical users to understand, if they could be viewed at all. Nowadays, learners can view visual displays of their data in dashboards, e.g., to monitor their learning progress relative to the cohort average. The early recognition of learners' personal needs and preferences (predictive analytics), including possible

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

performance deficits, and the resulting initiation of preventive remedial measures (timely interventions) increase the effectiveness of the learning process. The use of such learning analyses also supports learners with disabilities and helps identify accessibility deficits in learning opportunities.

Instructor (*meso level*)**.** Instructors can track the activities of their entire cohorts and detect progress and potential problem patterns, e.g., a lack of learner engagement early in the course. Consequently, instructors may initiate appropriate support through adaptive teacher response to observed learner needs and behaviour.

Educational Institution (*macro level*)**.** The educational institution can contribute to an improved holistic individual education strategy by enabling analyses, which can also reduce dropout rates in the long term. In addition, administration may benefit from information on the entire learner population, e.g., for future course planning. If comparisons of the data set across learning modules are permitted, similar patterns can be identified and indications derived, such as a potential accessibility problem.

Table I below summarises the various levels, stakeholders, data types, and their granularity to provide a quick overview. These three levels, along with the three stakeholders, run throughout the publication.

TABLE I. OVERVIEW OF LEVELS IN RELATION TO THE STAKEHOLDERS

Levels	Stakeholders	Type of data (granularity)
Micro	Learner	Personal data (in as much detail as possible)
Meso	Instructor	Pseudonymised data (standard)
Macro	Educational Institution	Anonymised data (abstracted)

E. Legal Principles and Classification

The implementation of learner models in educational institutions has various legal implications. For example, intellectual property is protected from the developer's perspective. On the other hand, learners' rights, especially their fundamental rights, must be taken into account. In the EU, the fundamental right to data protection, as stated in Article 8 of the EU Charter of Fundamental Rights (*CFR*) [20], is affected. The specific legal implementation of this fundamental right, the General Data Protection Regulation (*GDPR*) [21], is the primary law to be considered when using learner models. These provisions are considered when determining legal requirements.

IV. RELATED WORK

Learning offers must be modularised to enable the creation of individual learning paths based on a wide range of criteria. Many publications, particularly in the last decade [9], deal with adaptive teaching and learning settings. For this work, scientific publications on user models, particularly learner models, are the most relevant. Koch [22], [23] summarised the seven objectives of a user model: (1) supporting users in learning a specific topic; (2) providing users with customised information; (3) adapting the user interface to the user; (4) supporting users in searching for information; (5) providing users with feedback on their knowledge; (6) supporting collaborative work; (7) supporting the system's use.

Numerous publications address learner analyses, but mostly only as a means to the end of adaptive learning systems [27], [28], or only very superficially, without going into relevant details. Several publications also consider only individual information rather than aggregating it into a learner model [24], [25]. A couple of surveys attempt to create an overview of the field [26], [27], [28], [29], but mostly only compare the different models with each other, without providing direct insight into which learner characteristics can be used, how they are modelled, or where the required data comes from.

Therefore, we conducted two parallel systematic literature searches on learner models: on the one hand, to obtain a comprehensive picture of the current state of scientific research [9] and, on the other hand, to clarify the state of practice based on standards and norms [8]. We decided to conduct a systematic literature review (SLR) following the methodology described by Kitchenham and Charters [30], [31]. Ideally, this should pave the way for a standardised approach to designing and creating learner models, as well as a set of standardised subcomponents – regardless of their purpose or use. 868 standards were reviewed, 16 of which were classified as relevant to the structure and components of learner models. Three standards focus extensively on learner models. Among those, ISO/IEC 29140:2021 [19] describes a mobile learner model that considers specific attributes, such as the device used, connectivity, and the learner's location, to better describe the learning environment. The 1EdTech consortium [32] focuses on interoperability among internet-based information systems that support learners' interaction with other systems. It uses a data model that captures the essential characteristics of learners to monitor and manage their progress, goals, performance, and learning experiences. IEEE P1484.2 [33] attempted to specify the syntax and semantics of a learner model by centralising public and private learner information, which became known as PAPI Learner. This endeavour started in the 1990s, but was not pursued further. Extensive research into norms and standards in the field of learner models has shown that approaches in this area are rare and are either specialised [20], complex and more than just learner models [37], or abandoned [33]. No standards have been found that describe a generalised, realisable learner model that can be extended for specific use cases, or that address the creation of such learner models [8].

In parallel, scientific publications from relevant publishers (IEEE, ACM, Elsevier, pedocs) from 2014 to 2023 were systematically examined, yielding 197 papers deemed sufficiently relevant for detailed analysis [9]. Scientific publications on the design and development of learner models reveal a variety of different approaches. Many models integrate characteristics such as learning style, but do not specify why or how these characteristics influence learning, or whether it makes sense to take them into

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

account. There is also a lack of clear recommendations as to which combinations of characteristics should be included in a learner model. Standardised characteristics, such as demographic data (e.g., name, gender, age), are often used together with behavioural and learning data. The modelling of this data varies greatly in the literature. Knowledge characteristics can be modelled using, e.g., overlay models or fuzzy logic. This diversity makes it hard for developers to make decisions and implement the models. In addition, publications often lack details on data sources, data processing, and practical implementation of model components. Many of these topics are only touched upon in passing or omitted altogether, which makes it difficult to replicate learner models precisely [9].

The two systematic literature searches indicate that the descriptions of learner models are often rather superficial and lack detail. Many explanations refer to frequently used standards and then expand these to include individual aspects [34], [35], [36], [25], [37], but do not describe in detail how these are expanded or what the implementation or modelling of the learner model looks like. In summary, there are no ready-made, predefined learner models that can be downloaded and customised. Furthermore, the individual learner models described in the two literature reviews provide only very limited information regarding their implementation details.

Legal implications of learner models, specifically, have not yet been discussed. However, there is research regarding the requirements of the European data protection law in learning analytics, which can also be used for learner models [38], [39], [40], [41], [42], [43]. Our main objective is to contribute to this standardisation process by proposing an original reference model as a first step towards a reference architecture. This contribution is based on defining open software interfaces for each subsystem in the architecture, thereby avoiding any dependence on specific information models. We have already discussed elsewhere a possible reference architecture for an adaptive learning platform that integrates a learner model [44].

V. REQUIREMENTS

Based on the previously analysed publications and the resulting findings on possible components of learner models and their use in digital learning environments, we derive a list of requirements for a learner model and its application. The development of the requirements specification is based on a qualitative analysis. In separate publications, we have already examined the functional requirements for learner models in greater detail [44] and provided initial insights into the legal classification of these requirements [1]. We have also focused on the broader context of digital learning environments and the integration of learner models [18]. Based on the literature review, the following section will focus on the fundamental functional requirements, the non-functional requirements, and the legal basis for the use of learner models in Europe, especially in Germany. The following section is by no means comprehensive; rather, it collects and summarises the most relevant requirements, based on their occurrence and description in the literature, and links them to current legal restrictions in Europe, especially in Germany.

In addition to the analysed publications, the feedback and experiences of learners at our university play an essential role. It is important to emphasise again that learner models are structured models without any interference mechanisms [10]. However, it is also necessary to evaluate the effects of integrating a learner model into the overall eLearning environment, given the requirements for a learner model.

The following requirements are classified according to the *Kano model* [45] into so-called *must-be requirements*, one-dimensional requirements (*should characteristics*) and attractive requirements (*can characteristics*). The requirements are supplemented with classic examples to facilitate understanding and, where possible, to address any challenges or difficulties that may arise during implementation. The requirements were classified, on the one hand, based on the frequency of mentions in the publications analysed during the systematic literature reviews and, on the other hand, through the authors' assessment of which requirements are essential for operation, which requirements are desirable enhancements to the basic functionality, and which requirements would certainly be useful but are not strictly necessary for the implementation of a functional learner model. These two metrics were consistent with one another. This classification was validated by the authors' experience during the iterative development of their own learner model, as well as by feedback from reviewers of previously published papers and from discussions at conferences attended by experts in these fields.

A. Must-be Requirements

Must-be requirements are essential for the successful use of a learner model in future scenarios and form the basis for its design and subsequent development.

Collection and Management of Learning Data. The collection and aggregation of static and dynamic information about the learner (e.g., knowledge, skills, interests, preferences, behavioural data) is the initial step. The utilisation of data from various sources plays a significant role. In the second step, the collected data must be digitally processed, modelled, and persisted. While data about the learner is collected and stored at the micro level, such as (successfully) completed learning elements, behavioural data (in the learning environment, such as interaction data) and knowledge data (e.g., last degree, modules already taken, ...), this information and data are relevant to lecturers at the meso level in aggregated and pseudonymised or anonymised form. The same applies at the macro level, where the data in its entirety serves as the basis for all analyses. The challenge in collecting and persisting the data lies in its granularity and in selecting the data relevant to learning.

Legal Data Protection Requirements. The development and use of learner models raise a range of legal considerations. In particular, rights relating to data ownership under the EU Data Act may become relevant. Moreover, if learner models fall within the definition of artificial intelligence systems under the AI Act, additional

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

restrictions may apply. This is especially the case when such systems are classified as high-risk (Annex III No. 3(b)/(c) AI Act [46]), which would trigger extensive compliance obligations. However, the following section focuses on the most immediately relevant legal framework: The General Data Protection Regulation (*GDPR*). All data to be collected must at all times comply with the applicable legal requirements of the *GDPR* [21], [47]. This requirement is closely aligned with the legal constraints of designing and creating learner models. What are the applicable legal requirements for data collection, processing and storage? These questions cannot be generalised or answered in general terms; they must be legally considered and examined individually, depending on the purpose of use and the data to be collected. These questions also apply to data aggregation, i.e., the application of the learner model at the meso and macro levels. The challenge here is to check, or have it checked manually, the legal aspects of each application and the required data, to determine whether the project can be implemented legally as planned or what conditions apply. The concept of a data trust model might be applicable here: a neutral, trustworthy entity ensures that data is processed and used in accordance with defined data protection guidelines and is accessible only to authorised parties.

B. Should-be Requirements

Should-be requirements are important prerequisites for making optimal use of the learner model in future scenarios. They represent desirable but not mandatory features that increase the efficiency and flexibility of learner models. These requirements serve as an orientation for further development and improvement of the model.

Transparency / Traceability. The transparency in how the system reaches a result [53] creates trust and acceptance of certain subordinate recommendations, which form the basis of good support [48]. Therefore, learner models should be transparent at every step of the process, i.e., from the origin of the data to the individual processing steps the data went through, and what consequences this has for the result and its explanation. This is important to illuminate backgrounds and issues, such as potential discrimination [54], [55], [56], in categorisations. Transparency thereby relies on individual decisions so that they can be traced correctly. Traceability is achieved by involving the learner from the outset and by adopting a learner-centred approach to the design and development of learner models. This helps learners better understand how a decision is made by showing the individual steps involved, the data used, and the information that led to each decision. For example, in a subject area, repeated failures on knowledge tests – such as quizzes, exercises, or similar – could indicate that the learner has not yet internalised the material and therefore needs to repeat it. This information should then be made visible to the learner to improve self-reflection. The difficulty here, however, is that with many machine learning methods, the decisions made do not always allow conclusions to be drawn about the reasons for those decisions, as would be the case with a decision tree, whereas a neural network does. Technically speaking, there are already some approaches: Explainable Artificial Intelligence (XAI) [49], [50] encompasses methods and techniques that make the decisions of machine learning and artificial intelligence systems comprehensible and interpretable to humans [51], [52]. This transparency is also important at the meso and macro levels to deepen understanding among lecturers and educational institution management, and to initiate meaningful measures, above all.

Responsibility. Traceability is closely connected to responsibility. The data that may be collected must be left to users' choice, subject to applicable legal standards. That is, learners can decide which data to log and persist. This topic also includes compliance with ethical principles. In the digital domain, this means that if the learning management system demonstrates ethical behaviour, learners expect it to comply with ethical guidelines and principles. Briefly summarised, ethics is the view of moral values and their conception (based on Aristotle's *Nicomachean Ethics* and Immanuel Kant's *Categorical Imperative*). Specifically, from the learner's perspective, this means at the micro level that they alone want to decide what data they want to make available to the learning environment and that they can view and change this at any time, for example, whether, in addition to their learning history, the grades of modules and examinations already taken during their studies should and may be used. The system should then comply with this decision and not 'secretly' collect and use the data, as this would demonstrably weaken trust in the learning environment. Considering the other two levels, the responsibility lies in ensuring that this highly personal data is used only by the learner and cannot be viewed, let alone used, by third parties. This means that at the meso and macro levels, the data must not allow any conclusions to be drawn about individual learners, but should provide information about aggregated states, such as: ~80% of course participants have difficulties with the topic' complexity theory' in the learning element 'P-NP problems and NP completeness' (#0815), or the module' theoretical computer science' is very rarely taken voluntarily as an elective module, with only ~3% of learners choosing it.

Fairness & Ethics. Fairness affects many different steps within the learning process. For example, algorithms for decision-making and data processing must be checked for possible biases to ensure equal opportunities for all learners [53], [54].

Tamper-proof. The data managed and persisted by learner models must be protected from unauthorised intervention and changes. Any changes need to be logged in a traceable manner. This implies two interlinked requirements: that unauthorised data access must be prohibited, and that, if access is permitted, any changes to data shall be checked for plausibility and logged for tracking purposes. This requirement is closely linked to traceability. Learners must be able to be one hundred per cent certain that their data will only be used for the purposes for which they have given their consent, i.e., to assist them in their learning and in reaching their learning goals, and that this personal data cannot be viewed or altered by any unauthorised third party. At the meso level, for example, the lecturer should not

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

be able to identify individual learners based on their learning difficulties, but should only be able to obtain an overall picture of the cohort. The same applies at the macro level, of course. Any inference about individual learners must be prevented by all possible means. The only exception is the learner's prior consent, if they require external assistance, to temporarily disclose a selection of their data to specific third parties.

C. *Can Requirements*

Could-be requirements represent possible extensions that might optimise the learner model in certain scenarios, e.g., by providing additional benefits or improving user experience. If necessary, these requirements can be incorporated into future development phases to increase the model's flexibility and adaptability.

Openness & Visualisation. Learner models should be open [15], [55] to promote metacognitive behaviours, such as self-awareness and self-regulation. This means that learners may display and analyse their own instances of a learner model to better understand their learning progress and, e.g., misconceptions [14]. In this way, learners can see their current learning status and progress in any area at any time, compare it with their learning goals, and derive follow-up activities. Various visualisation options [56], [57], [58] are essential for presenting complex learning data to the learner in a clear and concise format. Different representations of the same data can bring learners closer to the various aspects and clarify them [59]. Leaner models designed in this way are called open learner models [60], [61]. The difficulty in opening the learner model is choosing the 'right' visualisation method, which depends on the available data. This is because the visualisation should not overwhelm the learner, but should provide further insights into their learning behaviour, ideally in a self-explanatory way. There are numerous visualisation options for the same data. For example, the knowledge learned can be represented using simple bar or column charts, tables, or radar plots (see Figure 2 for conceptual examples), to name just a few visualisation options [62], [63].

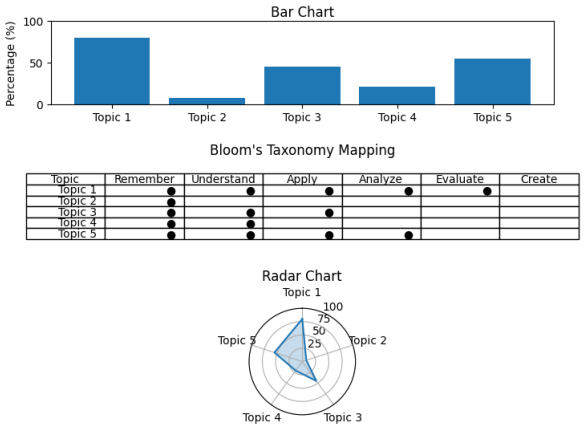


Figure 2. A few random examples of visualisation options

The same applies to the meso and macro levels. Visualisation options should provide both teachers and educational institutions with simpler, more easily understandable insights into the data and its interrelationships. For example, the lecturer could see, in pseudonymised form, where learners might be having problems (see Figure 3 as a conceptual example), or the management of the educational institution could see which subjects are popular based on enrolment figures over the last few years (see Figure 4 as a conceptual example).

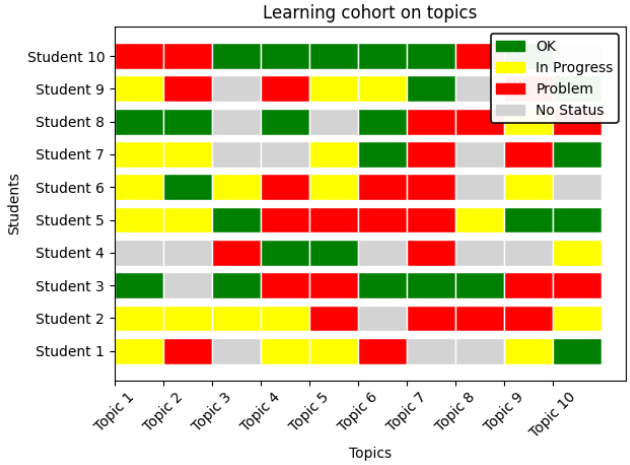


Figure 3. Possible view of a lecturer

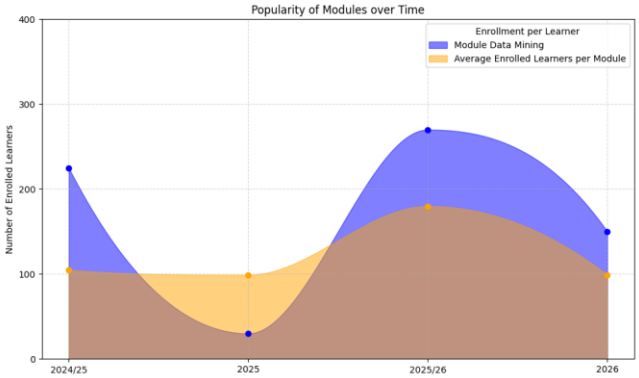


Figure 4. Possible view for the administration of the educational institution

Negotiation Options. In addition to the pure visualisation of personal data, some learner models also offer the option to interact with this data and, for example, negotiate with the model if the data shows inadequacies from the learner's perspective [14], [63]. In this way, (open) learner models give learners not only responsibility for their individual data and progress but also human control over their personal data. This requirement is particularly relevant at the micro level, as learners are allowed to correct data they believe is incorrect. For example, if the learner's current level of knowledge does not align with their self-reflection and they believe they are better at a particular topic, they may be

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

able to update this information. There are various approaches to adapting information, which can often be divided into four main types (*according to* [64]):

- inspectable, i.e., the learner can view some or all of their stored data using various visualisation options.
- co-operative, i.e., depending on how easily everyone can fulfil the respective modelling tasks, these are divided up.
- editable, i.e., the learner can change the content as desired.
- negotiated, i.e., the learner and the system discuss the stored contents of the model and agree on a joint solution.

Data Minimisation & Sustainability. Only data that is absolutely necessary for adaptation should be collected. Minimising the amount of collected data while maximising the value of the information avoids unnecessary strain on the infrastructure. In addition to the purely legal considerations of data economy and minimisation, unnecessary data places a strain on the existing infrastructure through the collection, transmission and storage of information. Of course, the initial challenge is to identify which data are relevant for learning. It can support causal inference, but once these questions have been largely clarified, the collection and retention of data that is not relevant for learning should be discontinued.

Modularity & Flexibility. A largely self-contained modular structure of the component of earner models enables flexibility, for example, to swap the technical infrastructure (learning management system) or to export the learner model and integrate and use it in other environments, for example, if the learner changes university after graduation. This requirement is independent of levels, as it concerns the overall system rather than the learner or the teacher; in the broadest sense, it also concerns the educational institution, as it involves a technically directional architectural decision and thus becomes less dependent on individual components.

Maintainability & Expandability. The learner model should have a modular structure (see previous requirement) to ease future adjustments or extensions. This offers the learner model a certain degree of secure prospects. Like the previous requirement, this one is also level-independent, as it concerns the overall system.

Standardised / Interoperability & Integration. Ideally, learner information is standardised so it can be easily exported and exchanged across different learning platforms (in line with the interoperability defined in [65], [66]) through standardised interfaces for data exchange and integration. Standard conformity is essential here, i.e., open standards to ensure compatibility with other systems should be supported. Standardisation also enables cross-platform integration, i.e., the model can be seamlessly integrated into existing learning management systems. As with data minimisation (Art. 5 lit. c GDPR), data portability is also a legal requirement (Art. 20 GDPR). The challenge clearly lies in providing a possible standard for so-called learner models. Education providers will have to continue addressing this issue over the next few years to arrive at a commonly recognised and accepted standard. However, aside from the challenge, the advantages of standardisation are enormous. For example, when changing educational institutions, learner data can be easily exported and reimported with minimal effort (at the macro level). From the learner's perspective (micro level), learning can continue directly where it left off, without having to start again from the beginning, and the current level of knowledge in various subject areas can be determined through pre-tests or learning preferences assessed via a questionnaire.

One option worth considering is to transfer the learner model topic to a so-called data trust model and treat it as an external entity that can be integrated and used with the learner's permission.

VI. LEGAL CONSTRAINTS

Legal considerations are crucial in the design and development of learner models, as lawful use is only possible if these models comply with applicable regulations. It is important to emphasise that the following legal analysis and the resulting practical recommendations are strictly oriented towards the European Union's regulatory landscape. Therefore, while the technical requirements may be universal, the legal compliance strategies discussed here have limited transferability outside the EU jurisdiction.

. Most data processed in learner models falls within the material scope of the GDPR pursuant to Art. 2(1) GDPR, as it constitutes personal data within the meaning of Art. 4 No. 1 GDPR. Personal data encompasses any information relating to an identified or identifiable natural person, including not only basic information such as names or email addresses, but also learning behaviour, performance metrics, and other characteristics that allow conclusions to be drawn about individual learners. The feedback from learner models also contains information about the learners, making it personal data.

Personal data plays a crucial role in enabling adaptive customisation of learning processes. Therefore, all data collected must always comply with the relevant legal requirements (data protection). However, data protection also encompasses other aspects. On the one hand, only data that is strictly necessary for meaningful adaptation should be collected (data minimisation & data economy). On the other hand, data should be collected only if there is a legal basis (lawfulness). All the aforementioned legal principles are enshrined in the General Data Protection Regulation (*GDPR* [47]). The collection and aggregation of static and dynamic information about the learner (e.g., knowledge, skills, interests, preferences) must be compliant with the GDPR. The utilisation of data from various sources plays a significant role. In the second step, this collected data must be digitally processed, modelled and persisted. Notably, the GDPR is technology-neutral (recital 15 GDPR), requiring the broad terms it employs to be defined explicitly in the context of learner models.

To ensure analytical precision throughout the subsequent legal evaluation, a fundamental distinction must be drawn between learner models and learning analytics, as their technical characteristics dictate varying levels of data protection risks. In their standard implementations, both

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

applications inherently process personal data within the meaning of Art. 4 No. 1 GDPR. However, a significant divergence exists regarding the feasibility of anonymisation: because it focuses on cohort-level patterns and aggregated behaviours (meso and macro levels), learning analytics often enables complete anonymisation. Once data is irreversibly anonymised, the personal reference is severed, rendering the GDPR entirely inapplicable (Recital 26 GDPR). Conversely, learner models are structurally tied to the individual (micro-level) to facilitate tailored pedagogical adjustments, making pseudonymization the primary safeguard while the data remain personal. Whenever anonymisation cannot be achieved, the full spectrum of GDPR obligations applies equally to both applications. Consequently, the requirements developed in this framework are fundamentally relevant to both domains. Ultimately, the precise operationalisation of these legal obligations remains a case-by-case determination, as the specific compliance risks heavily depend on the concrete system architecture and data categories involved. In general, however, learner models present a higher intensity of data processing, as they continuously track and model the nuanced learning status, competencies, and developmental trajectories of individual natural persons.

The GDPR's requirements are diverse, and due to the limited availability of case law and literature regarding the GDPR in relation to learner models, at the time of writing this publication, these legal obligations cannot yet be determined with a high degree of certainty. It is highly recommended that the local data protection officer be integrated into the process of designing and implementing learner models in educational institutions as soon as possible. Many obligations of the GDPR, for example, the data protection impact assessment in Art. 35 GDPR will be very hard to meet without professional support. This chapter aims to identify potential issues arising under the GDPR and highlight key considerations when designing and implementing a learner model. In providing an overview, we focus on the principles of the GDPR, which are concretised in the GDPR, and the arising challenges regarding learner models.

Lawfulness - Legal Bases. Any processing of personal data requires a legal base, as stated in Art. 5(1), Art. 6(1) GDPR. Art. 6(1)(a) GDPR establishes consent as a legal basis, which must be given voluntarily [67, p. 330]. In hierarchical contexts, such as teaching environments, the GDPR requires a strict interpretation. Learners often depend on lecturers for grading, which may pressure them to agree to the processing of their personal data to align with lecturers' expectations. However, this power asymmetry is not present in all educational scenarios. If the use of a learner model is a voluntary additional offer of the educational institution and is not linked to a specific course, voluntary consent seems possible. For state educational institutions, instead, Art. 6 (1)(e) GDPR serves as the legal basis for data processing when processing is necessary for the performance of a task carried out in the public interest, such as education. However, under Art. 6(3) GDPR, the legal basis of Art. 6 (1)(e) GDPR must be complemented by specific legal provisions by the member states that establish obligations and define tasks. Consequently, instructors must ensure that the applicable legal bases in their national laws include the processing of personal data for educational purposes. General requirements imposed by the GDPR on national law are discussed in [68].

Furthermore, the type and extent of data being processed should be carefully evaluated. On the one hand, this enables an assessment of the risks associated with potential data breaches. On the other hand, it highlights whether special categories of personal data are being processed in the learner model. Processing personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, biometric data for uniquely identifying individuals, health data, or information about a person's sex life or sexual orientation is subject to stringent requirements. Under Art. 9(2)(g) GDPR, alongside the requirements of Art. 6 (1)(e) GDPR, a legal basis in Union or Member State law is required, allowing such processing only if it is necessary for reasons of substantial public interest. Meeting this requirement is particularly challenging for learner models.

For learners, this means that, at the micro level, the learning environment is an additional learning opportunity that is often not mandatory but that supports the learner individually in their learning process. To this end, learners may revoke their consent at any time to the use of their personal data for individual learning support without incurring any disadvantages to their grades or final assessment by the lecturer or educational institution. This additional service is purely voluntary for learners who wish to receive it. Looking at the data to be collected in order to enable the most successful individual adaptation possible, this personal data is divided into two categories: personal data and sensitive personal data, Art. 9 GDPR. The latter includes data relating to health, religious, or philosophical beliefs. However, this data may be relevant to learning. For example, information about the learner's health can help tailor the programme to their needs, such as allowing more breaks if they have been diagnosed with attention deficit disorder, or omitting spelling from the assessment if they have been diagnosed with dyslexia. The same applies to religious data, for example. If this data influences learning behaviour, for instance, because prayer times interrupt active learning, this should be considered by the system, which means that the necessary data is also required. This data should be collected and processed with absolute caution only after explicit notification and, ideally, with the learner's explicit consent (Art. 9(2)(a) GDPR), if necessary to achieve the goal of individually tailored learning support. From a meso- or macro-level perspective, this issue is somewhat simplified when it concerns anonymised, aggregated data at most, which should not allow any conclusions to be drawn about individuals.

Purpose limitation. Developers of learner models must carefully evaluate the potential sources of personal data used for their development and operation. Often, data is collected for a different purpose — for example, when learners' exams, initially collected for grading, are fed into learner models. In addition, the subsequent use of data generated by

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

the learner model should be carefully evaluated. Art. 5(1)(b) GDPR requires that personal data be collected for specific, explicit, and legitimate purposes and not further processed in ways incompatible with those purposes. However, Art. 6(4) GDPR provides an exception: if the new purpose for processing personal data is not based on the data subject's consent or authorised by Union or Member State law, its compatibility with the original purpose must be assessed using the criteria outlined in Art. 6(4)(a)–(e) GDPR. Scientific research, which is what learner models could be part of, is privileged. Art. 5(1)(b) assumes scientific research is "not be considered to be incompatible with the initial purposes".

This requirement becomes especially relevant when repurposing existing learner data that was initially collected for other purposes, such as administrative processes, examination grading, or assessment of learner work. As this data is often ideal for populating and utilising learner models, an evaluation is required to determine whether such a change of purpose falls within the privilege afforded to scientific research.

Transparency. Transparency is not merely a technical requirement but is also enshrined in Art. 5(1)(a) GDPR. Personal data shall be processed "in a transparent manner in relation to the data subject […]." This principle is guaranteed by primary law in Art. 8(2)(2) CFR by granting every person "the right of access to data which has been collected concerning him or her […]." The requirement of transparency is primarily detailed in Art. 12 et seq. GDPR. It stipulates that if personal data is collected, the controller must, pursuant to Art. 13 GDPR, inform the data subject about the details outlined in Art. 13(1) and (2) GDPR at the time of collection. This information must be provided in a privacy statement. It is recommended to explain the system in clear and accessible language in the privacy statement to help learners understand how the learner model functions.

When learners exercise their right of access to their personal data, it must be ensured that all processing of personal data is captured. This encompasses registration data (such as name, learner ID, and email address) used for login, metadata related to system usage, and, most importantly, all input data entered by learners into the system. Additionally, this includes the inferences the system draws from such data, as well as the feedback it provides.

Data Minimisation. Within learner models, various sensitive data will be processed, which is in contrast to the GDPR principle of data minimisation, which allows the processing of personal data only when strictly necessary. The principle does not require minimising the data itself but seeks to limit the connection of data to natural persons, thereby reducing infringements of fundamental rights [69] mn. 96. To comply with this principle and the requirements of Data Protection by Design (Art. 25 GDPR), institutions are legally obligated to implement appropriate technical and organisational measures. While the GDPR does not dictate one specific technical solution, pseudonymisation, as mentioned in Art. 25(1) GDPR is strongly recommended as a primary safeguard to minimise risks. According to Art. 4(5) GDPR, pseudonymisation is the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information. *EDPD*, the European Data Protection Board, provides a detailed guideline for pseudonymisation [70]. As an organisational matter, a role concept should be implemented that specifies who can access the data collected in the learner model and who can edit it. For complete guidance on technical and organisational measures according to Art. 25, see [71]. In particular, the scope of people who can undo pseudonymisation needs to be kept small.

Also, the use of a data trustee can be discussed [18]. A data trustee is a neutral third party that acts as a steward for sensitive data, ensuring its secure handling, responsible use, and protection of individuals' privacy while facilitating data-driven innovation. Data to be collected and processed must be carefully selected, i.e. only for the purpose of optimising the learning process – in other words, adaptive learning. Data that is not required in this context may not be collected or processed further. Pseudonymisation is typically feasible at the meso level. At the macro level, the identity of individual learners becomes irrelevant, allowing personal information to be minimised or even completely removed through anonymisation.

Storage Limitation. The principle of storage limitation, as outlined in Art. 5(1)(e) GDPR requires that personal data be retained in a form permitting the identification of data subjects only for as long as necessary to achieve the purposes for which it is processed. Temporal storage limitation is a subset of the overarching principle of necessity [69] mn. 122. Educational institutions and lecturers must determine the appropriate retention period for the data. Developers of learning models must ensure that the complete deletion of learners' personal data is technically feasible. Typically, learners' personal data should be deleted once they leave the educational institution. If the data is intended to be retained for further optimisation of the learner models, the link between the data and the individual learners can be erased entirely and irreversibly (anonymisation) [72]. At the micro level, the data collected and processed should be periodically reviewed to determine whether it is still needed or can be reduced to the bare minimum. Nothing further needs to be done at the meso and macro levels, if only pseudonymised or even anonymised data is available at any given time, making it impossible to identify individual learners.

Integrity & Confidentiality. Art. 5(1)(f) GDPR states that personal data must be "processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction, or damage, using appropriate technical or organisational measures." This principle underscores the importance of systemic data protection. It is primarily implemented through Art. 25 GDPR, "Data protection by design," and Art. 32 GDPR, "Security of processing". Whenever possible, data should be stored in an encrypted format on a trustworthy server, ideally one operated by the educational institution. The use of servers in an EU member state is unobjectionable, as the GDPR establishes a uniform standard for data protection across member states. The use of servers outside the EU is

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

possible; however, data transfers to such servers must comply with Art. 45 et seq. GDPR, which aims to ensure continuity of the level of data protection [73] mn. 6.

Use of Processors. If an external service provider is engaged in processing personal data on behalf of an institution, the requirements of Art. 28 GDPR must be fulfilled. Processors are required to provide sufficient guarantees that they will implement appropriate technical and organisational measures to ensure compliance with the GDPR. These requirements must be formalised in a contract between the educational institution and the processor, so-called Data Processing Agreement (*DPA*). Whenever possible, the processor should store data on servers located within the EU.

No automatic Decision-Making with a Legal Impact. Once the learning model can analyse learners' input, it could be used for automated assessment, for example. However, it is important to emphasise that the learner model itself does not make any evaluations or decisions; it functions strictly as a structural data representation. Consequently, any legal restrictions or prohibitions under Art. 22(1) GDPR apply solely to the downstream system or processing logic that operates upon this data. Art. 22(1) GDPR states that decisions based solely on automated processing, including profiling, which produce legal effects, are prohibited. Art. 22(2) GDPR provides exemptions. Still, these are unlikely to apply to the use of learner models unless a member state establishes a legal basis explicitly permitting automated decision-making and defines suitable measures to safeguard the data subject's rights, freedoms, and legitimate interests. At least for Germany, no such legal basis is known at the time of writing this publication.

VII. The Relationship between Technical and Legal Requirements

Many requirements from technical and legal perspectives are quite similar, although these requirements should be viewed from both angles. Transparency serves as a key example. From a technical perspective, it is essential to design the learner model as transparent and traceable as possible to build trust in the system. Therefore, it might make sense to limit the displayed data, such as learning progress, to prevent information overload. The transparency required under Art. 5(1)(a) GDPR and Art. 12 et seq. GDPR, on the other hand, aim to give the data subject an extensive overview of all their processed data. Thus, a progress bar showing learners' learning progress may fulfil design transparency requirements, but it does not satisfy GDPR transparency obligations. Similar analytical tensions emerge across other core requirements. Regarding openness and negotiability, pedagogical design often allows learners to manually alter their skill levels to foster self-reflection and motivation. However, the legal principle of accuracy (Art. 5(1)(d) GDPR) and the right to rectification (Art. 16 GDPR) demand that institutional data objectively reflect reality, especially when tracking prerequisites. Consequently, to maintain data integrity, a system must store both values separately if a learner overrides a calculated skill state. Furthermore, interoperability reveals a conflict between de jure rights and de facto technical fragmentation. While Art. 20 GDPR grants learners the right to data portability in a machine-readable format, the lack of widely adopted, semantically unified learner-modelling standards means that mere technical transferability does not guarantee pedagogical utility in a new learning management system. Finally, ensuring fairness and ethics highlights an inherent 'accessibility paradox'. To guarantee technical fairness and automatically provide accessible learning materials (e.g., tailored formatting for learners with dyslexia or ADHD), the system must inherently process highly sensitive health data. This immediately triggers the stringent legal restrictions of Art. 9 GDPR, demonstrating that pedagogical adaptivity is frequently constrained by necessary legal privacy safeguards. Overall, this demonstrates that while technological opportunities and data protection regulations can mutually reinforce each other, they often rely on divergent interpretations or result in inherent conflicts of interest.

The learner model must also be evaluated from an ethical perspective, which may not be fully represented in design and legal considerations. It might be legal to collect various personal information about learners over several semesters, but it remains questionable from an ethical perspective. Many higher education institutions have their own ethics commissions, which often need to agree to the use of new technologies for learners. Therefore, both the internal ethics committee and the legal department, together with the data protection officer and, of course, the learners, should be involved in the process from the outset.

To make it easier to understand, we have summarised the technical requirements—both functional and non-functional—in Table II below. We have categorised these requirements by classification and affected level and, where there are legal implications, included them as well; this serves as a synthesis of the two sections described in detail above.

TABLE II. Overview of levels in relation to the stakeholders

Levels	Requirement	Primary Legal Principles	Technical Implications
Micro, Meso, Macro	Collection and Management of Learning Data	Art. 5(1)(c), Art. 5(1)(b) GDPR	To work with data, it must first be collected and stored. In the case of personal data, this always requires the data subject's consent.
Micro	Transparency / Traceability	Art. 5(1)(a) (Transparency), Arts. 12–14 (Information obligations), Art. 15 (Right of access) GDPR	Technical transparency optimises UX by reducing cognitive overload (e.g., simple progress bars). Legal transparency requires a comprehensive, granular disclosure of all processed data, including backend

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Levels	Requirement	Primary Legal Principles	Technical Implications
			algorithmic weights and profiling inferences.
Micro	Respon-sibility	---	The responsibility is to abide by the learner's decision and not to do anything without prior agreement that could undermine trust. Even though there is no legal basis for this, it is important to meet this requirement.
Micro	Fairness & Ethics	Art. 5(1)(a) (Fairness), Art. 9 (Processing of special categories of data) GDPR	The Accessibility Paradox: To ensure technical fairness and automatically provide accessible material (e.g., for dyslexia or ADHD), the system must know about the condition. This requires processing highly sensitive health data, which is subject to stringent legal restrictions.
Micro	Tamper-proof	Arts. 25, 32 GDPR	The data managed and persisted by learner models must be protected from unauthorised intervention and changes. Any changes need to be logged in a traceable manner.
Micro, Meso, Macro	Openness & Visualisation	Art. 5(1)(a),(d) (Accuracy), Art. 16 (Right to rectification) GDPR	Pedagogical negotiability allows learners to alter their skill levels for self-reflection or motivation. Legal accuracy requires that institutional data objectively reflect reality (especially when used to track prerequisites). If a learner manually overrides a system-calculated skill state, the database should store both values separately.
Micro	Negotiation Options	---	The various negotiation options involve purely technical interaction and are not based on any legal framework.
Micro	Data Minimisation & Sustainability	Art. 5(1)(c) GDPR	Only data that is absolutely necessary for adaptation should be collected. Minimising the amount of collected data while maximising the value of the information avoids unnecessary strain on the infrastructure.
Micro, Meso, Macro	Modularity & Flexibility	---	This involves a wide range of technical aspects and is not subject to any legal requirements.
Micro, Meso, Macro	Maintain-ability & Expandability	---	This involves a wide range of technical aspects and is not subject to any legal requirements.
Micro, Meso, Macro	Standardised / Interoper-ability & Integration	Art. 20 (Right to data portability) GDPR	The GDPR grants a legal right to transfer data in a machine-readable format; however, due to the lack of widely adopted, semantically unified learner-modelling standards, technical transferability does not guarantee pedagogical utility in a new LMS.

VIII. Recommended Action and Guidelines

Now you are a reader and possibly a future designer or developer of learner models, and you want to take the first step towards adaptive learning environments with learner models, or to integrate a learner model into an existing learning environment. However, you do not know how to get started. With many vague requirements, the GDPR can be hard to comply with, especially without legal expertise. The following section aims to provide a short roadmap for tackling GDPR compliance in a real-world scenario when learner models are to be integrated into higher education. Since a list for full compliance would be too extensive for this work, we focus on the essentials.

A. Common Process

However, as a first step, any processing of personal data must be based on a valid legal basis — typically either informed consent or a legal authorisation, such as the public task basis for educational institutions. Once the legal basis is clarified, data minimisation and storage limitation should become primary focus areas. However, this does not mean that the amount of data per se must be reduced; rather, the identifiability of individuals must be minimised. Accordingly, techniques such as pseudonymisation and, where feasible, anonymisation should be considered to reduce the risk of privacy breaches. In addition, the amount of data to be collected should still be critically examined to determine whether it is necessary and can make a meaningful contribution, and thus be used.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

In addition to legal regulations, technical and non-legal requirements must be considered to design learner models that are practical and effective. The variety of possible requirements analysed is wide. It has not yet been aggregated in a form that makes it easier for designers and developers of learner models to get started. Open learner models offer users transparency about their data, thereby promoting self-reflection and independent learning. Visualisations of complex data help users understand their learning status and continue working in a targeted manner. In addition, these can help increase transparency and build trust in the learning environment, as the data collected (in excerpts) is visualised, providing insight into what data is being collected and how it contributes to learning support.

Modular and interoperable structures ensure that learner models can be flexibly integrated and transferred between different learning environments. Maintainability and extensibility allow for long-term adaptation to new learning contexts. Overall, these technical requirements illustrate that a well-designed reference model forms the basis for the practical, transparent and fair design of adaptive learning environments.

Another key priority is transparency and traceability in all processing steps: learners must be able to understand how their personal data is processed, who has access to it, and how the system derives its outputs or recommendations. Only then can trust in adaptive learning technologies be established. In addition, fairness and the avoidance of algorithmic bias ensure that all learners have equal opportunities. These questions must be addressed in the next steps, including the design, implementation, and evaluation of a prototype learner model that complies with the outlined legal requirements.

B. Recommended Strategy

The following recommendations for action are intended to simplify the introduction for non-lawyers and provide an overview of the most important topics and issues. To this end, we have summarised the top five most important legal points.

1) Assessment of Personal Data

First, developers of learner models must consider the data required to develop and use the model. For instance, one could need data from learners to fill the model. Once the needed data is determined, developers must ask themselves whether this data will give any information about the persons. If so, these data fall within the material scope of the GDPR (Art. 2(1) GDPR). Developers should reconsider whether this specific personal data is really needed. Could the learner model be developed and used without this information linked to an identifiable person? The use of learner models primarily involves processing personal data. All this data should be listed, as it will be needed later, e.g., for the privacy policy.

2) Legal Basis for Processing the Personal Data

Afterwards, it must be ensured that there is a legal ground for processing the personal data, as described under VI. The legal grounds can be discussed with the entity's data protection officer later. If the learning model is to be used at a later date to grade learners, it falls within the scope of Art. 22 GDPR and is generally prohibited.

3) Safeguards

In the following step, developers must consider safeguards, referred to as appropriate technical and organisational measures in Art. 25 GDPR. In most cases, pseudonymisation of personal data is recommended. Regarding data access, a role-based access model should be implemented, distinguishing between user groups such as teachers and researchers. Access rights must be granted in accordance with this role concept, ensuring that each user can only access data necessary for their specific purpose. Furthermore, encryption should be applied both in transit and at rest to protect the data from unauthorised access. Finally, a deletion concept must be established to ensure that personal data is erased when it is no longer required for the purposes for which it was collected, or, at least, anonymised if it remains relevant to the system.

4) Feedback by the Data Protection Officer / Documentation

Data protection officers usually provide a form to document the aforementioned aspects of data protection. The elaborated concept can be discussed with the data protection officer at the higher education institution to obtain final approval. Higher Education institutions must prepare comprehensive documentation for learners and other data subjects, including a privacy policy that explains the AI system in an accessible language and informs users about their rights under the Arts. 15-22 GDPR, especially the right to object under Art. 21 GDPR for learning analytics.

5) Next Steps

In addition, they should be in close contact with their data protection officer to prevent data breaches. At the same time, they should always obtain ethics committee approval before beginning implementation, even if they already have legal approval.

C. Overview and Checklist

In the appendix, we have compiled and summarised some key information to help non-lawyers get started. This is intended as an initial introduction and is in no way a substitute for binding legal advice; rather, it is designed to raise awareness of this highly complex subject.

The checklist in Appendix A provides a quick overview of the most important laws and their classifications. It serves as an initial introduction and working basis. It must then be adapted and expanded to suit your own needs in consultation with specialists (data protection officers and ethics committees).

⇨ *See Appendix A*

We have compiled and briefly summarised key legal principles in Appendix B to ensure that no important principles are missed. This list is not exhaustive and does not cover all cases; rather, it serves as a simplified introduction and quick overview.

⇨ *See Appendix B*

As many legal terms in data protection are not self-explanatory to non-lawyers, we have provided Appendix C,

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

a glossary of the most important legal terms in data protection, as a reference guide. These are grouped into categories and briefly explained.

⇨ *See Appendix C*

IX. Conclusion and Future Work

A. Conclusion

Learning models have great potential to make education more individualised, efficient and equitable, but they are still a long way from being implemented in a standardised and legally compliant manner. The key challenge is to reconcile technical innovations with strict legal requirements, particularly those of the GDPR, in order to subsequently be able to use it in Europe, specifically in Germany.

The GDPR sets out only very general and technology-neutral requirements, offering few concrete implementation guidelines for learner models. Due to the lack of case law and specific regulatory guidance, legal obligations for learner models remain vague and difficult to apply in practice. This underscores the need for stronger support from data protection authorities in clarifying how GDPR principles can be operationalised in educational technology contexts. The guidelines and recommendations for action in Section VIII are to simplify the introduction to this topic.

B. Challenges and Invitations

Even though the legal requirements within the EU on this topic are not easy to understand, further steps could be taken to develop a reference model for learner models, which would enable other educational institutions to get started with adaptive learning environments more quickly, as learner models have been proven to make a decisive contribution to accommodate the heterogeneity of learners and support their learning processes individually and sustainably. However, an open standard for the use of learning models would be important and useful not only from this perspective, for example, to enable educational institutions to differentiate themselves from other providers by offering targeted, individualised support to learners, thereby minimising dropout rates in the long term, but also, of course, from the perspective of learners, who would be more motivated to engage with the content as they are neither under- nor overchallenged, but are individually supported according to their needs [74], [75]. This well-known and scientifically proven flow experience has a demonstrably positive effect on learners and the content [76], [77]. Even though people with disabilities make up only a small part of our society, they should not be excluded but included. The first step is to create an inclusive learning environment that automatically recognises learners with learning disabilities and fully supports them in their personal learning journey through accessible offerings. Unfortunately, it is often overlooked or forgotten that, according to the Charter of Fundamental Rights of the European Union, “everyone [...] has the right to education and to have access to vocational and continuing training” [78] Art. 14, mn. 3. Another major legal advantage of an open standard would be that learner data is standardised. This makes it easier to export and import data to and from another learning environment. From a purely legal point of view, learners have already had this right since the GDPR came into force with Art. 20 (1) GDPR' Right to data portability', which enables them to “receive their personal data [...] in a structured, commonly used and machine-readable format”. Apart from the legal framework, this enables easier, closer exchange between (state) educational institutions and provides a remedy when learners switch between courses at different universities, as they could transfer their digital data as a learner model through machine learning. Learners can also benefit from this, as they do not have to answer the same questions every time they register for a new learning environment. Instead, if their personal data were available externally, for example, through a data trust model, and they gave their permission to use individual services, their learner model would be maintained centrally, without redundant information across different educational institutions or learning environments. After completing their studies, this could include further professional training in a work context.

C. Future Work

The next technical steps are to conceptualise the legal constraints described in this publication with the help of the local data protection authority and the functional requirements [44] into an interoperable and legally compliant learning model, then to implement this model and integrate it into the learning platform [18]. We have already developed an initial prototype through an iterative process, refining it in successive cycles whilst implementing the various requirements in accordance with their classification (see Section V) [79]. The aim of this initial Minimum Viable Product (MVP) was to demonstrate correct functionality within a legal framework. To this end, this prototype was integrated into a proof-of-concept implementation of the digital learning environment and subjected to initial validation tests [80]. This initial prototype of a learner model must be evaluated and further developed in compliance with the legal conditions so that the needs of learners are met, and the learning process is improved in a sustainable and prosperous manner. An (empirical) assessment can be divided into two aspects: (1) legal validation and (2) an evaluation of the learner model and the associated requirements that have already been implemented. Regarding (1), it can be said that a classic, schematic validation based on a catalogue of criteria is not possible, as the interpretation of legal standards is carried out using legal methodology and, in the event of a dispute, is determined in a binding manner by the courts. The situation is somewhat more difficult when it comes to assessing learner models (2). There is not really much published on this topic in the literature. The learning environment is often validated as an overall system, but not its individual components; if so, most likely the inference component, as shown by a further systematic literature review of possible methods for evaluating learner models [81]. We are therefore currently working on the design and implementation of possible assessments for learner models.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

ACKNOWLEDGMENT

This work is a collaboration between the VoLL-KI project and the Komp-HI project and funded by the German Ministry of Education and Research (*Bundesministerium für Bildung und Forschung*) under grants 16DHBKI090 and 16DHBKI073.

REFERENCES

[1] F. Böck, H. Link, and D. Landes, 'Learner Models: Requirements and Legal Issues for the Development and Application of Learner Models', presented at the AISyS 2025, The Second International Conference on AI-based Systems and Services, Sep. 2025, pp. 27–35. Accessed: Oct. 24, 2025. [Online]. Available: https://www.thinkmind.org/library/AISyS/AISyS_2025/aisys_2025_1_40_80019.html

[2] N. Tagalidou, M. Bui, and M. Vukelić, 'Finding the Sweet Spot: Assessing Skill–Challenge Balance and Flow Using a Novel Game-based Task Paradigm', in *Proceedings of the Mensch und Computer 2025*, Chemnitz Germany: ACM, Aug. 2025, pp. 594–598. doi: 10.1145/3743049.3748556.

[3] S. Deterding and J. Cutting, 'Objective Difficulty-Skill Balance Impacts Perceived Balance but Not Behaviour: A Test of Flow and Self-Determination Theory Predictions', *Proc. ACM Hum.-Comput. Interact.*, vol. 7, no. CHI PLAY, pp. 1179–1205, Sep. 2023, doi: 10.1145/3611065.

[4] S. Zhang and S. Wang, 'Modeling Learner Heterogeneity: A Mixture Learning Model With Responses and Response Times', *Front. Psychol.*, vol. 9, p. 2339, Dec. 2018, doi: 10.3389/fpsyg.2018.02339.

[5] T. Kärner, J. Warwas, and S. Schumann, 'A Learning Analytics Approach to Address Heterogeneity in the Classroom: The Teachers' Diagnostic Support System', *Tech Know Learn*, vol. 26, no. 1, pp. 31–52, Mar. 2021, doi: 10.1007/s10758-020-09448-4.

[6] E. A. Edmonds, 'Adaptive man-computer interfaces', *Computing skills and the user interface*, vol. 122, pp. 389–426, 1981.

[7] A. Johnson and N. Taatgen, 'User Modeling', in *Handbook of human factors in Web design*, R. W. Proctor and K.-P. L. Vu, Eds, Mahwah, N.J: Lawrence Erlbaum Associates, 2005, pp. 424–439.

[8] F. Böck, D. Landes, and Y. Sedelmaier, 'Learner Models: A Systematic Literature Research in Norms and Standards', in *Proceedings of the 16th International Conference on Computer Supported Education*, Angers, France: SCITEPRESS - Science and Technology Publications, 2024, pp. 187–196. doi: 10.5220/0012556100003693.

[9] F. Böck, M. Ochs, A. Henrich, D. Landes, J. L. Leidner, and Y. Sedelmaier, 'Learner models: design, components, structure, and modelling: A systematic literature review', *User Model User-Adap Inter*, vol. 35, no. 15, p. 81, Nov. 2025, doi: 10.1007/s11257-025-09434-4.

[10] F. Böck, 'A Research Agenda for Learner Models: For adaptive Educational Digital Learning Environments', in *Innovation via Collaborative Learning in Engineering Education*, vol. 1847, M. E. Auer and P. Toth, Eds, in Lecture Notes in Networks and Systems, vol. 1847. , Cham: Springer Nature Switzerland, 2026, pp. 3–13. doi: 10.1007/978-3-032-18885-4_1.

[11] Society for Learning Analytics Research, '1st International Conference on Learning Analytics and Knowledge (LAK11) – Call for Papers'. [Online]. Available: https://tekri.athabascau.ca/analytics/

[12] P. Long and G. Siemens, 'Penetrating the fog: Analytics in learning and education', *EDUCAUSE Review*, vol. 46, no. 5, Oct. 2011, Accessed: May 26, 2026. [Online]. Available: https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education

[13] P. Long and G. Siemens, 'Penetrare la nebbia: tecniche di analisi per l'apprendimento', *TD Tecnologie Didattiche*, vol. 22, no. 3, Dec. 2014, doi: 10.17471/2499-4324/195.

[14] S. Bull and J. Kay, 'Open Learner Models', in *Advances in Intelligent Tutoring Systems*, R. Nkambou, J. Bourdeau, and R. Mizoguchi, Eds, in Studies in Computational Intelligence. , Berlin, Heidelberg: Springer, 2010, pp. 301–322. doi: 10.1007/978-3-642-14363-2_15.

[15] S. Bull and J. Kay, 'Student Models that Invite the Learner In: The SMILI:() Open Learner Modelling Framework', *International Journal of Artificial Intelligence in Education*, vol. 17, no. 2, pp. 89–120, 2007.

[16] R. Bodily *et al.*, 'Open learner models and learning analytics dashboards: a systematic review', in *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, Sydney New South Wales Australia: ACM, Mar. 2018, pp. 41–50. doi: 10.1145/3170358.3170409.

[17] J. Kay and S. Bull, 'New Opportunities with Open Learner Models and Visual Learning Analytics', in *Artificial Intelligence in Education*, vol. 9112, C. Conati, N. Heffernan, A. Mitrovic, and M. F. Verdejo, Eds, in Lecture Notes in Computer Science, vol. 9112. , Cham: Springer International Publishing, 2015, pp. 666–669. doi: 10.1007/978-3-319-19773-9_87.

[18] F. Böck, A. Deuerling, and D. Landes, 'Adaptive Learning Environment Reference Architecture for an Optimised Learning Process', in *2025 IEEE Global Engineering Education Conference (EDUCON)*, London, United Kingdom: IEEE, Apr. 2025, pp. 1–10. doi: 10.1109/EDUCON62633.2025.11016496.

[19] *[ISO/IEC TR 20748-1:2016] Information technology for learning, education and training — Learning analytics interoperability — Part 1: Reference model*, ISO/IEC TR 20748-1:2016, Feb. 2016. [Online]. Available: https://www.iso.org/standard/68976.html

[20] *Charter of Fundamental Rights of the European Union*, vol. C 326/391. 2016, p. 17. [Online]. Available: http://data.europa.eu/eli/treaty/char_2012/oj

[21] *Regulation (EU) 2016/679 of the European Parliament and of the Council*, vol. 02016R0679-20160504. 2016, p. 78. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02016R0679-20160504

[22] N. P. de Koch, 'Software Engineering for Adaptive Hypermedia Systems', Ludwig-Maximilians-Universität München, München, 2000. [Online]. Available: https://www.pst.ifi.lmu.de/~kochn/PhDThesisNoraKoch.pdf

[23] O. Abdennour, H. Kemouss, M. Erradi, and M. Khaldi, 'From Learner Profile Management to Learner Model Modeling', in *Advances in Educational Technologies and Instructional Design*, M. Khaldi, Ed., IGI Global, 2023, pp. 71–93. doi: 10.4018/978-1-6684-7634-5.ch004.

[24] X. Li and S. He, 'Research and Analysis of Student Portrait Based on Campus Big Data', in *International Conference on Big Data Analytics (ICBDA)*, Mar. 2021, pp. 23–27. doi: 10.1109/ICBDA51983.2021.9403070.

[25] Y. Jun-min, X. Song, L. Da-Xiong, W. Zhi-Feng, H. Peng-Wei, and X. Chen, 'Research on the Construction and Application of Individual Learner Model', in *International Conference on Computing and Information Technology*, Procedia Computer Science, Jun. 2018, pp. 88–92. doi: 10.1016/j.procs.2018.04.189.

[26] D. Hooshyar, M. Pedaste, K. Saks, Ä. Leijen, E. Bardone, and M. Wang, 'Open learner models in supporting self-regulated

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

learning in higher education: A systematic literature review', *Computers & Education*, vol. 154, p. 103878, Sep. 2020, doi: 10.1016/j.compedu.2020.103878.

[27] S. Ouf, M. Abd Ellatif, S. E. Salama, and Y. Helmy, 'A proposed paradigm for smart learning environment based on semantic web', *Computers in Human Behavior*, vol. 72, pp. 796–818, Jul. 2017, doi: 10.1016/j.chb.2016.08.030.

[28] X. Wang, Y. Maeda, and H.-H. Chang, 'Development and techniques in learner model in adaptive e-learning system: A systematic review', *Computers & Education*, vol. 225, p. 105184, Feb. 2025, doi: 10.1016/j.compedu.2024.105184.

[29] A. Zanellati, D. D. Mitri, M. Gabbrielli, and O. Levrini, 'Hybrid Models for Knowledge Tracing: A Systematic Literature Review', *IEEE Trans. Learning Technol.*, vol. 17, pp. 1021–1036, 2024, doi: 10.1109/TLT.2023.3348690.

[30] B. Kitchenham, 'Procedures for Undertaking Systematic Reviews', Keele University (TR/SE0401) and National ICT Australia Ltd. (0400011T.1), Joint Technical Report, Jul. 2004.

[31] B. A. Kitchenham and S. Charters, 'Guidelines for performing Systematic Literature Reviews in Software Engineering', Keele University, EBSE 2007-001, Jul. 2007. [Online]. Available: https://www.elsevier.com/__data/promis_misc/525444systematicreviewsguide.pdf

[32] *1EdTech Learner Information Package Specification*, Standard, Mar. 09, 2001. [Online]. Available: https://site.imsglobal.org/standards/sldm

[33] *[IEEE P1484.2] - Standard for Information Technology — Learning Systems — Learner Model*, PAR IEEE P1484.2, Oct. 09, 1997. Accessed: Aug. 02, 2003. [Online]. Available: http:/ltsc.ieee.org:80/wg2/

[34] W. QIU, J. DU, and F. LI, 'Knowledge Description Frame of Learning Resources for Recommendation System: From the Perspectives of Learning Psychology', in *2020 15th International Conference on Computer Science & Education (ICCSE)*, Aug. 2020, pp. 438–440. doi: 10.1109/ICCSE49874.2020.9201686.

[35] X. Zhao, 'Research on learner model of adaptive learning system', in *2021 16th International Conference on Computer Science & Education (ICCSE)*, Lancaster, United Kingdom: IEEE, Aug. 2021, pp. 906–909. doi: 10.1109/ICCSE51940.2021.9569673.

[36] S. Wang, L. Yuan, and H. Yang, 'User Model Construction of Chinese Learners Based on Learning Style', in *International Conference on Computer Science and Educational Informatization (CSEI)*, Jun. 2021, pp. 70–75. doi: 10.1109/CSEI51395.2021.9477714.

[37] T. Fei Zhou, Y. Qing Pan, and L. R. Huang, 'Research on Personalized E-Learning Based on Decision Tree and RETE Algorithm', in *2017 International Conference on Computer Systems, Electronics and Control (ICCSEC)*, Dec. 2017, pp. 1392–1396. doi: 10.1109/ICCSEC.2017.8446741.

[38] A. N. Cormack, 'A Data Protection Framework for Learning Analytics', *Learning Analytics*, vol. 3, no. 1, Apr. 2016, doi: 10.18608/jla.2016.31.6.

[39] Q. Liu and M. Khalil, 'Understanding privacy and data protection issues in learning analytics using a systematic review', *Brit J Educational Tech*, vol. 54, no. 6, pp. 1715–1747, Nov. 2023, doi: 10.1111/bjet.13388.

[40] T. Hoel and W. Chen, 'Privacy and data protection in learning analytics should be motivated by an educational maxim—towards a proposal', *Research and Practice in Technology Enhanced Learning*, vol. 13, no. 1, p. 20, Dec. 2018, doi: 10.1186/s41039-018-0086-8.

[41] M. Paludi, 'The Right to Privacy and Data Protection for High School Students in the Context of Digital Learning Models and Learning Analytics', in *Proceedings of the Doctoral Consortium of the Learning Analytics Summer Institute Europe 2024 (LASI Europe 2024 DC)*, Jerez de la Frontera: Dimitriadis, Yannis; Cerezo, Rebeca; Balderas, Antonio; Martínez-Monés, Alejandra; Spikol, Daniel, Dec. 2018, pp. 1–10.

[42] H. Drachsler and W. Greller, 'Privacy and analytics: it's a DELICATE issue a checklist for trusted learning analytics', in *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16*, Edinburgh, United Kingdom: ACM Press, 2016, pp. 89–98. doi: 10.1145/2883851.2883893.

[43] J. Doveren, B. Heinemann, and U. Schroeder, 'Towards Guidelines for Data Protection and Privacy in Learning Analytics Implementation', in *Online-Labs in Education*, D. Uckelmann, G. Romagnoli, J. Baalsrud Hauge, and V. Kammerlohr, Eds, Nomos Verlagsgesellschaft mbH & Co. KG, 2022, pp. 45–52. doi: 10.5771/9783957104106-45.

[44] F. Böck, 'Learner Models: Requirements Analysis for Application in Adaptive Learning Environments and Recommendation Systems', in *2025 IEEE Digital Education and MOOCS Conference (DEMOcon)*, Boca Raton, FL, USA: IEEE, Oct. 2025, pp. 1–7. doi: 10.1109/DEMOcon65705.2025.11282849.

[45] N. Kano, N. Seraku, F. Takahashhi, and S. ichi Tsuji, 'Attractive Quality and Must-Be Quality', *Journal of the Japanese Society for Quality Control*, vol. 14, no. 2, pp. 147–156, Apr. 1984, doi: 10.20684/quality.14.2_147.

[46] *Regulation (EU) 2024/1689 of the European Parliament and of the Council*, vol. PE/24/2024/REV/1. p. 144. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ%3AL_202401689

[47] Bundesministerium der Justiz, *Bundesdatenschutzgesetz*. 2017, p. 45. [Online]. Available: https://www.gesetze-im-internet.de/bdsg_2018/BDSG.pdf

[48] F. Ricci, L. Rokach, and B. Shapira, Eds, *Recommender systems handbook*, Third edition. New York, NY: Springer, 2022.

[49] F. Doshi-Velez and B. Kim, 'Towards A Rigorous Science of Interpretable Machine Learning', *arXiv*, 2017, [Online]. Available: https://arxiv.org/abs/1702.08608

[50] C. Rudin, 'Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead', *Nat Mach Intell*, vol. 1, no. 5, pp. 206–215, May 2019, doi: 10.1038/s42256-019-0048-x.

[51] M. T. Ribeiro, S. Singh, and C. Guestrin, '"Why Should I Trust You?": Explaining the Predictions of Any Classifier', in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco California USA: ACM, Aug. 2016, pp. 1135–1144. doi: 10.1145/2939672.2939778.

[52] S. M. Lundberg and S.-I. Lee, 'A unified approach to interpreting model predictions', in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, in NIPS'17. Red Hook, NY, USA: Curran Associates Inc., 2017, pp. 4768–4777.

[53] K. Kitto and S. Knight, 'Practical ethics for building learning analytics', *Br J Educ Technol*, vol. 50, no. 6, pp. 2855–2870, Nov. 2019, doi: 10.1111/bjet.12868.

[54] Y.-S. Tsai, C. Perrotta, and D. Gašević, 'Empowering learners with personalised learning approaches? Agency, equity and transparency in the context of learning analytics', *Assessment & Evaluation in Higher Education*, vol. 45, no. 4, pp. 554–567, May 2020, doi: 10.1080/02602938.2019.1676396.

[55] S. Bull and J. Kay, 'SMILI☺: a Framework for Interfaces to Learning Data in Open Learner Models, Learning Analytics and Related Fields', *Int J Artif Intell Educ*, vol. 26, no. 1, pp. 293–331, Jan. 2016, doi: 10.1007/s40593-015-0090-8.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

[56] S. Bull, B. Ginon, C. Boscolo, and M. Johnson, 'Introduction of Learning Visualisations and Metacognitive Support in a Persuadable Open Learner Model', in *International Conference on Learning Analytics and Knowledge*, in LAK '16. New York, NY, USA: Association for Computing Machinery, Apr. 2016, pp. 30–39. doi: 10.1145/2883851.2883853.

[57] H. Ferreira, G. Oliveira, R. Araújo, F. Dorça, and R. Cattelan, 'An Open Model for Student Assessment Visualization', in *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*, Jul. 2019, pp. 375–379. doi: 10.1109/ICALT.2019.00117.

[58] J. A. Ruipérez-Valiente, P. J. Muñoz-Merino, D. Leony, and C. D. Kloos, 'ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform', *Computers in Human Behavior*, vol. 47, pp. 139–148, Jun. 2015, doi: https://doi.org/10.1016/j.chb.2014.07.002.

[59] C.-Y. Law, J. Grundy, R. Vasa, and A. Cain, 'An empirical study of user perceived usefulness and preference of open learner model visualisations', in *2016 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, Sep. 2016, pp. 49–53. doi: 10.1109/VLHCC.2016.7739663.

[60] S. Bull and J. Kay, 'Open Learner Models', in *Advances in Intelligent Tutoring Systems*, vol. 308, R. Nkambou, J. Bourdeau, and R. Mizoguchi, Eds, in Studies in Computational Intelligence, vol. 308. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 301–322. doi: 10.1007/978-3-642-14363-2_15.

[61] J. Kay, K. Bartimote, K. Kitto, B. Kummerfeld, D. Liu, and P. Reimann, 'Enhancing learning by Open Learner Model (OLM) driven data design', *Computers and Education: Artificial Intelligence*, vol. 3, p. 100069, May 2022, doi: 10.1016/j.caeai.2022.100069.

[62] C.-Y. Law, J. Grundy, R. Vasa, and A. Cain, 'An empirical study of user perceived usefulness and preference of open learner model visualisations', in *2016 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, Sep. 2016, pp. 49–53. doi: 10.1109/VLHCC.2016.7739663.

[63] S. Bull, B. Ginon, C. Boscolo, and M. Johnson, 'Introduction of Learning Visualisations and Metacognitive Support in a Persuadable Open Learner Model', in *International Conference on Learning Analytics and Knowledge*, in LAK '16. New York, NY, USA: Association for Computing Machinery, Apr. 2016, pp. 30–39. doi: 10.1145/2883851.2883853.

[64] S. Bull, 'Supporting learning with open learner models', in *Proceedings of fourth Hellenic Conference on Information and Communication Technologies in Education*, Athen, Griechenland, 2004, pp. 47–61.

[65] *[ISO/IEC 25010:2011] Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models*, 35.080 Software 1, Mar. 2011. [Online]. Available: https://www.iso.org/standard/35733.html

[66] *[ISO/IEC 25002:2024] Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Quality model overview and usage*, 35.080 Software 1, Mar. 2024. [Online]. Available: https://www.iso.org/standard/78175.html

[67] W. Kotschy, 'Art. 6 GDPR', in *The EU General Data Protection Regulation (GDPR): a commentary*, C. Kuner, C. Docksey, L. A. Bygrave, and L. Drechsler, Eds, Oxford, United Kingdom: Oxford University Press, 2019.

[68] J. T. Helmke and H. Link, 'Das Verhältnis von Art. 6 Abs. 2 und 3 DSGVO', *Datenschutz Datensich*, vol. 47, no. 11, pp. 708–714, Nov. 2023, doi: 10.1007/s11623-023-1848-z.

[69] A. Roßnagel and P. Richer, 'Art. 5', in *General data protection regulation: article-by-article commentary*, First edition., I. Spiecker Döhmann, E. Papakōnstantinu, G. Hornung, and P. de Hert, Eds, in Beck-online Bücher. , Baden-Baden, Germany: Nomos, 2023.

[70] The European Data Protection Board, 'Guidelines 01/2025 on Pseudonymisation'. Jan. 16, 2025. [Online]. Available: https://www.edpb.europa.eu/system/files/2025-01/edpb_guidelines_202501_pseudonymisation_en.pdf

[71] The European Data Protection Board, 'Guidelines 4/2019 on Article 25 Data Protection by Design and by Default, Version 2.0'. Oct. 20, 2020. [Online]. Available: https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_en.pdf

[72] Article 29 Data Protection Working Party, 'Opinion 4/2007 on the concept of personal data'. Jun. 20, 2007. [Online]. Available: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf

[73] P. Schantz, 'Art. 44', in *General data protection regulation: article-by-article commentary*, First edition., I. Spiecker Döhmann, E. Papakōnstantinu, G. Hornung, and P. de Hert, Eds, in Beck-online Bücher. , Baden-Baden, Germany: Nomos, 2023.

[74] M. Csikszentmihalyi, *Flow: the psychology of optimal experience*, 1st ed. New York: Harper & Row, 1990.

[75] M. Csikszentmihalyi, 'The flow experience and its significance for human psychology', in *Optimal Experience*, 1st edn, M. Csikszentmihalyi and I. S. Csikszentmihalyi, Eds, Cambridge University Press, 1988, pp. 15–35. doi: 10.1017/CBO9780511621956.002.

[76] M. A. Guadagnoli and T. D. Lee, 'Challenge Point: A Framework for Conceptualizing the Effects of Various Practice Conditions in Motor Learning', *Journal of Motor Behavior*, vol. 36, no. 2, pp. 212–224, Jun. 2004, doi: 10.3200/JMBR.36.2.212-224.

[77] C. J. Fong, D. J. Zaleski, and J. K. Leach, 'The challenge–skill balance and antecedents of flow: A meta-analytic investigation', *The Journal of Positive Psychology*, vol. 10, no. 5, pp. 425–446, Sep. 2015, doi: 10.1080/17439760.2014.967799.

[78] C. Calliess and M. Ruffert, Eds, *EUV/AEUV: das Verfassungsrecht der Europäischen Union mit Europäischer Grundrechtecharta: Kommentar*, 6. Auflage. München: C.H. Beck, 2022.

[79] F. Böck and D. Landes, 'PLM - an Open Learner Model for Adaptive Learning Environments – Specification', in *2025 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, Macao: IEEE, Dec. 2025, pp. 1–7. doi: 10.1109/TALE66047.2025.11346719.

[80] F. Böck, A. Deuerling, and D. Landes, 'Towards a Personalised Digital Learning Environment of Tomorrow':, in *Proceedings of the 18th International Conference on Computer Supported Education*, Benidorm, Spain: SCITEPRESS - Science and Technology Publications, 2026, pp. 1479–1490. doi: 10.5220/0014738000004021.

[81] F. Böck, 'Learner Models: Assessment Methods - A Literature Analysis', submitted & under review, 2026.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

APPENDIX

APPENDIX A

Overview of the most important GDPR Articles

The following overview table serves merely as a guide to help you get started with the most important provisions of the GDPR and to shed light on the highly complex subject of data protection. This working aid does not claim to be fully compliant, as achieving such a standard would exceed the scope of this paper. It is therefore recommended that you involve your data protection officer or in-house data protection authority right from the start and discuss this complex topic with specialist staff.

☑	Question	Consequenses ☑	☒
☐	Am I processing **personal data** (name, ID, performance data, learning behaviour, etc.)? Do the data allow a person to be identified directly or indirectly? Can I fully anonymise data if they are to be reused?	Art. 2(1) GDPR (*filing system*), Art. 4(1) GDPR (*personal data*) → GDPR fully applicable	GDPR not applicable
☐	Do I have a **legal basis** for the processing? Can the processing be based on voluntary consent? Is there a legal basis in the applicable national law?	Art. 6(1) GDPR (l*awfulness of processing*) [consent lit. a or public task lit. e]	Processing unlawful
☐	Am I processing **special categories of personal data** (e.g., health, ethnic origin)?	Art. 9 GDPR (p*rocessing of special categories of personal data*) → only with an explicit statutory basis	No additional requirements
☐	Am I using the data for a **purpose other than originally collected**?	Art. 5(1)(b) GDPR (*purpose limitation*), Art. 6(4) GDPR (*purpose of data collection*) → check purpose compatibility	No additional review required
☐	Are the data used for **automated decision-making with legal effect** (e.g., grading)?	Art. 22 GDPR (a*utomated individual decision-making*) → generally prohibited	No restriction under Art. 22
☐	Is there **profiling** within the meaning of GDPR?	Art. 4(4) GDPR (*profiling*), Art. 22 GDPR (*Automated individual decision-making*) → review required	No specific requirements
☐	Am I collecting **more data than strictly necessary**?	Art. 5(1)(c) GDPR (*data minimisation*) → reduce	No action required
☐	Am I storing data **longer than necessary**?	Art. 5(1)(e) GDPR (*storage limitation*) → deletion concept required	Ok
☐	Is there a **deletion policy** (e.g., upon exmatriculation)?	Art. 5(1)(e) GDPR (*storage limitation*)	Create one
☐	Do I inform data subjects **transparently** about processing?	Art. 5(1)(a) GDPR (*lawfulness, fairness and transparency*), Arts. 12 – 14 GDPR (*information obligations)*	Implement information obligations
☐	Can users exercise their **data subject rights** (access, erasure, objection)?	Arts. 15 – 22 GDPR (d*ata subject rights*)	Establish processes
☐	Is a **Data Protection Impact Assessment (DPIA)** required? (high risk, extensive profiling)	Art. 35 GDPR (*data protection impact assessment*)	No DPIA required
☐	Do I use **processors** (cloud providers, external vendors)?	Art. 28 GDPR (*DPA/AV contract required*)	No action required
☐	Are data processed **outside the EU**?	Arts. 45 – 49 GDPR (*Processor*) [review third-country transfer]	Ok
☐	Are data **pseudonymised**, where possible?	Art. 4(5) GDPR (*pseudonymisation*), Art. 25 GDPR (*privacy by design*)	Check if possible
☐	Do I have a **role and access control concept**?	Art. 25 GDPR (*privacy by design*), Art. 32 GDPR (*technical & organisational measures*)	Implement
☐	Are data **encrypted** (at rest & in transit)?	Art. 32 GDPR (*security of processing*)	Add security measures
☐	Do I enable **data portability** (export of learner model)?	Art. 20 GDPR (r*ight to data portability*)	Add function
☐	Was the **Data Protection Officer** involved early?	Art. 38 GDPR (*data protection officer)*	Urgently catch up
☐	Is there a risk of **algorithmic discrimination/bias**?	Art. 5(1)(a) GDPR (*fairness*)	Conduct a fairness assessment

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

APPENDIX B

Key Legal Principles (*Brief Overview*)

The most important data protection laws are summarised briefly below to facilitate entry and serve as a checklist to ensure they are not overlooked. This list does not claim to be complete or to cover all casualties, but merely serves as a simplified introduction and quick overview.

☑	Law	Relevant Content
☐	Art. 5 GDPR	Basic principles (*lawfulness, transparency, purpose limitation, data minimisation, storage limitation, integrity*)
☐	Art. 6 GDPR	Legal basis
☐	Art. 9 GDPR	Special categories of personal data
☐	Art. 12 – 14 GDPR	Information obligations
☐	Art. 15 – 22 GDPR	Rights of data subjects
☐	Art. 22 GDPR	Prohibition of automated decisions with legal effect
☐	Art. 25 GDPR	Privacy by design and by default
☐	Art. 28 GDPR	Processing by a processor
☐	Art. 32 GDPR	Security of processing
☐	Art. 35 GDPR	Data protection impact assessment
☐	Art. 45 et seq. GDPR	Transfer to third countries.
☐	Art. 20 GDPR	Data portability
☐	Art. 8 EU ChFR	Fundamental right to data protection
☐	Art. 14 EU ChFR	Right to education

In addition to the key data protection rights listed above, there are, however, other laws that play a role in conjunction with the aforementioned laws and which may therefore be worth considering in this context:

- EU Data Act
- AI Act
- Art. 14 EU ChFR (Right to education)

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

APPENDIX C

Key Legal Terms in Data Protection (Brief Overview)

This glossary serves as a reference work. The most important data protection laws are summarised briefly below to facilitate entry and serve as a checklist to ensure they are not overlooked. This list does not claim to be complete or to cover all casualties, but merely serves as a simplified introduction and quick overview.

Term	Explanation
Core Principles of Data Protection Law	
Lawfulness of Processing (Art. 5 (1)(a) GDPR)	Processing of personal data must be based on a valid legal ground under data protection law.
Purpose Limitation Art. 5 (1)(b) GDPR	Personal data must be collected for specified purposes and not used incompatibly with those purposes.
Data Minimisation Art. 5 (1)(c) GDPR	Only the personal data necessary for the intended purpose should be processed.
Transparency Art. 5 (1)(a) GDPR	Data subjects must be clearly informed about how and why their data is processed.
Storage Limitation Art. 5 (1)(e) GDPR	Personal data should be kept only for as long as necessary for the processing purpose.
Integrity and Confidentiality Art. 5 (1)(f) GDPR	Personal data must be protected against unauthorised access, loss, or damage.
Accountability (Art. 5(2) GDPR	Controllers are responsible for and must be able to demonstrate compliance with data protection principles.
Data Portability (*Art. 20 GDPR*)	Individuals have the right to receive their personal data in a structured and transferable format.
Legal Bases and Justifications for Processing	
Lawfulness Requirement (*Art. 5(1), Art. 6 GDPR*)	Personal data may only be processed when a valid legal basis applies, and principles are respected.
Consent as Legal Basis (*Art. 6(1)(a) GDPR*)	Processing is lawful if the data subject has freely given informed and explicit consent.
Public Interest Processing (*Art. 6(1)(e) GDPR*)	Processing is permitted when necessary to perform a task carried out in the public interest or official authority.
Legal Authorisation under Member State Law	National laws may provide specific legal bases that permit certain data processing activities.
Categories of Personal Data	
Personal Data (Art. 4(1) GDPR	Any information relating to an identified or identifiable natural person.
Sensitive Personal Data / Special Categories of Data (*Art. 9 GDPR*)	– racial and ethnic origin, – political opinions, – religious or philosophical beliefs, – trade union membership, – genetic data, – biometric data, – health data, – data concerning sex life or sexual orientation
Inferred Data and Profiling Results	Data derived from analysis or automated evaluation is used to predict or assess personal aspects.
Data Protection Safeguards and Technical-Legal Measures	
Pseudonymisation (Art. 4(5) GDPR	Processing personal data so it cannot be attributed to a person without additional information.
Anonymisation	Irreversibly removing identifiers so that individuals can no longer be identified.
Encryption of Personal Data	Transforming data into a coded format to prevent unauthorised access.
Role-Based Access Control	Restricting system access based on the user's role and responsibilities.
Data Protection by Design and by Default (*Art. 25 GDPR*)	Systems must integrate data protection measures from the outset and use privacy-friendly default settings.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Term	Explanation
Security of Processing (*Art. 32 GDPR*)	Appropriate technical and organisational measures must ensure a level of security appropriate to the risk.
Rights of Data Subjects	
Right of Access to Personal Data (Art. 15 GDPR)	Individuals can obtain confirmation and a copy of the personal data processed about them.
Right to Object (*Art. 21 GDPR*)	Individuals may object to processing based on specific legal grounds, especially legitimate interests.
Right to Withdraw Consent (Art. 7(3) GDPR	Individuals can revoke previously given consent at any time with future effect.
Right to Data Portability (Art. 20 GDPR)	Individuals can receive and transfer their personal data between controllers.
Compliance and Governance Mechanisms	
Data Protection Impact Assessment (*Art. 35 GDPR*)	A risk assessment is required for high-risk processing to evaluate impacts on data subjects' rights.
Data Processing Agreement (*Art. 28 GDPR*)	A contract that governs processing activities carried out by a processor on behalf of a controller.
Use of Processors (*external service providers*)	Engaging third parties to process data under the controller's instructions and safeguards.
Role of the Data Protection Officer (Art. 37 GDPR)	An independent advisor who monitors compliance with data protection law within an organisation.
Documentation and Privacy Policies	Written records and notices that demonstrate and communicate compliance with data protection rules.
Emerging Legal-Technical Concepts	
Data Trust Model (*trusted intermediary for data governance*)	A governance framework where an intermediary manages data access and use in the interest of stakeholders.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Converter-Level Modeling and Coordinated Control of a Wind–Battery Powered Surface Water Heat Pump DC Microgrid: Case Study of Ouistreham, France

Joyce Assaf, Mamadou-Baïlo Camara, Damien Guilbert, Brayima Dakyo
Groupe de Recherche en Electrotechnique et Automatique du Havre
Université Le Havre Normandie
Le Havre, Normandie, France
e-mail: {joyce.assaf | camaram | damien.guilbert | brayima.dakyo}@univ-lehavre.fr

***Abstract*—This paper presents the design and validation of a DC microgrid integrating two parallel 30 kW Surface Water Heat Pumps (SWHPs), a 35 kW Wind Turbine (WT), and a 35 kWh Battery Energy Storage System (BESS). The subsystems are interconnected through a regulated 1000 V DC–bus and coordinated via a hierarchical control strategy. A detailed MATLAB/Simulink model is developed using measured wind speed data from the Le Cano Ouistreham, France pilot site and realistic thermal demand data from an operational aquathermal installation in Dijlemolens, Belgium. A refined Region-III power limitation strategy for the WT and cascaded voltage–current control loops for the BESS are implemented to ensure stable operation under fluctuating wind conditions and high thermal demand. Simulation results demonstrate tight DC–bus regulation, with a maximum voltage deviation below 0.62% and a Root Mean Square Error (RMSE) of 0.033%. Under peak thermal demand of 52.32 kW, coordinated WT–BESS power sharing maintains supply continuity without violating operational constraints. The proposed DC microgrid architecture provides a validated and scalable framework for renewable-powered aquathermal heating systems.**

Keywords-DC microgrid; Wind Turbine; Surface Water Heat Pump; Battery Energy Storage System; Energy Management Strategy.

I. Introduction

This journal paper extends our previous conference contribution presented at GREEN 2025 [1], where the preliminary DC microgrid architecture and initial simulation results were introduced. The present version significantly expands the modeling framework, refines the Wind Turbine (WT) Region-III control strategy, incorporates a cascaded dual-loop DC–bus voltage and battery current control strategy for the Battery Energy Storage System (BESS), and provides an extended system-level performance analysis under realistic pilot-site operating conditions.

Aquathermal Energy (AE) is the extraction, storage, and distribution of heat from water bodies, including surface water, wastewater, and groundwater. It is an emerging Renewable Energy Source (RES) with strong potential for heating and cooling applications. In AE systems using surface water, heat is extracted from the water body using a Surface Water Heat Pump (SWHP) via a heat exchanger, which then increases or decreases the temperature as needed for domestic hot water supply and other thermal energy applications [2]. AE systems reduce carbon emissions in the heating and cooling sector and decrease its reliance on imported energy [3].

According to the latest IRENA report, 87 million Heat Pumps (HPs) are projected to be installed in buildings, with wind and solar expected to contribute 91% of the total Renewable Energy (RE) capacity by 2050 [4]. When coupled with RESs, such as wind or Photovoltaics (PVs) to power their compressors, SWHPs offer a promising pathway to decarbonize the building sector [5]. In Europe, the deployment of WTs is spreading rapidly, with wind energy accounting for almost 39% of the total electricity generated from RESs in 2024 [6]. However, due to the intermittent nature of WTs, BESSs are integrated to effectively manage short-term fluctuations, while appropriate control strategies can significantly improve system reliability and efficiency [7].

Power electronics play an essential role in enabling the efficient integration of HPs with RESs. DC-links, incorporating AC–DC and DC–DC converters, allow for improved energy management, voltage regulation, and efficient power conversion across hybrid systems. Compared to traditional AC microgrids, DC-based systems offer lower conversion losses, improved energy management, and more straightforward control, particularly advantageous for hybrid configurations with large shares of DC loads or storage units [8].

In the literature, several studies have explored hybrid systems for HP applications [9]–[12]. Although these studies laid essential groundwork, they often overlooked power electronics integration, realistic thermal demand modeling, and dynamic power management between sources and loads. Few studies addressed power electronics control strategies and converter topologies for RE-powered HPs. For instance, in [13], the authors proposed an energy management strategy (EMS) for a microgrid that combines a PV, a BESS, and an air-source HP. The main goal of the study was to reduce the cost of the BESS by leveraging the thermal storage of the HP, incorporating a double fuzzy logic that coordinates power fluctuation stabilization between the BESS and the air-source HP. However, HP operation was mainly governed by RE availability and system efficiency, not by actual, time-varying thermal demand.

In [14], the authors presented a hybrid system that includes PV, WT, and BESS powering a DC-HP. However, the approach is limited as it only accounts for the electrical power consumption of the HP, with the thermal load profile being derived from thermodynamic equations without incorporating its actual dynamics. As a result, the control and performance

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

of the system are not realistically reflective of actual HP operation, which is typically governed by unpredictable and varying thermal loads over time. Furthermore, in [15], the authors proposed an advanced control strategy based on the model predictive control combined with a fractional short-circuit current approach to optimize power extraction from the PV system interfaced with a SWHP. However, this study modeled the SWHP as a resistive load, so results were only analyzed on the PV–BESS side. Alternatively, in [16], the authors proposed a high-voltage-gain DC–DC converter to improve PV-HP coupling efficiency. Although their converter showed promising results in handling variable solar irradiance, the system also lacked realistic thermal modeling of the HP. Although their work offers valuable insights, it remains unclear whether the results are scalable to real-world scenarios.

This paper contributes to the growing body of research on HPs, particularly SWHPs, by presenting a realistic simulation conducted for an AE harvesting pilot site within the "Water-Warmth" project, funded by Interreg North Sea Region, in Le Cano Ouistreham, France. This pilot site is currently in the preparatory phase for constructing a SWHP system, making it an ideal case study to simulate and validate RE integration and control strategies under realistic site conditions. The proposed system employs a 35 kW WT and a 35 kWh BESS to power two identical 30 kW SWHPs operating in parallel, all connected via a DC–bus with realistic control architecture, using both measured wind speeds from Ouistreham pilot site and thermal load profile from an operational site in Belgium. The originality of this study lies in its comprehensive simulation of component interactions between fluctuating RE generation and dynamic thermal demand, and its validation of a coordinated control strategy that ensures voltage stability under realistic fluctuating wind and load conditions. The proposed model provides a more accurate reflection of operational conditions and offers a pathway toward scalable, site-adapted AE solutions.

Unlike previous works that neglect realistic thermal demand dynamics or simplified power electronics modeling, this study integrates full electromechanical and converter-level modeling under measured site data.

In summary, the novelty of this work lies in three main contributions: (i) the use of real measured climatic and thermal demand profiles, preserving realistic variability in wind speed and heating demand; (ii) the coordinated control of the WT and the BESS to guarantee DC–bus voltage stability under high-load and fluctuating wind conditions, including a refined Region-III wind power limitation strategy; and (iii) the implementation of a DC microgrid that minimizes energy conversion stages, thereby improving structural efficiency and controllability.

The main objective of this study is to design and validate, through MATLAB/Simulink under real-site operating conditions, an EMS that ensures stable, demand-driven operation of a hybrid WT–BESS–SWHP DC microgrid.

This paper is organized as follows. Section I briefly reviews the state of the art in RE-powered heating systems. Section II presents the DC microgrid configuration, including the WT, BESS, SWHP units, and associated power electronic converters. Section III details the control framework implemented on the WT and BESS DC–DC converters, including voltage regulation and power-limiting strategies. Section IV discusses the performance analysis conducted under realistic pilot-site operating conditions within the framework of the WaterWarmth project. Finally, Section V concludes the paper and outlines key findings.

II. DC Microgrid Configuration

The rated power of the WT (35 kW) and the energy capacity of the BESS (35 kWh) were selected to ensure supply continuity under worst-case operating scenarios, particularly during simultaneous high thermal demand and low wind availability. The objective of this study is not to optimize energy capacity sizing, but rather to validate the coordinated control strategy and DC microgrid stability under realistic operating conditions.

A. Wind Turbine

The system shown in Figure 1 is based on a variable-speed WT coupled with a Permanent Magnet Synchronous Generator (PMSG), which converts mechanical energy to electrical energy. The rotating rotor blades extract kinetic energy from the wind and transform it into shaft torque, which is then converted into electricity by the generator. To maximize the energy extracted from the varying wind speeds, the system relies on control strategies that regulate the WT's rotor speed and blade pitch angle. A detailed description of the control strategy adopted is provided in Section III.

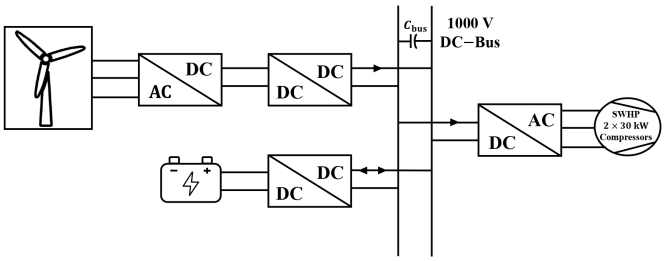


Figure 1. Overall configuration of the WT-BESS-SWHP system.

The wind speed data, illustrated in Figure 2, were collected for Le Cano Ouistreham, covering the period from March 7, 2024 at 11:00 AM to May 7, 2024 at 2:00 AM, with a sample time of one hour. The wind speed frequently drops below the 10.9 m/s nominal speed and occasionally falls near or below cut-in speed of the WT (3 m/s), reflecting the intermittent nature of wind energy. These fluctuations justify the need for both WT speed control to maximize energy capture under partial load conditions and the integration of a BESS to stabilize the system's power supply.

For computational feasibility, the original time-series datasets were temporally compressed and mapped onto a 1456 s simulation horizon while preserving their statistical variability and dynamic characteristics.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

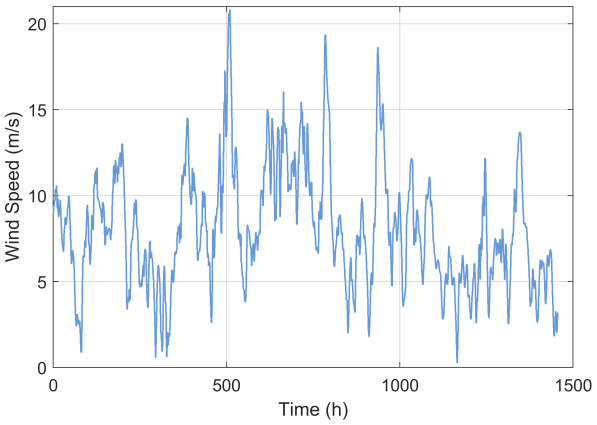


Figure 2. The wind speed profile in Le Cano Ouistreham, France.

1) Wind Turbine Aerodynamic Modeling: The mechanical power extracted from the wind by a horizontal-axis WT is expressed as:

$$P_{\text{aero}} = \frac{1}{2}\rho A C_p(\lambda, \beta) V_w^3 \tag{1}$$

where ρ is the air density (1.225 kg/m^3), $A = \pi R^2$ is the swept area of the rotor blades, C_p is the power coefficient, V_w is the wind speed, λ is the tip-speed ratio, and β is the blade pitch angle.

The tip-speed ratio is defined as:

$$\lambda = \frac{\omega_m R}{V_w} \tag{2}$$

where ω_m is the mechanical rotor speed and R is the turbine radius ($R = 5$ m).

The aerodynamic efficiency of the WT is characterized by the power coefficient $C_p(\lambda, \beta)$, which depends on the tip-speed ratio λ and the pitch angle β. In this study, a semi-empirical analytical approximation of the C_p surface is adopted to enable continuous simulation of aerodynamic behavior [17]:

$$C_p(\lambda, \beta) = (0.5 - 0.00167(\beta - 2)) \sin\left(\frac{\pi(\lambda + 0.1)}{18.5 - 0.3(\beta - 2)}\right) - 0.00184(\lambda - 3)(\beta - 2) \tag{3}$$

This formulation captures the nonlinear dependence of aerodynamic efficiency on both rotor speed and blade pitch angle. It enables realistic modeling of Region-II maximum power extraction and Region-III power limitation through pitch control.

For $\beta = 0°$, the function exhibits a maximum value $C_{p,\max} = 0.5276$ at the optimal tip-speed ratio $\lambda_{\text{opt}} = 9.7205$, which are the parameters used in this study.

In Region-II, the WT operates at the optimal tip-speed ratio λ_{opt} that maximizes the power coefficient:

$$C_p = C_{p,\max}, \quad \lambda = \lambda_{\text{opt}} \tag{4}$$

Under this condition, the wind speed can be expressed as:

$$V_w = \frac{\omega_m R}{\lambda_{\text{opt}}} \tag{5}$$

Substituting (5) into (1) yields:

$$P_{\text{opt}} = \frac{1}{2}\rho\pi R^2 C_{p,\max} \left(\frac{\omega_m R}{\lambda_{\text{opt}}}\right)^3 \tag{6}$$

Rearranging the terms leads to:

$$P_{\text{opt}} = \frac{1}{2}\rho\pi R^5 \frac{C_{p,\max}}{\lambda_{\text{opt}}^3} \omega_m^3 \tag{7}$$

For simplicity, the constant term is defined as:

$$K_{\text{opt}} = \frac{1}{2}\rho\pi R^5 \frac{C_{p,\max}}{\lambda_{\text{opt}}^3} \tag{8}$$

Thus, the reference power used for Maximum Power Point Tracking (MPPT) in Region-II becomes:

$$P_{\text{m,ref}} = K_{\text{opt}}\, \omega_m^3 \tag{9}$$

This formulation corresponds to the classical Optimal Torque Control (OTC) method used in variable-speed WTs. Unlike wind-speed-based MPPT strategies, this approach relies solely on the measured rotor speed, eliminating the need for wind sensors and improving robustness against measurement noise. It also provides a direct cubic relationship between rotor speed and extracted power, ensuring that the WT operates continuously at the maximum power coefficient under below-rated wind conditions.

In this study, the parameters used are $\rho = 1.225$ kg/m^3, $R = 5$ m, $C_{p,\max} = 0.5276$, and $\lambda_{\text{opt}} = 9.7205$, from which K_{opt} is computed according to Eq. (8).

2) Permanent Magnet Synchronous Generator Electromagnetic Modeling: The WT is coupled to a three-phase PMSG modeled in the rotor-oriented dq reference frame using the standard Park transformation. In this study, a surface-mounted (round rotor) machine with sinusoidal back electromotive force (back-EMF) is considered. Under this assumption, the d- and q-axis inductances are equal ($L_d = L_q = L_s$), and magnetic saliency effects are neglected.

The stator voltage equations in the synchronous dq reference frame are expressed as [18]:

$$\begin{aligned} v_d &= R_s i_d + L_s \frac{di_d}{dt} - \omega_e L_s i_q \\ v_q &= R_s i_q + L_s \frac{di_q}{dt} + \omega_e \left(L_s i_d + \lambda_m\right) \end{aligned} \tag{10}$$

where v_d and v_q are the stator voltages, i_d and i_q are the stator currents in the rotating dq frame, R_s is the stator phase resistance, L_s is the synchronous inductance, λ_m is the permanent magnet flux linkage, and ω_e is the electrical angular speed.

The electrical angular speed is related to the mechanical rotor speed by:

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

$$\omega_e = p\,\omega_m \tag{11}$$

where p denotes the number of pole pairs and ω_m is the mechanical angular speed of the rotor.

For a non-salient surface-mounted rotor ($L_d = L_q$), the electromagnetic torque expression simplifies to:

$$T_e = \frac{3}{2} p \lambda_m i_q \tag{12}$$

Since magnetic saliency is absent, the reluctance torque component is zero, and the electromagnetic torque depends solely on the q-axis current. This property enables direct torque control through regulation of i_q.

The mechanical dynamics of the generator shaft are governed by:

$$J\frac{d\omega_m}{dt} = T_m - T_e - F\omega_m \tag{13}$$

where J represents the total rotational inertia of the turbine–generator assembly, F is the viscous friction coefficient, and T_m is the mechanical torque supplied by the WT.

The mechanical torque supplied by the WT is derived from the aerodynamic power as:

$$T_m = \frac{P_{\text{aero}}}{\omega_m} \tag{14}$$

In Region-II maximum power operation, where $P_{\text{aero}} = K_{\text{opt}}\omega_m^3$, the mechanical torque becomes:

$$T_m = K_{\text{opt}}\omega_m^2 \tag{15}$$

Substituting (12) and (15) into (13) leads to the coupled nonlinear electromechanical model:

$$J\frac{d\omega_m}{dt} = K_{\text{opt}}\omega_m^2 - \frac{3}{2} p \lambda_m i_q - F\omega_m \tag{16}$$

The rotor position is obtained from:

$$\frac{d\theta_m}{dt} = \omega_m \tag{17}$$

The main parameters of the PMSG used in the simulation model are summarized in Table I.

TABLE I. PARAMETERS OF THE PMSG USED IN THE WT MODEL.

Parameter	Symbol	Value
Stator phase resistance	R_s	0.05 Ω
Synchronous inductance	L_s	1.55 mH
Permanent magnet flux linkage	λ_m	0.98 Wb
Number of pole pairs	p	12
Total inertia	J	0.01993 kg·m^2
Viscous damping coefficient	F	0.0002 N·m·s

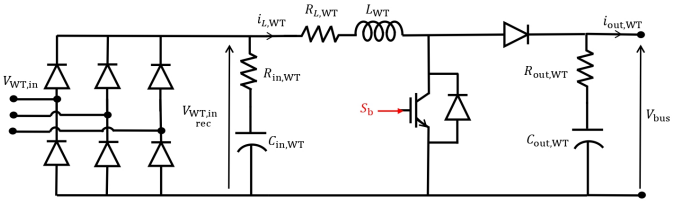


Figure 3. WT AC–DC–DC converter topology.

3) AC–DC–DC Converter: The three-phase output of the PMSG is rectified through a diode bridge and interfaced with the DC–bus using a DC–DC boost converter as shown in Figure 3. Boost converters are commonly used in RE applications to raise the input voltage generated by RESs, particularly when interfacing low-voltage generators such as WTs with a higher-voltage DC–bus. In this study, the DC–bus voltage reference is set to 1000 V, which is a typical design choice in medium-power DC microgrids and industrial DC link systems to ensure compatibility with grid-connected inverters and battery storage systems [19][20].

The three-phase output of the PMSG is first rectified through a diode bridge, producing an unregulated DC voltage that varies with rotor speed. The boost converter then elevates this rectified voltage to the regulated DC–bus level, enabling efficient power transfer under fluctuating wind conditions.

Assuming continuous conduction mode (CCM), the steady-state voltage conversion ratio of the boost converter is given by:

$$\frac{V_{\text{bus}}}{V_{\text{WT, in rec}}} = \frac{1}{1-D} \tag{18}$$

where D is the duty cycle. This relationship shows that the DC–bus voltage can be regulated by adjusting the duty cycle, which is generated by the MPPT-based control loop described in Section III.

The choice of the boost topology is motivated by its simplicity, high efficiency, and reduced component count compared to isolated or multi-stage alternatives [21]. Furthermore, the 10 kHz switching frequency adopted in this study represents a compromise between switching losses and current ripple minimization. The inductance L_{WT} and DC–bus capacitance C_{bus} are selected to ensure stable operation with limited current ripple and acceptable voltage deviation under dynamic wind and load variations.

Assuming negligible mechanical losses and approximates generator efficiency through a constant factor, the electrical power delivered at the rectifier output is expressed as:

$$P_{\text{elec}} = \eta_g\, P_{\text{aero}} \tag{19}$$

where η_g represents the generator efficiency.

The main parameters of the WT boost converter used in the simulation are summarized in Table II.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

TABLE II. PARAMETERS OF THE WT BOOST CONVERTER.

Parameter	Symbol	Value
Input capacitance (rectifier output)	$C_{\text{in,WT}}$	100 μF
Output capacitance (boost output)	$C_{\text{out,WT}}$	500 μF
Input inductance	L_{WT}	1 mH
Inductor resistance	$R_{L,\text{WT}}$	0.01 Ω

B. Battery Energy Storage System

1) Battery Modeling: Due to the intermittent nature of wind energy, the DC microgrid requires an auxiliary energy storage element to ensure power balance and voltage stability. The BESS acts as a dynamic buffer between generation and load, absorbing surplus power during high wind periods and supplying energy when the wind generation is insufficient to meet the thermal demand of the SWHPs.

A lithium-ion BESS is integrated into the DC microgrid to compensate for wind power fluctuations and maintain DC–bus voltage stability. The battery is modeled using the generic lithium-ion model available in Simulink, which captures the nonlinear voltage–State-of-Charge (SOC) characteristics and internal resistance effects.

The battery operates between a fully charged voltage of 407.4 V and a cut-off voltage of 262.5 V. The nominal voltage is 350 V with a rated capacity of 100 Ah, corresponding to an energy storage capacity of 35 kWh.

The main parameters of the BESS used in the simulation are summarized in Table III.

TABLE III. PARAMETERS OF THE LITHIUM-ION BESS.

Parameter	Symbol	Value
Nominal voltage	V_{nom}	350 V
Rated capacity	Q_{nom}	100 Ah
Internal resistance	R_{int}	0.035 Ω
Fully charged voltage	V_{full}	407.4 V
Cut-off voltage	V_{cut}	262.5 V

2) DC-DC Bidirectional Converter Modeling: To ensure stable operation of the DC–bus, a buck-boost Bidirectional Converter (BDC) is integrated to interface the BESS with the DC microgrid. BDCs are pivotal components for energy management in RE applications, as they enable controlled charge and discharge operation of the battery while regulating the DC–bus voltage. During periods of excess wind generation, the converter operates in buck mode, transferring energy from the DC–bus to the battery. Conversely, when wind generation is insufficient to meet the load demand, the converter operates in boost mode, injecting energy from the battery into the DC–bus. Therefore, their integration is essential for maximizing the utilization of RESs, as they ensure a consistent and reliable power supply.

In the context of HPs, a few studies are found in the literature, including a BESS [22]. The BDC topology adopted in this study is illustrated in Figure 4 [23].

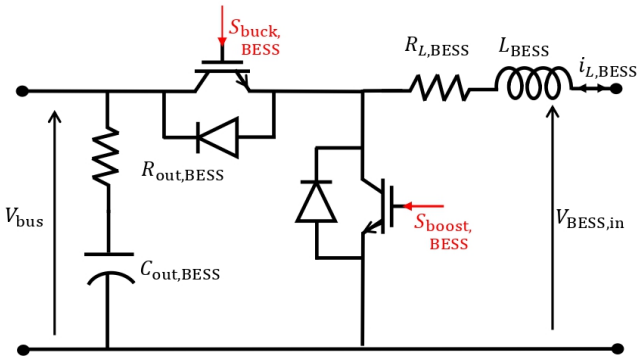


Figure 4. Bidirectional DC–DC converter topology interfacing the BESS.

The main parameters of the BESS BDC used in the simulation are summarized in Table IV.

TABLE IV. PARAMETERS OF THE BESS BDC CONVERTER.

Parameter	Symbol	Value
Output capacitance (boost output)	$C_{\text{out,BESS}}$	500 μF
Input inductance	L_{BESS}	1 mH
Inductor resistance	$R_{L,\text{BESS}}$	0.01 Ω

The equations governing the working principle of the buck and boost modes of this BDC are presented in Eq. (20) [24].

$$\begin{cases} \dfrac{di_{L,\text{BESS}}}{dt} = \dfrac{V_{\text{BESS,in}} - S_{\text{boost,BESS}}\, V_{\text{bus}}}{L_{\text{BESS}}} \\ \dfrac{dV_{\text{bus}}}{dt} = \dfrac{S_{\text{buck,BESS}}\, i_{L,\text{BESS}} - i_{\text{bus}}}{C_{\text{out,BESS}}} \end{cases} \tag{20}$$

where $i_{L,\text{BESS}}$ denotes the inductor current exchanged with the battery (positive during discharge), V_{bus} is the DC–bus voltage, $S_{\text{boost,BESS}}$ and $S_{\text{buck,BESS}}$ represent the binary switching states of the BDC generated by the control algorithm described in further details in Section III, L_{BESS} and $R_{L,\text{BESS}}$ represent the inductance and its parasitic resistance, respectively, and $C_{\text{out,BESS}}$ is the DC–bus side capacitance associated with the BESS converter.

C. Surface Water Heat Pump

In this model, two identical SWHPs (CIAT DYNACIAT LG 300 A water-to-water HP) are used as parallel loads, providing a nominal heat capacity of 90.3 kW, a cooling capacity of 61.5 kW, and a rated electrical consumption of 29.4 kW.

The primary inputs of the model were the thermal power generated by the SWHPs (Figure 5(a)) and its coefficient of performance (COP) (Figure 5(b)), based on measurements from an installed system at an operational AE harvesting-site in Dijlemolens, Belgium. This site also uses a backup gas boiler to satisfy the total building heat demand. The data covers the period from March 7, 2024 at 11:00 AM to May 7, 2024 at 2:00 AM, with a sample time of five seconds.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

The electrical power consumption for the two SWHPs is calculated based on the thermal load profile of the building, as well as the SWHP's technical datasheet. It is given by Eq. (21):

$$P_{\text{e,SWHP}} = \frac{P_{\text{th,SWHP}}}{COP} \tag{21}$$

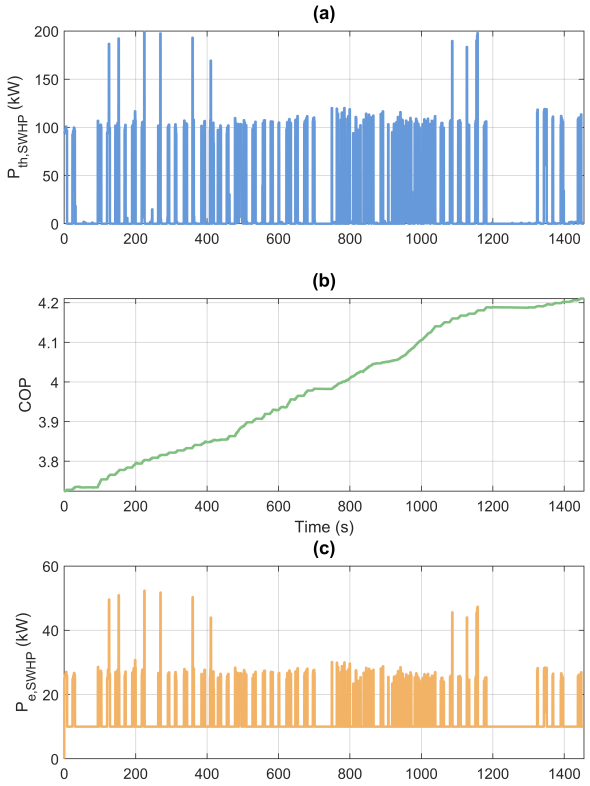


Figure 5. (a) Thermal output power, (b) COP, and (c) electrical consumption of the SWHPs over the simulation period.

As shown in Figure 5, the thermal load profile exhibits frequent peaks in high demand and varying on/off cycles, reflecting realistic building heating demand. Operating in a fully demand-driven mode, these dynamics emphasize the importance of a rapid control response from the supply side and a robust power balance between generation, storage, and load. During periods of low thermal demand, a minimum electrical power of 10 kW is imposed to account for auxiliary electrical equipment and standby operation.

III. Control Strategy of the DC Microgrid

The proposed control framework is designed to regulate the DC–bus voltage, manage bidirectional battery power flow, and ensure safe WT operation across the different operating regions.

A. Wind Turbine Control Strategy

1) Region-Based Power Control: Based on the aerodynamic and electromechanical models presented in Section II, the WT control architecture is structured into three hierarchical layers: (i) operating region management and power reference generation, (ii) boost converter current regulation, and (iii) pitch angle control in the above-rated regime [25].

The WT operates in three distinct regions defined by wind speed and rotor speed conditions, as illustrated in Figure 6.

Region-I (Below Cut-in): For wind speeds $v \leq 3$ m/s, power production is inhibited, and the reference power is set to zero. This prevents inefficient operation under very low wind conditions.

Region-II (Maximum Power Point Tracking): For wind speeds above cut-in ($V_{\text{cut-in}} = 3$ m/s) and rotor speeds below the rated value $\omega_{\text{rated}} = 21.6$ rad/s (corresponding to the rated wind speed $v_{\text{rated}} = 11.1$ m/s), the turbine operates in maximum power point tracking (MPPT) mode.

In this region, the reference mechanical power $P_{\text{m,ref}}$ follows the optimal torque principle as defined in Eq. (9), ensuring operation at the optimal tip-speed ratio and maximum aerodynamic efficiency. Power regulation is achieved exclusively through generator torque control, while the pitch angle is maintained at $\beta = 0°$ to maximize the aerodynamic power coefficient.

Region-III (Above Rated): When the rotor speed exceeds the rated value $\omega_m > 21.6$ rad/s, the extracted power is limited to the rated turbine power $P_{\text{rated}} = 35$ kW. In this operating regime, the control objective shifts from maximum energy capture to power and speed regulation in order to protect the generator and power electronics.

For wind speeds exceeding the cut-out threshold (20 m/s), the turbine is safely shut down and power production is set to zero to ensure safe operation.

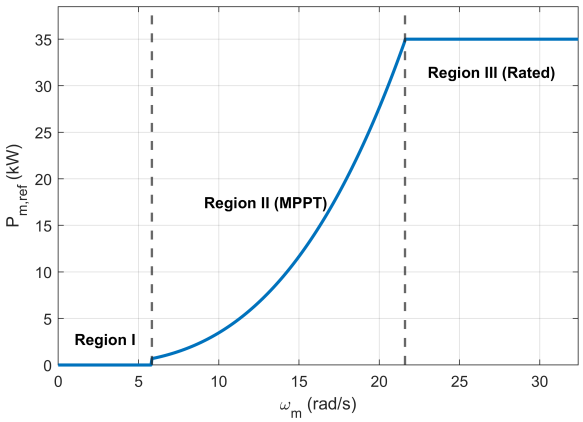


Figure 6. WT operating regions and corresponding power reference as a function of rotor speed.

In Region-III, rotor speed regulation is achieved through blade pitch control. The pitch controller, displayed in Figure 7, is activated when the rotor speed exceeds ω_{rated}, and the speed error is defined as:

$$e_\omega = \omega_m - \omega_{\text{rated}} \tag{22}$$

A proportional–integral (PI) controller generates the pitch β, constrained within the mechanical limits of the turbine:

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

$$0^\circ \leq \beta \leq 35^\circ \tag{23}$$

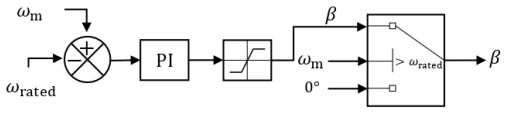


Figure 7. Pitch-based rotor speed regulation.

Increasing the pitch angle reduces the aerodynamic power coefficient $C_p(\lambda, \beta)$, thereby limiting the aerodynamic torque T_m and stabilizing the rotor speed around its rated value. Defining the control error as in Eq. (22) ensures that a positive overspeed condition produces a positive pitch command, enabling effective power limitation in the above-rated region.

2) Boost Converter Current Regulation: The electrical power reference is translated into an inductor current reference for the WT boost converter. Assuming that the mechanical power reference is approximately equal to the electrical power reference (i.e., $P_{\mathrm{m,ref}} \approx P_{\mathrm{elec,ref}}$), the current reference can be obtained from the power–voltage relationship as:

$$i_{L,\mathrm{WT,ref}} = \frac{P_{\mathrm{elec,ref}}}{V_{\mathrm{WT,in\ rec}}} \tag{24}$$

As shown in Figure 8, an inner current control loop regulates the measured inductor current i_L using a PI controller. The control error is defined as:

$$e_i = i_{L,\mathrm{WT,ref}} - i_{L,\mathrm{WT}} \tag{25}$$

The PI controller generates the duty cycle of the boost converter, ensuring fast dynamic tracking of the reference power. Current saturation and voltage-based enabling mechanisms are included to guarantee safe converter operation under transient conditions.

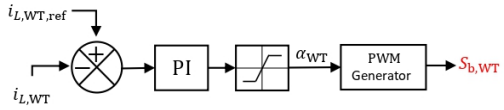


Figure 8. Inductor current control loop of the WT boost converter.

B. BESS Control Strategy

The control of the battery-side BDC is based on maintaining the DC–bus voltage at its nominal value of 1000 V while managing the direction of power flow (buck or boost mode) according to the instantaneous power balance between the renewable source and the load. To achieve this objective, a cascaded dual-loop control structure is implemented, combining DC–bus voltage regulation with bidirectional current control.

An outer voltage control loop regulates the measured DC–bus voltage V_{bus} around its reference value $V_{\mathrm{bus,ref}}$. The corresponding voltage error is defined as:

$$e_v = V_{\mathrm{bus,ref}} - V_{\mathrm{bus}} \tag{26}$$

A PI controller processes this error and generates the reference battery current $i_{\mathrm{BESS,ref}}$, which determines whether the battery should charge or discharge.

An inner current control loop then regulates the measured inductor current $i_{\mathrm{BESS,in}}$ according to:

$$e_i = i_{L,\mathrm{BESS,ref}} - i_{L,\mathrm{BESS}} \tag{27}$$

The current PI controller produces the duty cycle command α_{BESS}, which is forwarded to the converter mode management logic. The complete control architecture is illustrated in Figure 9.

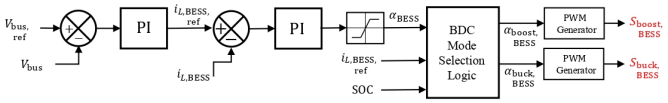


Figure 9. PWM control signals for buck and boost switches for the BESS BDC.

The bidirectional operating mode is selected based on the sign of $i_{L,\mathrm{BESS,ref}}$ and the battery state of charge $\mathrm{SOC}(t)$, as detailed in the flowchart of Figure 10. When $i_{L,\mathrm{BESS,ref}} \geq 0$ and $\mathrm{SOC}(t) > \mathrm{SOC}_{\min}$, the converter operates in boost mode (battery discharging). In this case, the corresponding duty cycles are assigned as:

$$\alpha_{\mathrm{boost,BESS}} = \alpha_{\mathrm{BESS}}, \qquad \alpha_{\mathrm{buck,BESS}} = 0 \tag{28}$$

Conversely, when $i_{L,\mathrm{BESS,ref}} < 0$ and $\mathrm{SOC}(t) < \mathrm{SOC}_{\max}$, the converter operates in buck mode (battery charging), with:

$$\alpha_{\mathrm{boost,BESS}} = 0, \qquad \alpha_{\mathrm{buck,BESS}} = 1 - \alpha_{\mathrm{BESS}} \tag{29}$$

If neither condition is satisfied, both duty cycles are set to zero to prevent operation outside the allowable SOC limits.

This hierarchical control structure ensures stable DC–bus voltage regulation, proper bidirectional power flow, and safe battery operation within predefined SOC constraints, limited to 20% and 80% for $\mathrm{SOC}_{\min}$ and $\mathrm{SOC}_{\max}$, respectively, to preserve battery lifetime and operational safety.

IV. RESULTS | DISCUSSION

The system configuration previously defined in Figure 1 is simulated in MATLAB/Simulink to validate the architecture under realistic operating conditions.

Figure 11 illustrates the evolution of the blade pitch angle β and the aerodynamic power coefficient C_p over the entire simulation period (1456 s). During Region-II operation (below rated wind conditions), the pitch angle remains at $\beta = 0^\circ$, and the power coefficient stabilizes around its optimal value $C_{p,\max} \approx 0.53$, confirming proper MPPT behavior. In this regime, power regulation is achieved solely through generator torque control without aerodynamic intervention.

When the wind speed exceeds its rated value ($v_{\mathrm{rated}} = 11.1$ m/s), corresponding to a rotor speed threshold of $\omega_{\mathrm{rated}} = 21.6$ rad/s, the pitch controller is activated. The pitch angle increases progressively, reaching values up to approximately

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

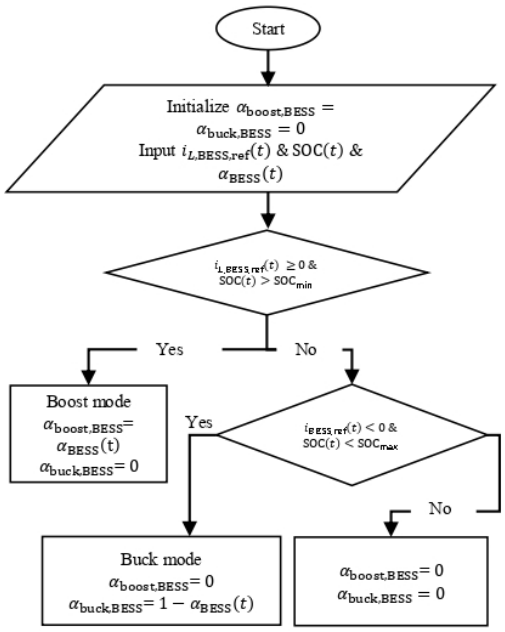


Figure 10. Flowchart of the BDC mode selection logic for the BESS.

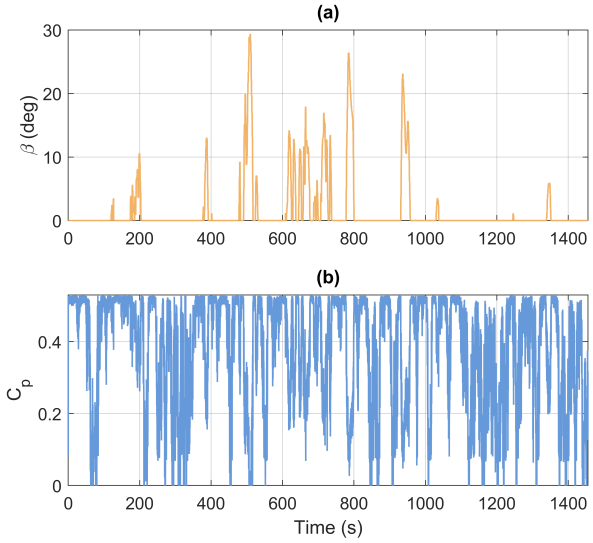


Figure 11. Evolution of (a) blade pitch angle β and (b) aerodynamic power coefficient C_p during WT operation over the simulation period.

29.37° during high wind events. As a consequence, the aerodynamic power coefficient decreases from its optimal value to below $C_p \approx 0.35$, effectively reducing the aerodynamic torque and limiting the mechanical power transferred to the generator.

Under extreme wind conditions, the effectiveness of the pitch-based regulation and the cut-out protection mechanism is clearly demonstrated. At the maximum recorded wind speed of 20.83 m/s (at $t = 509$ s), the pitch controller is strongly activated, with the blade angle reaching values close to its upper operating range (up to $\beta = 29.37°$). As a result, the aerodynamic power coefficient decreases significantly, dropping to $C_p = 0.0414$ and approaching zero during peak pitch activation. This reduction in C_p confirms that the WT is effectively de-loaded aerodynamically under high wind conditions, substantially limiting the extracted mechanical power.

Moreover, once the wind speed exceeds the cut-out threshold of 20 m/s, the electrical reference power is automatically forced to zero ($P_{e,ref} = 0$ W), validating the correct activation of the shutdown logic. This two-stage protection mechanism—progressive aerodynamic de-loading through pitch regulation followed by complete shutdown above cut-out—ensures safe WT operation while preventing excessive energy conversion and generator overloading in severe wind events.

Figure 12 presents the closed-loop tracking performance of the WT current loop, the BESS current loop, and the DC–bus voltage over the entire simulation horizon.

For the WT inductor current, the peak value reaches 91.6 A while the root-mean-square (RMSE) tracking error remains limited to 6.04 A, corresponding to approximately 6.6% of the peak current. The larger instantaneous deviations occur during rapid power transitions, whereas steady-state tracking remains accurate and well damped.

The BESS current loop exhibits stable bidirectional regulation during both charging and discharging phases. The peak current reaches 138.16 A, with a maximum instantaneous error of 53.59 A. The RMSE is 7.85 A (approximately 5.7% of the peak value), indicating consistent current control despite dynamic power exchange within the microgrid.

At the system level, the DC–bus voltage is tightly regulated around the nominal 1000 V. The voltage varies between 993.78 V and 1003.47 V, yielding a peak-to-peak deviation of 9.70 V (0.97%). The maximum absolute voltage deviation from its reference remains below 0.62%, while the RMS voltage error is only 0.332 V, corresponding to 0.033% of nominal voltage, demonstrating robust voltage stability under fluctuating wind input and high thermal demand.

The SOC trajectory, displayed in Figure 13, remains well within the predefined operational limits of 20% and 80% throughout the simulation period. The SOC varies between 69.06% and 71.55%, with an average value of 70.23%, corresponding to a total swing of approximately 2.5%. This limited variation indicates that the wind generation and load demand remain largely balanced over the studied period, while the BESS effectively compensates short-term power mismatches. Moreover, the absence of boundary saturation confirms that the proposed energy management strategy successfully prevents deep discharge and overcharge conditions, thereby ensuring safe battery operation and enhanced lifetime.

Figure 14 presents the instantaneous power exchange at the DC–bus between the WT, BESS, and the two SWHPs. The BESS power is considered positive when discharging to the DC–bus and negative during charging.

The BESS operates fully bidirectionally, with a maximum discharge power of 51.38 kW and a maximum charging

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

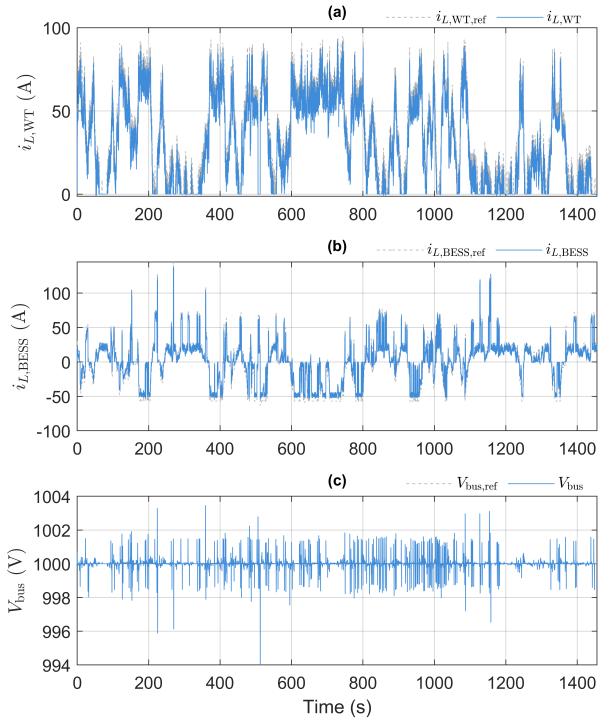


Figure 12. Closed-loop control performance: (a) $i_{L,\mathrm{WT}}$ tracking, (b) $i_{L,\mathrm{BESS}}$ tracking, and (c) V_{bus} regulation.

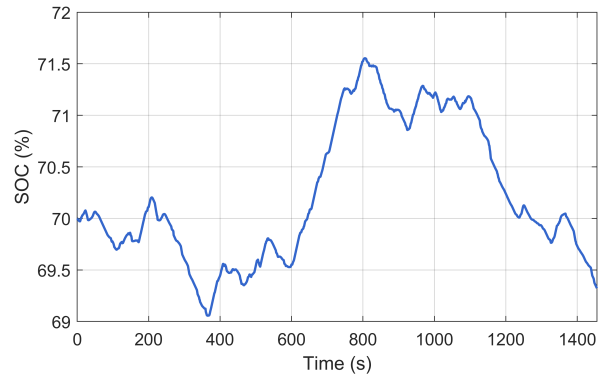


Figure 13. SOC evolution of the BESS over the simulation period.

power of 21.29 kW. The converter remains in discharge mode during 50.55% of the simulation time and in charge mode during 49.45%, indicating balanced alternating power buffering behavior driven by wind-load mismatch dynamics.

The instantaneous DC–bus power balance was evaluated as

$$P_{\mathrm{WT}} + P_{\mathrm{BESS}} - P_{\mathrm{SWHPs}}$$

The mean absolute power imbalance remains limited to 2.46 kW (approximately 4.7% of the peak load), confirming that short-term mismatches are effectively absorbed by the DC–bus capacitor without compromising system stability. This result is consistent with the previously observed voltage deviation below 1%, demonstrating dynamic consistency of the model.

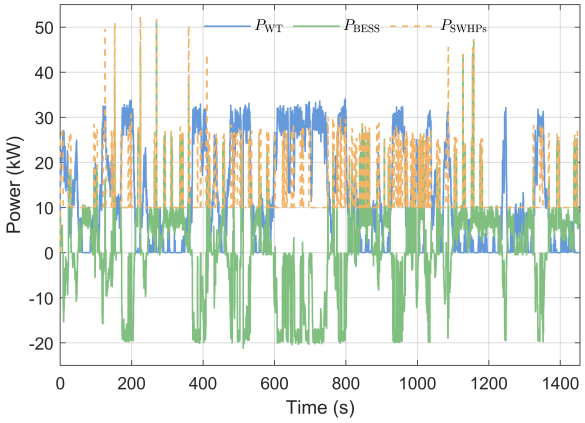


Figure 14. DC–bus power balance between the WT, BESS, and SWHPs over the simulation period.

A critical operating condition occurs at $t = 225$ s, where the SWHPs demand reaches its maximum value of 52.32 kW. At this instant, the WT supplies only 6.33 kW, while the BESS delivers 42.80 kW to the DC–bus. The combined generation therefore reaches 49.13 kW, resulting in a transient deficit of 3.19 kW (approximately 6.1% of the peak load). This mismatch is temporarily compensated by the DC–bus capacitance, confirming the robustness of the coordinated control strategy.

A complementary operating condition occurs at $t = 512.40$ s, corresponding to the maximum battery charging event. At this instant, the WT produces 29.85 kW while the load demand is only 10.00 kW, resulting in a surplus of 19.85 kW. The BESS absorbs 21.29 kW, thereby preventing renewable curtailment and maintaining DC–bus stability. The instantaneous residual is limited to 1.44 kW (approximately 4.8% of the WT power), confirming effective transient buffering by the DC–bus capacitance. This event validates the proper activation of the buck operating mode and demonstrates the ability of the BESS to capture excess wind energy under low-demand conditions.

The selected BESS capacity therefore provides sufficient power support during peak demand events while maintaining operation within a narrow SOC range, thereby preserving battery lifetime and preventing deep charge/discharge cycling. The limited SOC variation observed during the simulation confirms that the system operates in a dynamic buffering regime, where the BESS primarily compensates short-term power imbalances rather than serving as long-term energy storage.

Overall, the results demonstrate that the proposed WT–BESS architecture ensures stable dynamic power sharing, enforces rated WT power limitation, and maintains supply continuity under both wind deficit and high-demand conditions. Thus, the proposed configuration enables simulated fully RE-powered operation under the studied operating conditions, without reliance on auxiliary fossil-fuel backup.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

V. Conclusion and Future Work

This paper presented the design and validation of a DC-coupled hybrid WT–BESS–SWHP microgrid based on realistic climatic data from Le Cano Ouistreham, France pilot site and thermal demand data from an operational aquathermal installation in Dijlemolens, Belgium. A coordinated hierarchical control strategy combining Region-II MPPT, refined Region-III pitch regulation, and cascaded voltage–current control loops was implemented to ensure stable demand-driven operation under fluctuating wind conditions.

Simulation results demonstrate robust system performance under dynamic operating scenarios. The DC–bus voltage remained tightly regulated around its nominal value of 1000 V, with a maximum deviation limited to 0.62%, and an RMSE of only 0.033%. During the most critical operating condition, corresponding to a peak thermal demand of 52.32 kW, coordinated power sharing between the 35 kW WT and the 35 kWh BESS ensured supply continuity without violating operational constraints. The BESS delivered up to 51.38 kW during discharge events while maintaining the SOC within a narrow 2.5% swing (69.06%–71.55%), confirming effective short-term buffering and safe operation.

The results validate the proposed DC microgrid architecture as a realistic and scalable framework for renewable-powered aquathermal heating systems. Future work will focus on the investigation of advanced DC–DC converter topologies, along with corresponding advanced control strategies, to further improve dynamic response, power quality, and overall system efficiency of the proposed DC microgrid.

ACKNOWLEDGMENT

This work was carried out at the GREAH laboratory, University of Le Havre Normandie, particularly by the "Renewable Energies and Storage Systems" (MERS) team, within the framework of the Interreg WaterWarmth project, co-funded by the European Union. The authors would like to thank the European Union and the University of Le Havre Normandie for their financial support.

References

[1] J. Assaf, M.-B. Camara, D. Guilbert, and B. Dakyo, "Energy management of a wind–battery powered surface water heat pump system", in *Proceedings of the Tenth International Conference on Green Communications, Computing and Technologies (GREEN 2025)*, Barcelona, Spain: IARIA, Oct. 2025, pp. 16–21, ISBN: 978-1-68558-311-8.

[2] J. D. Spitler and M. S. Mitchell, "Surface Water Heat Pump Systems", in *Advances in Ground-Source Heat Pump Systems*, Elsevier, 2016, pp. 225–246.

[3] E. Brasser, "Heating Houses Using Surface Water: A Sustainable Alternative to Natural Gas", Ph.D. dissertation, Delft University of Technology, Delft, Netherlands, 2020.

[4] International Renewable Energy Agency (IRENA), "Regional Energy Transition Outlook: European Union", IRENA, Abu Dhabi, Report, 2025.

[5] A. Żelazna and A. Pawłowski, "Review of the Role of Heat Pumps in Decarbonization of the Building Sector", *Energies*, vol. 18, no. 13, 2025, ISSN: 1996-1073. DOI: 10.3390/en18133255.

[6] Eurostat, "Electricity from renewable sources reaches 47% in 2024", 2025, Accessed: May 6, 2026. [Online]. Available: https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20250319-1.

[7] D. Rekioua, "Hybrid Renewable Energy Systems Overview", in *Hybrid Renewable Energy Systems: Optimization and Power Management Control*, Springer, 2019, pp. 1–37.

[8] N. Bayati, A. Hajizadeh, and M. Soltani, "Protection in DC Microgrids: A Comparative Review", *IET Smart Grid*, vol. 1, no. 3, pp. 66–75, 2018. DOI: 10.1049/iet-stg.2018.0024.

[9] S. Baraskar, D. Günther, J. Wapler, and M. Lämmle, "Analysis of the Performance and Operation of a Photovoltaic-Battery Heat Pump System Based on Field Measurement Data", *Solar Energy Advances*, vol. 4, p. 100 047, 2024. DOI: 10.1016/j.seadv.2023.100047.

[10] A. V. Klokov, A. S. Tutunin, E. S. Sharaborova, A. A. Korshunov, and E. Y. Loktionov, "Inverter Heat Pumps as a Variable Load for Off-Grid Solar-Powered Systems", *Energies*, vol. 16, no. 16, p. 5987, 2023, ISSN: 1996-1073. DOI: 10.3390/en16165987.

[11] M. Bojić, N. Nikolić, D. Nikolić, J. Skerlić, and I. Miletić, "Toward a Positive-Net-Energy Residential Building in Serbian Conditions", *Applied Energy*, vol. 88, no. 7, pp. 2407–2419, 2011, ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2011.01.015.

[12] C. Roselli, M. Sasso, and F. Tariello, "Dynamic Simulation of a Solar Electric Driven Heat Pump for an Office Building Located in Southern Italy", *International Journal of Heat and Technology*, vol. 34, no. S2, S496–S504, 2016, ISSN: 0392-8764. DOI: 10.18280/ijht.34S243.

[13] L. Yang, N. Tai, C. Fan, and Y. Meng, "Energy Regulating and Fluctuation Stabilizing by Air Source Heat Pump and Battery Energy Storage System in Microgrid", *Renewable Energy*, vol. 95, pp. 202–212, 2016, ISSN: 0960-1481. DOI: 10.1016/j.renene.2016.04.026.

[14] C. B. N. Fapi, M. L. Touré, M.-B. Camara, and B. Dakyo, "Control Strategy for DC Micro-Grids in Heat Pump Applications with Renewable Integration", *Electronics*, vol. 14, no. 1, p. 150, 2025, ISSN: 2079-9292. DOI: 10.3390/electronics14010150.

[15] C. B. N. Fapi, M. L. Touré, M.-B. Camara, and B. Dakyo, "MPPT Based Fractional Short-Circuit Current-Model Predictive Control for PV System in Real Weather Conditions for Heat-Pump Applications", in *Proceedings of the 2024 International Conference on Intelligent Systems and Computer Vision (ISCV 2024)*, IEEE, 2024, pp. 1–6. DOI: 10.1109/ISCV61718.2024.10694395.

[16] C. B. N. Fapi, M. L. Touré, M.-B. Camara, and B. Dakyo, "High Voltage Gain DC-DC Converter Based Maximum Power Tracking from Photovoltaic Systems for Heat-Pump Applications", in *Proceedings of the 2024 12th International Conference on Smart Grid (icSmartGrid 2024)*, IEEE, 2024, pp. 493–498. DOI: 10.1109/icSmartGrid61521.2024.10557289.

[17] B. Singh and S. Sharma, "Stand-alone wind energy conversion system with an asynchronous generator", *Journal of Power Electronics*, vol. 10, no. 5, pp. 538–547, 2010, ISSN: 1598-2092.

[18] E. Chavero-Navarrete, M. Trejo-Perea, J.-C. Jáuregui-Correa, R.-V. Carrillo-Serrano, and J.-G. Rios-Moreno, "Pitch angle optimization by intelligent adjusting the gains of a PI controller for small wind turbines in areas with drastic wind speed changes", *Sustainability*, vol. 11, no. 23, p. 6670, 2019, ISSN: 2071-1050. DOI: 10.3390/su11236670.

[19] K. B. Samal, S. Pati, and R. Sharma, "Power Management Using an Improved EMS Algorithm in a Stand-Alone Hybrid PV-PEMFC Microgrid with Reduced Converter Count", *Green Energy and Intelligent Transportation*, p. 100 302, 2025, ISSN: 2772-3755. DOI: 10.1016/j.get.2025.100302.

[20] K. Zafar, M. K. Kamaludeen, Y. Esa, A. A. A. Mohamed, and S. Odie, "Fault Analysis for DC Bus-Integrated Energy

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Storage System, Electric Vehicle Supply Equipment, and Photovoltaic Systems", *Electric Power Systems Research*, vol. 234, p. 110 837, 2024, ISSN: 0378-7796. DOI: 10.1016/j.epsr.2024.110837.

[21] J. Assaf, J. S. Menye, M. B. Camara, D. Guilbert, and B. Dakyo, "Power Converter Topologies for Heat Pumps Powered by Renewable Energy Sources: A Literature Review", *Electronics*, vol. 13, no. 19, p. 3965, 2024, ISSN: 2079-9292. DOI: 10.3390/electronics13193965.

[22] C. Lorenzo, L. Narvarte, R. H. Almeida, and A. B. Cristóbal, "Technical Evaluation of a Stand-Alone Photovoltaic Heat Pump System Without Batteries for Cooling Applications", *Solar Energy*, vol. 206, pp. 92–105, 2020, ISSN: 0038-092X. DOI: 10.1016/j.solener.2020.05.051.

[23] M. B. Camara, "Supercapacitors for Dynamic Energy Exchange Onboard Hybrid Electric Vehicles: Modeling, Converter Study, and Control", Ph.D. dissertation, Université de Franche-Comté, France, 2007.

[24] S. Agrawal, L. Umanand, and B. S. Reddy, "Bidirectional Current-Fed Converter for High Gain DC–DC and DC–AC Applications", in *Proc. of the Symposium on Power Electronic and Renewable Energy Systems Control (PERESC 2020)*, ser. Lecture Notes in Electrical Engineering, vol. 731, Singapore: Springer, 2021, pp. 101–111. DOI: 10.1007/978-981-33-4081-9_10.

[25] D. Ikni, M. B. Camara, A. Payman, and B. Dakyo, "Dynamic control of wind energy conversion system", in *Proceedings of the 8th International Conference and Exhibition on Ecological Vehicles and Renewable Energies (EVER 2013)*, IEEE, 2013, pp. 1–5. DOI: 10.1109/EVER.2013.6521561.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Soil Moisture Prediction and Benchmark Reliability: A Review with Multi-Site SMAP–ISMN Analysis

Kamrul Hasan, Cliff Mutegi, Arnold Muiruri, Puneeth Kumar Amudala, and Andrew Njiyo
College of Computing
Grand Valley State University
Allendale, MI, USA
e-mail: {hasanka | mutegic | muiruria | amudalap | njiyoa}@mail.gvsu.edu

***Abstract*—Unsustainable irrigation practices are accelerating groundwater depletion worldwide, making accurate soil moisture monitoring increasingly important for agricultural and water resource management. However, a fundamental challenge remains: the benchmark datasets used to validate soil moisture models often contain uncertainties that are not fully understood. This paper addresses this gap through two complementary contributions. First, we present a structured review of soil moisture prediction approaches, including physics-based land surface models, machine learning methods, and hybrid frameworks. We also provide a critical assessment of commonly used benchmark datasets, including ISMN, SCAN, FLUXNET, and satellite-derived products such as SMAP, SMOS, and ASCAT, together with recent advances in data fusion techniques. Second, we present an empirical case study comparing NASA SMAP satellite estimates with ISMN ground observations at two sites in the United States from January 2023 to January 2025. The results reveal strong site-dependent differences in performance. The Michigan site exhibits reasonable agreement for root-zone soil moisture estimates ($r = 0.63$, RMSE = $0.02\ \mathrm{m}^3/\mathrm{m}^3$), whereas the Indiana site shows near-zero correlation and errors more than five times larger (RMSE = $0.11\ \mathrm{m}^3/\mathrm{m}^3$). Moreover, SMAP retrievals are largely insensitive to the dynamic range of in-situ variability, a finding with important implications for model validation practices. Residual analysis further reveals structured, moisture-dependent bias patterns at both sites, suggesting that environmental factors such as row-crop agriculture, tile drainage systems, and sub-pixel heterogeneity are not adequately captured by current retrieval algorithms. Overall, these findings provide strong evidence against single-benchmark validation approaches and support the adoption of integrated, multi-source evaluation frameworks. We conclude by discussing the implications for precision agriculture, downscaling algorithm development, and the design of future soil moisture monitoring networks.**

***Keywords*-soil moisture prediction; remote sensing; international soil moisture network; data fusion; machine learning.**

I. Introduction

Groundwater levels are declining at an alarming rate across the globe due to various factors, with excessive irrigation practices being one of the primary ones [1][2][3]. According to the 2018 U.S. Census of Agriculture, approximately 50% of irrigated land in the United States depends exclusively on groundwater, while an additional 16% relies on a combination of groundwater and surface water [4]. Moreover, groundwater levels have declined substantially across many U.S. aquifers since 1980, highlighting increasingly unsustainable patterns of water use [5].

To address this growing crisis, it is imperative to optimize agricultural water consumption. An ongoing research project at Grand Valley State University (GVSU), Michigan, conducted under the Precision Agriculture Research Lab, aims to address this challenge. The focus of the project is on predicting soil moisture by integrating data from sparsely distributed in-situ moisture sensors with satellite-based observations, such as NASA's Soil Moisture Active Passive (SMAP) mission [6] and the European Space Agency's Climate Change Initiative (ESA CCI) [7]. An initial study from this project [1] introduced the problem of benchmark dataset reliability in soil moisture prediction and provided a preliminary review of key datasets and their limitations. The present paper substantially extends that earlier work in three respects: it broadens the review to encompass recent advances in machine learning, deep learning, and data fusion techniques; it introduces a formal research framework situating these contributions; and, most critically, it presents an original empirical case study comparing NASA SMAP satellite estimates against ISMN ground observations at two contrasting U.S. sites, providing quantitative evidence for the benchmark reliability concerns raised in [1].

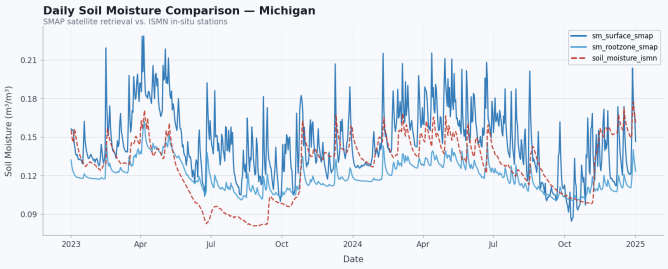


Figure 1. Average daily soil moisture by SMAP (surface-level and root-zone) vs. ISMN at Gaylord-9-SSW (Michigan, U.S.). Null values were imputed through forward-fill (rolling average, window-size = 3).

Soil moisture monitoring and prediction can play a pivotal role not only in minimizing water waste but also in enabling informed decision-making for farmers and policy makers. Effective soil moisture management supports long-term soil health, prevents erosion, and ensures sustained agricultural productivity. In addition to precision agriculture, it enables better drought monitoring, flood forecasting, and land-atmosphere interaction modeling [8][9]. Although soil moisture prediction has been widely investigated, the development of consistent and reliable benchmark datasets remains an ongoing challenge.

Figure 1 presents the aggregated daily average soil moisture

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

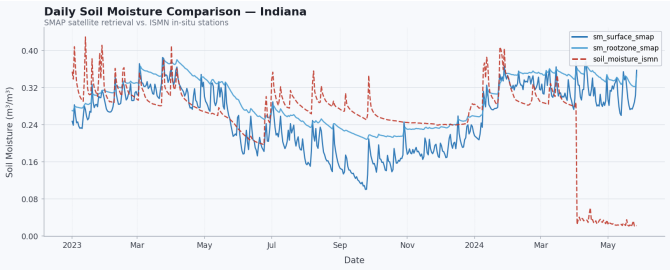


Figure 2. Average daily soil moisture by SMAP (surface-level and root-zone) vs. ISMN at Bedford-5-WNW (Indiana, U.S.). Null values were imputed through forward-fill (rolling average, window-size = 3).

measurements from January 2023 to January 2025 at the Gaylord-9-SSW station in Michigan, USA, an example site within the International Soil Moisture Network (ISMN), a publicly accessible global database that consolidates in-situ soil moisture observations from numerous monitoring networks. By offering standardized data formats and automated quality control protocols, the ISMN serves a vital role in validating satellite-derived soil moisture products and land surface models, and has become a widely adopted reference in hydrological and climate research due to its comprehensiveness and accessibility [10].

To assess the consistency between ground-based and satellite-derived soil moisture measurements, we compare average daily values from NASA's SMAP products with corresponding ISMN data. As illustrated in Figure 1, both datasets exhibit similar seasonal trends, with the primary differences occurring in magnitude rather than overall pattern. A comparable analysis at a second ISMN site, Bedford-5-WNW in Indiana, is shown in Figure 2. In this case, the discrepancy between SMAP and ISMN measurements is more pronounced than at the Gaylord-9-SSW station. These findings raise important questions regarding the reliability of these data as a benchmark for soil moisture modeling: *To what extent can ISMN be trusted for model evaluation? What are its inherent strengths and limitations? And are there viable alternatives that offer improved consistency or coverage?*

This paper addresses these questions through the following contributions, which together extend the preliminary investigation in [1]:

- A structured review of recent advances and key challenges in soil moisture prediction, spanning physics-based models, machine learning, deep learning, and hybrid frameworks.
- A critical evaluation of existing benchmark datasets for soil moisture monitoring and prediction, including ISMN, SCAN, FLUXNET, and satellite-derived products, with particular emphasis on their strengths and limitations in supporting robust model development.
- An original empirical case study comparing SMAP satellite estimates against ISMN observations at two contrasting U.S. sites, providing quantitative, site-specific evidence of the benchmark reliability concerns identified in [1] and motivating the need for multi-source validation frameworks.

The remainder of this paper is organized as follows. Section II presents the research framework guiding the study. Section III discusses recent advances and key challenges in soil moisture prediction. Section IV provides a review of ISMN data, highlighting its strengths, limitations, and applications. Section V examines alternative soil moisture datasets and data fusion approaches. Section VI presents a comparative analysis of SMAP and ISMN soil moisture data for two representative sites introduced in Section I: Gaylord-9-SSW (Michigan, U.S.) and Bedford-5-WNW (Indiana, U.S.). Finally, Section VII summarizes the key findings and provides concluding remarks.

II. Research Framework

To situate the contributions of this paper within a coherent analytical structure, Figure 3 presents an overview of the research framework guiding this work. The framework comprises four interconnected components: a critical review of soil moisture prediction approaches and benchmark datasets, an empirical comparative analysis of SMAP and ISMN observations, a rigorous statistical evaluation of agreement and bias structure, and a synthesis of findings into actionable recommendations for multi-source validation practice. It is worth noting that these components are not strictly sequential; the review informed site selection for the case study, while the empirical results in turn strengthen and substantiate the review's critique of ISMN as a universal benchmark.

Component 1: Literature Review

The first component systematically examines existing soil moisture prediction methodologies across three categories: physics-based land surface models (LSMs) such as the Noah LSM and Community Land Model; machine learning and deep learning approaches including random forests, support vector machines, convolutional neural networks (CNNs), and recurrent neural networks (RNNs); and hybrid frameworks that integrate physical modeling with data-driven learning. Alongside these prediction approaches, this component critically evaluates benchmark datasets used for model training and validation, including ground-based repositories (ISMN, SCAN, FLUXNET), satellite-derived products (SMAP, SMOS, ASCAT), and regional networks such as OzNet and REMEDHUS. Recent advances in data fusion techniques that integrate satellite and in-situ observations are also reviewed to establish the conceptual context within which the empirical case study is situated.

Component 2: Empirical Case Study

The second component presents an original comparative analysis drawing on two years of daily soil moisture observations from January 2023 to January 2025 at two contrasting sites in the United States. The sites were deliberately selected to represent opposing land surface conditions: Gaylord-9-SSW in Michigan, characterized by a comparatively stable, naturally vegetated temperate landscape, and Bedford-5-WNW in Indiana, dominated by intensive row-crop agriculture, artificial tile drainage infrastructure, and spatially variable soil

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

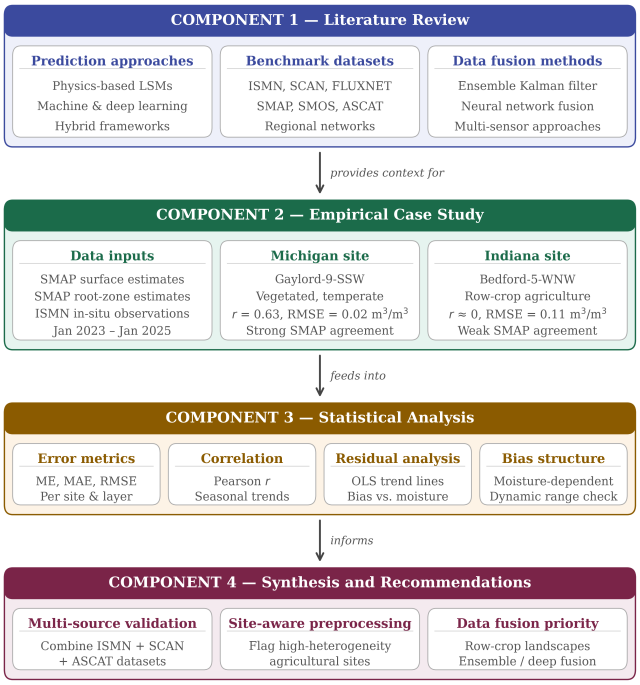


Figure 3. Research framework illustrating the four interconnected components of this study: literature review, empirical case study, statistical analysis, and synthesis of recommendations. Arrows indicate the primary direction of analytical flow; note that Components 1 and 2 also share a feedback relationship, as the review informed site selection and the empirical results substantiate the review's critique of ISMN benchmark reliability.

organic matter. For each site, NASA SMAP satellite-derived estimates at both surface and root-zone layers are compared against corresponding ISMN in-situ measurements, enabling a controlled contrast between a favorable and a challenging retrieval environment.

Component 3: Statistical Analysis

The third component applies a suite of statistical metrics to quantify the agreement between SMAP and ISMN observations across both sites and soil layers. Metrics include mean error (ME), mean absolute error (MAE), root mean square error (RMSE), and Pearson correlation coefficient (r). Critically, the analysis extends beyond aggregate summary statistics to include ordinary least squares (OLS) residual analysis, which examines bias as a function of in-situ moisture magnitude rather than treating error as uniformly distributed. This approach reveals moisture-dependent bias structures that aggregate metrics alone would obscure, and provides the empirical basis for distinguishing systematic retrieval limitations from random measurement noise.

Component 4: Synthesis and Recommendations

The fourth component integrates insights from the review and the case study to evaluate the reliability of ISMN as a universal benchmark and to propose concrete, actionable guidance for practitioners. Specific recommendations address three areas: the adoption of multi-source validation frameworks that combine ISMN with complementary datasets such as SCAN and ASCAT; the application of site-aware preprocessing protocols that flag high-heterogeneity agricultural environments for additional scrutiny; and the prioritization of data fusion approaches for row-crop dominated landscapes where passive microwave retrievals alone are insufficient. These recommendations are grounded directly in the empirical findings of Component 3 rather than derived from general principles alone.

Table I
MAJOR SOIL MOISTURE PREDICTION APPROACHES AND THEIR CHARACTERISTICS.

Method	Strengths	Limitations	Refs.
Physics-based LSMs	Physically interpretable; models water and energy fluxes; widely used	Parameter uncertainty; scale mismatch; computational cost	[11]–[13]
Traditional ML	Captures nonlinear relationships; flexible data fusion; good predictive accuracy	Large training data requirements; limited extrapolation; low interpretability	[14], [15]
Deep Learning	Learns complex spatial and temporal patterns; highly scalable	Data- and computation-intensive; black-box behavior	[16]
Hybrid Physics–ML	Combines physical and data-driven models; improved generalization	Complex integration; limited evaluation frameworks	[17], [18]
Satellite-based Estimation	Global coverage; long-term observations; large-scale monitoring	Coarse resolution; vegetation and atmospheric effects	[19]–[21]

III. SOIL MOISTURE PREDICTION: APPROACHES, ADVANCES, AND OPEN CHALLENGES

Soil moisture monitoring and prediction research has evolved significantly over the past two decades, primarily driven by advances in remote sensing, data assimilation, and machine learning. Traditional approaches generally relied on physics-based land surface models (LSMs), such as the Noah LSM and the Community Land Model (CLM), to simulate water and energy fluxes at the land-atmosphere interface [11][12]. These models use meteorological inputs and land surface parameters, but their performance is often constrained by uncertainties in input data, parameterization, and the scale mismatches between model outputs and observational datasets [13].

Machine Learning (ML) and Deep Learning (DL) methods have recently emerged as powerful alternatives or complements to traditional models. Data-driven algorithms, including random forests, support vector machines, and artificial neural networks, have been employed to estimate soil moisture from remote sensing and meteorological data [14][15]. Deep learning architectures, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have demonstrated strong capabilities in modeling complex spatiotemporal patterns in soil moisture dynamics [16]. Additionally, hybrid approaches that integrate physical modeling with ML have gained attention for improving generalizability and interpretability [17][18]. To illustrate these evolutions, Table I compares the key approaches.

Satellite missions such as NASA's SMAP, ESA's Soil Moisture and Ocean Salinity (SMOS), and the AMSR series

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

have facilitated the development of predictive models at multiple spatial scales, contributing to applications from global hydrological assessment to localized precision farming [19][20]. However, most satellite-derived products are available at coarse spatial resolutions (e.g., 1 km or greater), limiting their usefulness in field-level agricultural decision-making [21].

Despite these technological advancements, several key challenges hinder the development of accurate and reliable soil moisture prediction models. A major issue is the scarcity and spatial sparsity of high-quality ground truth data, which is critical for both model training and validation [22]. The heterogeneity of environmental variables, such as soil properties, vegetation cover, land use patterns, and topography, further complicates model generalization across different regions [23]. Equally critical are the challenges associated with data collection. In-situ soil moisture measurements, such as those provided by ISMN, offer valuable ground truth but are often spatially sparse and unevenly distributed, particularly in under-monitored regions [24]. Variations in sensor type, calibration, and installation practices introduce inconsistencies, while sensor failure or communication issues can lead to temporal gaps. Satellite-based data, while offering broader coverage, are impacted by cloud cover, vegetation, and surface roughness, reducing measurement reliability in many settings [25][26]. Arid and semi-arid regions, where accurate soil moisture monitoring is most crucial, are particularly affected due to low signal-to-noise ratios [27].

Addressing these multifaceted challenges calls for multi-disciplinary strategies involving improved sensor networks, data harmonization, uncertainty quantification, and interpretable modeling frameworks. The integration of adaptive machine learning algorithms with heterogeneous data sources is critical to developing high-resolution, accurate soil moisture predictions that can transform sustainable water resource management and data-driven agriculture.

IV. The International Soil Moisture Network: Capabilities, Limitations, and Applications

The ISMN has emerged as a critical resource for collecting and harmonizing in-situ soil moisture data across global observation networks. It serves as a foundational resource for validating, calibrating, and benchmarking satellite- and model-derived soil moisture datasets. Its importance lies in the harmonized collection and open dissemination of in-situ soil moisture data from a wide array of monitoring networks across different climate zones, land cover types, and soil structures [10][24]. The ISMN enables intercomparison of remote sensing products (e.g., SMAP, SMOS, AMSR2) by providing a global standard against which these data sources can be evaluated [22]. It also supports the assessment and development of downscaling algorithms and machine learning models by offering high-quality ground truth measurements [28]. Moreover, the temporal consistency and metadata richness of ISMN facilitate long-term hydrological studies and trend detection, which are crucial for climate resilience planning and agricultural decision-making. By improving the accuracy and robustness of predictive models, ISMN plays a critical role in the advancement of soil moisture science and its practical applications in water resource management, agriculture, and disaster mitigation.

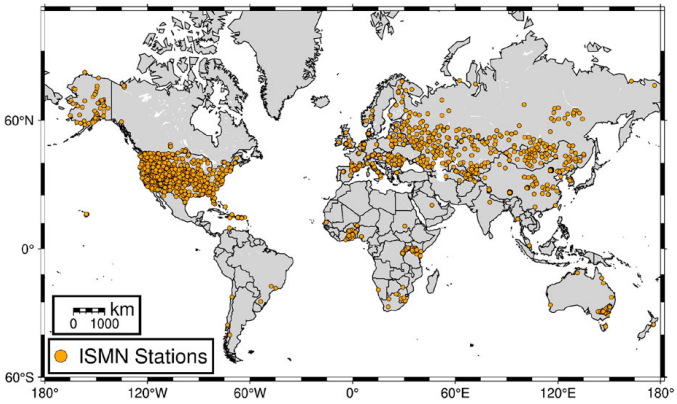


Figure 4. Worldwide distribution of ISMN Stations – an extract from [29].

ISMN aggregates soil moisture measurements from a variety of sources, standardizing and applying quality control procedures to improve accessibility and usability [10]. However, the ISMN data can still exhibit inconsistencies due to differences in sensor types, installation depths, and environmental heterogeneity [22]. Draper et al. emphasized the importance of preprocessing ISMN data before using it for validation or modeling tasks [30].

Despite its value, ISMN presents several challenges when used in predictive soil moisture modeling. The spatial distribution of ISMN stations, as shown in Figure 4, is highly uneven, with denser coverage in North America and Europe and sparse representation in Africa, Asia, and South America. This limits global-scale modeling and regional calibration, especially in underrepresented ecosystems. Station metadata, including soil depth, vegetation, and land use, is sometimes incomplete or inconsistent, complicating efforts to standardize data inputs for machine learning and physical models [22]. Discrepancies also arise from heterogeneity in sensor types, calibration protocols, and measurement depths across networks, introducing uncertainty into inter-station comparisons and satellite validation studies [24]. Moreover, data gaps due to sensor maintenance or environmental interference pose problems for time series continuity. These limitations necessitate pre-processing steps such as harmonization, gap-filling, and filtering, which introduce additional complexity into model development pipelines. Despite these challenges, ISMN remains a critical benchmark for validating satellite retrievals and downscaling methods, though its shortcomings highlight the importance of complementing it with other data sources and standardization frameworks.

The increasing availability of ISMN data has enabled its integration into machine learning and deep learning models for high-resolution soil moisture estimation. Xu et al. [31] used ISMN data to validate a wide and deep neural network that improved the spatial resolution of SMAP satellite data

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

across the U.S. Similarly, Celik et al. [32] and Lee et al. [33] developed deep learning models incorporating ISMN observations to improve performance in heterogeneous landscapes by reducing dependency on physical modeling assumptions. In the agricultural domain, Custódio and Prati [34] applied ensemble machine learning models to IoT-supported irrigation systems, using soil moisture as a key variable. Their results, validated with real-time field data, support the use of AI for operational water resource management.

While the ISMN is the most prominent repository for in-situ soil moisture measurements, several alternative datasets and platforms also play crucial roles in soil moisture research and modeling. One key alternative is the USDA Soil Climate Analysis Network (SCAN), which provides high-resolution, near-real-time soil moisture data across agricultural zones in the United States [35]. Similarly, the FLUXNET network offers point-based data through eddy covariance towers, which include soil moisture as part of broader ecosystem flux measurements [36]. In terms of satellite-derived products, SMAP and ESA's SMOS missions provide global, gridded soil moisture datasets at regular intervals [37]. The Advanced Scatterometer (ASCAT) onboard EUMETSAT MetOp satellites also offers a long-term record of soil moisture estimations with relatively high temporal resolution [38]. Additionally, regional in-situ networks such as the OzNet (Australia), REMEDHUS (Spain), and ARM Southern Great Plains (USA) serve as valuable sources for local model calibration and validation. These alternatives, while often complementary to ISMN, highlight the diversity of data sources available for soil moisture modeling and reinforce the importance of integrated approaches that combine satellite, in-situ, and model-based observations.

V. Alternative Datasets and Data Fusion Strategies for Soil Moisture Monitoring

Recent studies have evaluated the performance of alternative soil moisture datasets relative to the International Soil Moisture Network (ISMN). These datasets provide complementary advantages in spatial coverage, temporal continuity, and measurement accuracy. For example, the USDA Soil Climate Analysis Network (SCAN) often demonstrates lower root mean square error (RMSE), approximately $0.04\ \mathrm{m}^3/\mathrm{m}^3$, in agricultural validation studies across the United States compared to global ISMN averages of approximately $0.06\ \mathrm{m}^3/\mathrm{m}^3$. This improvement is largely attributed to denser station coverage and standardized sensor installations [39]. FLUXNET provides valuable integration of soil moisture measurements with ecosystem flux observations, enabling analysis of land–atmosphere exchanges. However, its global network consists of roughly 200 sites, significantly fewer than the approximately 2200 stations in ISMN, which limits its applicability for high-resolution spatial modeling [40].

Satellite-derived products such as ASCAT offer improved temporal continuity, typically providing daily observations compared with the occasional temporal gaps present in in-situ datasets. Nevertheless, these products operate at coarser spatial resolutions (approximately 25 km), with validation studies reporting unbiased RMSE values around $0.05\ \mathrm{m}^3/\mathrm{m}^3$ when compared with in-situ observations [39]. Regional soil moisture networks, including OzNet and REMEDHUS, can achieve high site-specific accuracy, with validation coefficients of determination (R^2) exceeding 0.85. However, their geographic coverage remains limited, restricting global applicability [40]. Table II summarizes key alternative soil moisture datasets and their complementary roles relative to ISMN.

Table II
Comparison of representative soil moisture datasets relative to ISMN.

Dataset	Coverage	Resolution	Strengths	Limitations
USDA SCAN	U.S.	Point-based	Standardized sensors and rich metadata	Limited geographic coverage
FLUXNET	Global	Point-based	Combines ecosystem flux and soil moisture measurements	Sparse station distribution
SMAP / SMOS	Global	9–36 km	Broad spatial coverage and frequent observations	Coarse resolution and vegetation effects
ASCAT	Global	∼25 km	Long-term record and high temporal frequency	Limited to surface soil moisture
Regional Networks	Local / Regional	Point-based	High-quality site-specific measurements	Limited scalability and coverage

Leveraging these complementary datasets can help mitigate spatial biases in ISMN, particularly in underrepresented environments such as arid and semi-arid regions where satellite products such as ASCAT may provide more reliable temporal coverage [39]. Recent research increasingly focuses on data fusion techniques to integrate satellite observations with in-situ measurements. Approaches such as ensemble Kalman filtering and artificial neural networks have demonstrated substantial improvements in soil moisture estimation accuracy. For instance, integrating SMAP satellite observations with ground measurements has reduced RMSE by approximately 31–50% in semi-arid grassland ecosystems [41].

More advanced techniques incorporating deep learning, ensemble modeling, and multi-sensor satellite fusion have further enhanced spatial resolution and predictive performance [42]–[44]. Emerging distributed learning frameworks, including federated learning approaches, also show promise for soil moisture prediction in large-scale agricultural Internet-of-Things networks [45]. These developments highlight the growing importance of integrating heterogeneous datasets and advanced machine learning techniques to improve soil moisture monitoring and prediction capabilities [46]. These advances align closely with ongoing efforts at the Precision Agriculture Research Lab at GVSU to develop integrated predictive frameworks for precision agriculture applications.

VI. Multi-Site Case Study: SMAP Satellite Estimates Versus ISMN Ground Observations

This section extends the comparative analysis of SMAP and ISMN soil moisture data for the two sites introduced

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Table III
COMPARISON OF STATISTICAL METRICS BETWEEN SMAP AND ISMN SOIL MOISTURE OBSERVATIONS AT THE MICHIGAN AND INDIANA SITES. ME: MEAN ERROR; MAE: MEAN ABSOLUTE ERROR; RMSE: ROOT MEAN SQUARE ERROR; r: PEARSON CORRELATION COEFFICIENT. ALL ERROR METRICS IN m^3/m^3.

Site	Layer	ME	MAE	RMSE	r
Michigan	Surface	0.019	0.024	0.031	0.48
Michigan	Root-zone	-0.009	0.017	0.020	0.63
Indiana	Surface	-0.002	0.077	0.111	-0.01
Indiana	Root-zone	0.039	0.070	0.112	-0.09

in Section I: Gaylord-9-SSW (Michigan, U.S.) and Bedford-5-WNW (Indiana, U.S.). The objective is to evaluate the consistency, bias, and variability between satellite-derived and in-situ measurements across multiple temporal scales, and to identify systematic discrepancies arising from spatial resolution differences, measurement techniques, and environmental heterogeneity.

A. Cross-Site Statistical Overview

Table III summarizes the key statistical metrics for both sites and soil layers. A clear contrast emerges between the two locations: Michigan exhibits strong agreement between SMAP and ISMN observations, particularly for the root-zone product, while Indiana shows substantially weaker agreement, with higher errors and near-zero or negative correlations across both layers.

The site-dependent performance of SMAP visible in Table III is most evident in two contrasts. First, root-zone RMSE at the Indiana site (0.112 m^3/m^3) is more than five times larger than at the Michigan site (0.020 m^3/m^3), indicating that retrieval error is not uniformly distributed across environments but is strongly conditioned by local land surface characteristics. Second, while Michigan yields meaningful positive correlations at both layers ($r = 0.48$ surface, $r = 0.63$ root-zone), Indiana produces near-zero or slightly negative values ($r = -0.01$ surface, $r = -0.09$ root-zone), suggesting that SMAP fails to track day-to-day moisture variability at that site entirely, irrespective of the depth layer considered.

These contrasts are attributable to the distinct land surface characteristics of the two sites [23]. The Indiana site is heavily influenced by extensive row-crop agriculture, highly variable soil organic matter, and the presence of artificial tile drainage infrastructure. Together, these factors introduce sub-pixel spatial heterogeneity that significantly complicates passive microwave moisture retrievals [25][26]. The Michigan site, by contrast, presents a comparatively stable, naturally vegetated landscape that aligns more closely with the parameterizations underlying SMAP's retrieval algorithms, yielding higher correlations and lower overall error [13].

It is important to note that the near-zero correlation observed at Indiana does not imply that SMAP is globally unreliable; rather, it underscores the strong site dependency of satellite soil moisture retrievals and cautions against using a single ground-based benchmark such as ISMN to draw general conclusions about retrieval performance. This distinction motivates the site-specific residual analyses presented next.

B. Michigan Site: Gaylord-9-SSW

For the Michigan site ($n = 732$ daily observations), SMAP and ISMN exhibit strong seasonal consistency, with both datasets capturing wet and dry periods effectively (Figure 5).

At the daily scale, SMAP surface soil moisture shows a positive bias ($ME \approx 0.019$), indicating overestimation relative to ISMN observations. Error levels are moderate ($MAE \approx 0.024$, $RMSE \approx 0.031$). In contrast, the root-zone product demonstrates improved accuracy with lower errors ($MAE \approx 0.017$, $RMSE \approx 0.020$) and a slight negative bias ($ME \approx -0.009$). **Correlation analysis** further confirms stronger agreement for the root-zone product ($r \approx 0.63$) compared to the surface product ($r \approx 0.48$), suggesting that deeper-layer estimates better capture soil moisture variability. **Residual analysis** (Figures 6 and 7) indicates that surface estimates exhibit greater variability and occasional extreme deviations, whereas root-zone residuals are more stable and tightly distributed. Temporal aggregation (Figure 5) reduces high-frequency variability and further improves agreement, indicating that discrepancies are partly driven by daily-scale variability and spatial mismatch. Overall, SMAP performs well at this site, particularly for root-zone soil moisture, demonstrating strong capability in capturing temporal dynamics.

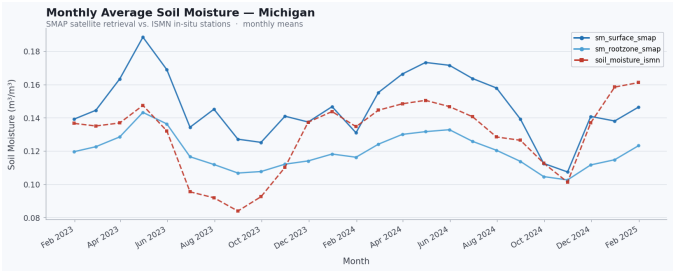


Figure 5. Monthly average soil moisture comparison between SMAP (surface and root-zone) and ISMN observations at the Michigan site.

The residual analysis in Figure 6 for the Michigan site reveals that SMAP surface retrievals exhibit a consistent positive wet bias relative to ISMN in-situ measurements, with a mean bias of $+0.0186$ m^3/m^3 and an RMSE of 0.0313 m^3/m^3. Critically, the residuals are not randomly distributed around zero, the OLS trend line (slope $= -0.463$, $R^2 = 0.185$) indicates a statistically meaningful moisture-dependent bias structure, whereby SMAP most strongly overestimates soil moisture at low in-situ values (approximately 0.08–0.10 m^3/m^3) and progressively converges toward, and ultimately crosses, zero bias at higher moisture levels near 0.17–0.18 m^3/m^3. The density-coloured scatter shows that the bulk of observations cluster between 0.12 and 0.15 m^3/m^3, where SMAP and ISMN are in closest agreement, and the ± 1 SD band of 0.0251 m^3/m^3 confirms moderate but non-negligible spread. This moisture-dependent overestimation at drier conditions is consistent with known SMAP retrieval limitations in temperate, vegetated

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

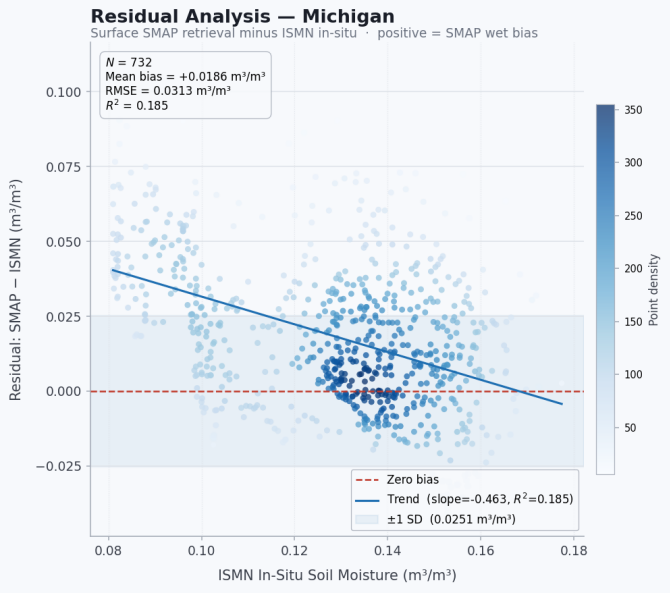


Figure 6. Residuals (SMAP surface minus ISMN) over time for Michigan.

landscapes such as Michigan, where canopy opacity and soil heterogeneity can inflate passive microwave-derived surface moisture estimates when true soil moisture is low.

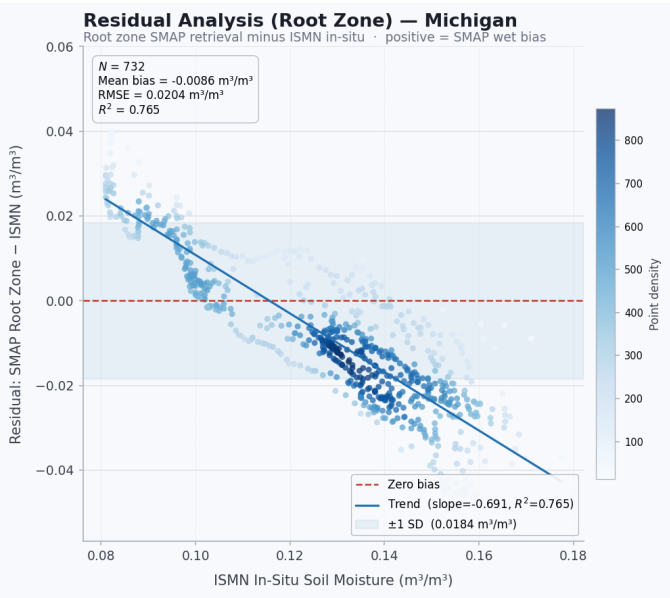


Figure 7. Residuals (SMAP root-zone minus ISMN) over time for Michigan.

The root zone residual analysis for the Michigan site in Figure 7 reveals a strongly moisture-dependent bias structure between SMAP root zone retrievals and ISMN in-situ measurements, with the OLS trend line (slope $= -0.691$, $R^2 = 0.765$) explaining over three-quarters of the residual variance, a markedly stronger relationship than observed for the surface layer. The mean bias is small and slightly negative (-0.0086 m^3/m^3), indicating that SMAP root zone estimates are near-unbiased on average, yet this aggregate statistic masks a pronounced systematic pattern: SMAP overestimates at low in-situ moisture values (roughly 0.08–0.10 m^3/m^3) and increasingly underestimates as in-situ moisture rises above approximately 0.10 m^3/m^3, with the zero-crossing occurring near 0.10 m^3/m^3. The tight clustering of high-density points along the trend line, particularly visible between 0.12 and 0.15 m^3/m^3, and the relatively narrow ± 1 SD band of 0.0184 m^3/m^3 together confirm that this is a systematic, structured bias rather than random noise. The RMSE of 0.0204 m^3/m^3 is notably lower than that of the surface layer (0.0313 m^3/m^3), suggesting that while the root zone product is more precise, its strong moisture-dependent slope points to a scaling or model assumption within the SMAP root zone algorithm that does not fully capture the dynamic range of soil moisture conditions observed across Michigan's in-situ network.

C. Indiana Site: Bedford-5-WNW

For the Indiana site ($n = 732$ daily observations), SMAP captures the general seasonal patterns observed in ISMN data (Figure 8), but with substantially weaker agreement compared to Michigan.

At the daily scale, SMAP exhibits smoother temporal behavior, reflecting spatial averaging effects. The surface product shows a negligible bias ($ME \approx -0.002$), but significantly larger errors ($MAE \approx 0.077$, $RMSE \approx 0.111$) compared to Michigan. The root-zone product shows a positive bias ($ME \approx 0.039$) with similarly high error levels ($MAE \approx 0.070$, $RMSE \approx 0.112$). Unlike Michigan, **correlation** between SMAP and ISMN is extremely weak, with $r \approx -0.01$ (surface) and $r \approx -0.09$ (root-zone). This indicates that SMAP fails to capture day-to-day variability at this site, despite reproducing general seasonal trends. **Residual analysis** (Figures 9 and 10) reveals structured discrepancies, suggesting the influence of environmental factors such as precipitation events, vegetation, and soil properties. Temporal aggregation (Figure 8) improves visual agreement, indicating that discrepancies are most pronounced at finer temporal scales. Overall, SMAP performance at the Indiana site is notably weaker, with higher errors and poor correlation, highlighting strong site dependency in satellite soil moisture accuracy.

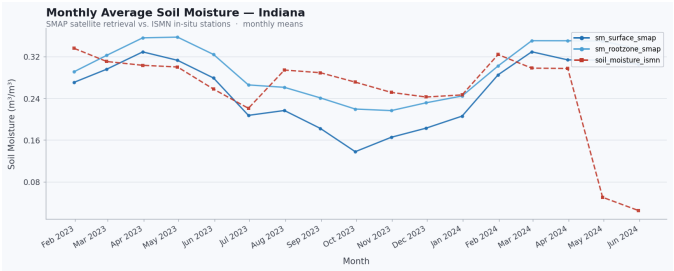


Figure 8. Monthly average soil moisture comparison between SMAP and ISMN at the Indiana site.

The surface residual analysis for Indiana in Figure 9 reveals a pattern of SMAP retrieval error that closely mirrors the root zone results for the same state, yet with important distinctions.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

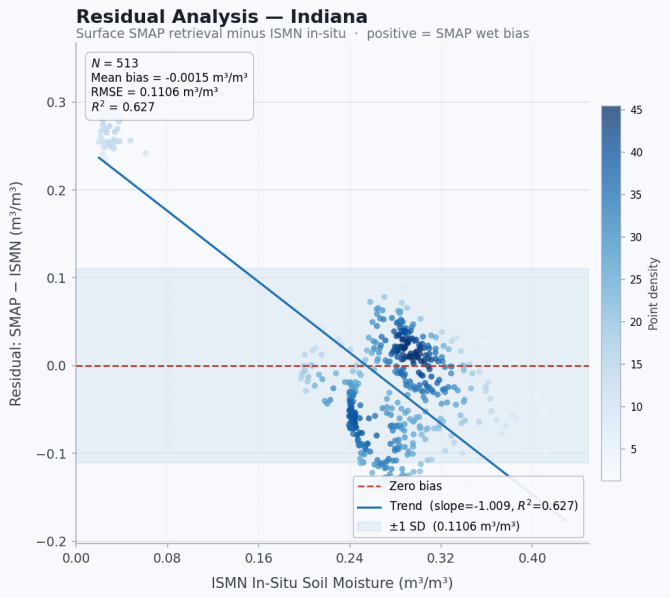


Figure 9. Residuals (SMAP surface minus ISMN) over time for Indiana.

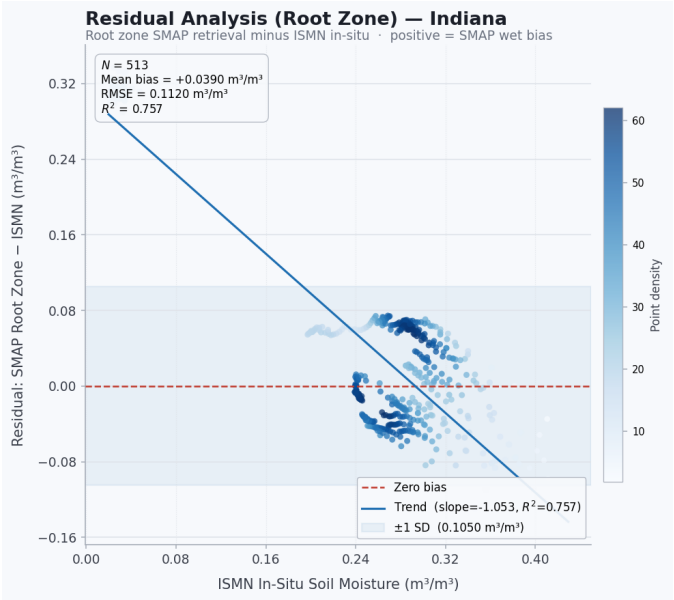


Figure 10. Residuals (SMAP root-zone minus ISMN) over time for Indiana.

The mean bias is near-zero at -0.0015 m^3/m^3, suggesting that SMAP surface retrievals are essentially unbiased on average over Indiana, a deceptively favourable statistic that obscures a severe moisture-dependent structure captured by the OLS trend line (slope $= -1.009$, $R^2 = 0.627$). A slope of approximately -1 implies that SMAP retrievals are almost entirely insensitive to the dynamic range of in-situ moisture variability, effectively returning a near-constant value while the ground truth varies substantially across the 0.24–0.38 m^3/m^3 range observed by the ISMN network. The RMSE of 0.1106 m^3/m^3 is essentially identical to the root zone RMSE (0.1120 m^3/m^3) and more than five times larger than the Michigan surface RMSE (0.0313 m^3/m^3), confirming that Indiana represents a persistently challenging retrieval environment at both depth layers. The scatter plot further reveals a structurally bimodal residual distribution, with a dense high-positive cluster near 0.24 m^3/m^3 and a secondary negative cluster near 0.30 m^3/m^3, which together drive the wide ± 1 SD band of 0.1106 m^3/m^3 and point toward systematic station-level or land-surface heterogeneity, likely tied to Indiana's extensive row-crop agriculture, artificial tile drainage, and spatially variable soil organic matter, that the SMAP surface retrieval algorithm does not adequately resolve.

The root zone residual analysis for the Indiana site in Figure 10 exposes a severe and systematic disagreement between SMAP root zone retrievals and ISMN in-situ measurements that is qualitatively distinct from the Michigan results. The OLS trend line (slope $= -1.053$, $R^2 = 0.757$) carries a slope exceeding unity in magnitude, indicating that the residual grows faster than the in-situ moisture signal itself, a hallmark of a fundamental dynamic range mismatch rather than a simple offset. The mean bias of $+0.0390$ m^3/m^3 is considerably larger than Michigan's (-0.0086 m^3/m^3), and the RMSE of 0.1120 m^3/m^3 is more than five times higher (0.0204 m^3/m^3 for Michigan), reflecting the much wider scatter visible in the figure. Notably, the in-situ observations are confined to a narrow and relatively high moisture range of approximately 0.24–0.38 m^3/m^3, yet the residuals span nearly ± 0.10 m^3/m^3 within that window, and the data appear to cluster into two distinct lobes straddling the zero-bias line, suggestive of station-level heterogeneity, multi-depth sampling artefacts, or land cover discontinuities across the Indiana network. The wide ± 1 SD band of 0.1050 m^3/m^3 further underscores the large unexplained variability. Taken together, these results suggest that the SMAP root zone algorithm performs substantially worse over Indiana than Michigan, possibly reflecting differences in soil texture, agricultural land use, or tile drainage infrastructure that are not adequately parameterized in the SMAP retrieval model.

VII. Conclusions and Recommendations

Accurate soil moisture estimation is essential for sustainable water resource management, yet the reliability of the benchmark datasets used to evaluate predictive models has received relatively limited attention. This study addressed this gap through a combined review and empirical analysis, providing insights with direct implications for the development, validation, and deployment of soil moisture prediction systems.

The review showed that recent advances in machine learning, deep learning, and data fusion have substantially improved soil moisture prediction capabilities. However, progress has been constrained in part by limitations in the benchmark datasets themselves. Although ISMN remains an indispensable source of global ground observations, it suffers from several well-documented challenges, including uneven spatial distribution, sensor heterogeneity, inconsistent metadata, and temporal gaps. These issues represent structural limitations that may

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

systematically bias model evaluation, particularly when a single dataset is used as the sole benchmark.

The empirical case study further reinforced this concern. At the Michigan site, SMAP root-zone estimates exhibited reasonable agreement with ISMN observations ($r = 0.63$, RMSE $= 0.02\ \mathrm{m}^3/\mathrm{m}^3$), and residual analysis revealed moisture-dependent biases consistent with known retrieval limitations in vegetated temperate environments. In contrast, performance at the Indiana site was considerably poorer, with both surface and root-zone estimates exhibiting RMSE values exceeding $0.11\ \mathrm{m}^3/\mathrm{m}^3$ and near-zero correlation with in-situ measurements. Moreover, the ordinary least squares residual slope approached -1, indicating that SMAP retrievals were largely insensitive to variations in ground moisture conditions and instead produced nearly constant estimates while observed moisture values varied substantially. This behavior reflects a fundamental dynamic-range mismatch rather than a simple calibration offset. Environmental characteristics such as extensive row-crop agriculture, subsurface tile drainage, and heterogeneous soil organic matter likely contribute to these discrepancies and are not adequately represented in current retrieval algorithms.

Taken together, these findings suggest three practical recommendations for researchers and practitioners:

1) **Adopt multi-source validation as standard practice.** Reliance on ISMN alone, or any single benchmark dataset, may lead to misleading conclusions regarding model performance. Instead, multiple complementary datasets should be incorporated. For example, SCAN can provide agricultural measurements, ASCAT offers temporally continuous observations, and regional networks such as OzNet and REMEDHUS provide valuable site-specific information. Agreement across independent sources provides stronger evidence of model robustness.
2) **Apply site-aware preprocessing before using ISMN for validation.** Sites characterized by intensive agriculture, artificial drainage systems, or high sub-pixel heterogeneity should receive additional scrutiny. Residual analysis, which examines bias as a function of soil moisture magnitude rather than relying solely on aggregate metrics, can reveal systematic errors that may be overlooked by summary statistics.
3) **Prioritize data fusion approaches in heterogeneous agricultural landscapes.** The Indiana case study suggests that passive microwave retrievals alone may be insufficient in row-crop environments. Integrating SMAP observations with higher-resolution ancillary information, including soil texture maps, crop calendars, and drainage infrastructure data, through ensemble Kalman filtering or deep learning-based fusion frameworks represents a promising direction for improving field-scale soil moisture estimation.

Future work at the Precision Agriculture Research Lab at GVSU will extend this analysis to additional ISMN sites encompassing diverse land cover types and climatic regimes. Planned efforts include the development of site-classification protocols to guide benchmark selection and the construction of integrated predictive frameworks that combine satellite, in-situ, and model-based observations for precision agriculture applications. Ultimately, the reliability of soil moisture prediction systems depends not only on the sophistication of the underlying models, but also on the rigor with which the datasets used for training and evaluation are characterized and understood.

References

[1] K. Hasan and A. Muiruri, "Predictive modeling of soil moisture: A review of benchmark datasets, their strengths, and limitations", *The Second International Conference on Sustainable and Regenerative Farming (CSRF)*, vol. 978-1-68558-327-9, pp. 9–12, 2025.

[2] L. F. Konikow, "Groundwater depletion in the united states (1900–2008)", *US Geological Survey Scientific Investigations Report*, vol. 2013, no. 5079, p. 63, 2005.

[3] S. Foster and J. Chilton, "Groundwater: The processes and global significance of aquifer degradation", *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, vol. 358, no. 1440, pp. 1957–1972, 2003.

[4] U.S. Department of Agriculture, National Agricultural Statistics Service, "2018 irrigation and water management survey", USDA-NASS, Tech. Rep., 2019.

[5] S. Jasechko, D. Perrone, K. M. Befus, *et al.*, "Rapid groundwater decline and some cases of recovery in aquifers globally", *Nature*, vol. 625, pp. 715–721, 2024. DOI: 10.1038/s41586-023-06879-8.

[6] D. Entekhabi *et al.*, "The soil moisture active passive (SMAP) mission", *Proceedings of the IEEE*, vol. 98, no. 5, pp. 704–716, 2010. DOI: 10.1109/JPROC.2010.2043918.

[7] W. Dorigo *et al.*, "Esa cci soil moisture for improved earth system understanding: State-of-the art and future directions", *Remote Sensing of Environment*, vol. 203, pp. 185–215, 2017.

[8] I. Eniang, A. Edet, and E. Anwan, "Importance of soil moisture content in agriculture: A review", *International Journal of Agriculture Innovations and Research*, vol. 7, no. 2, pp. 2319–1473, 2018.

[9] L. Alletto, Y. Coquet, P. Benoit, D. Heddadj, and E. Barriuso, "Soil moisture monitoring for optimizing irrigation management in a maize cropping system", *Agricultural Water Management*, vol. 148, pp. 1–13, 2015.

[10] W. A. Dorigo *et al.*, "The international Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements", *Hydrology and Earth System Sciences*, vol. 15, no. 5, pp. 1675–1698, 2011. DOI: 10.5194/hess-15-1675-2011.

[11] M. Rodell *et al.*, "The global land data assimilation system", *Bulletin of the American Meteorological Society*, vol. 85, no. 3, pp. 381–394, 2004.

[12] K. W. Oleson *et al.*, "Technical description of version 4.0 of the community land model (clm)", NCAR, Tech. Rep., 2010.

[13] T. Enemark, I. Sandholt, R. Fensholt, and K. H. Jensen, "Evaluating the skill of land surface models to represent the soil moisture variability over africa", *Hydrology and Earth System Sciences*, vol. 24, no. 7, pp. 3717–3738, 2020.

[14] Y. Zhang, Z. Li, H. Shen, and E. Chen, "Machine learning for downscaling remotely sensed soil moisture using multisource data", *Remote Sensing*, vol. 10, no. 3, p. 431, 2018.

[15] D. Jeong, H. Lee, and S. Lee, "Machine learning-based soil moisture prediction using remote sensing and weather data", *Environmental Modelling & Software*, vol. 148, p. 105 248, 2022.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

[16] S. Khaki *et al.*, "Cnn-rnn deep learning framework for spatiotemporal soil moisture estimation", *Water Resources Research*, vol. 56, no. 2, 2020.

[17] M. Reichstein *et al.*, "Deep learning and process understanding for data-driven earth system science", *Nature*, vol. 566, no. 7743, pp. 195–204, 2019.

[18] J. Yin, Q. Zhuang, H. Tian, and M. Pan, "A physics-informed machine learning framework for hydrological modeling: Application to the mississippi river basin", *Water Resources Research*, vol. 57, no. 8, 2021.

[19] S. Chan, P. O'Neill, E. Njoku, T. Jackson, and R. Bindlish, "Soil moisture active passive (smap) enhanced level 3 radiometer global daily 9 km ease-grid soil moisture, version 1", *NASA National Snow and Ice Data Center Distributed Active Archive Center*, 2018.

[20] M. Saberi, H. Moradkhani, W. Bardsley, and X. He, "Multiscale soil moisture modeling for agricultural drought monitoring using remote sensing and land surface modeling data", *Agricultural and Forest Meteorology*, vol. 308, p. 108 548, 2021.

[21] F. Greifeneder, L. Brocca, T. Pellarin, and W. Wagner, "Machine learning in soil moisture retrievals: Progress, challenges, and future directions", *Reviews of Geophysics*, vol. 59, no. 2, 2021.

[22] A. Gruber, W. T. Crow, W. A. Dorigo, and W. Wagner, "Characterizing coarse-scale representativeness of in situ soil moisture measurements from the international soil moisture network", *Vadose Zone Journal*, vol. 12, no. 2, 2013.

[23] N. Nicolai-Shaw, L. Gudmundsson, M. Hirschi, and S. I. Seneviratne, "Improving soil moisture modeling in heterogeneous landscapes", *Hydrology and Earth System Sciences*, vol. 23, pp. 1217–1235, 2019.

[24] W. A. Dorigo *et al.*, "The international soil moisture network: Serving earth system science for over a decade", *Hydrology and Earth System Sciences*, vol. 25, no. 2, pp. 5749–5804, 2021. DOI: 10.5194/hess-25-5749-2021.

[25] S. Kim, T. J. Jackson, R. Bindlish, M. H. Cosh, and V. Lakshmi, "Validation of smos soil moisture retrievals over the continental us using ground-based observations", *Remote Sensing of Environment*, vol. 121, pp. 383–398, 2012.

[26] Y. Ma *et al.*, "Validation of smap soil moisture products with in situ observations from the international soil moisture network", *Remote Sensing*, vol. 8, no. 6, p. 524, 2016.

[27] H. Wu, X. Li, and J. Peng, "Evaluating soil moisture retrievals from amsr2 using ground-based observations over semiarid regions in china", *Remote Sensing*, vol. 9, no. 5, p. 499, 2017.

[28] Y. Liu, R. M. Parinussa, W. A. Dorigo, R. A. M. de Jeu, and W. Wagner, "A comparative study of four algorithms for downscaling remotely sensed soil moisture over the contiguous us", *Remote Sensing of Environment*, vol. 179, pp. 1–19, 2016.

[29] S. K. Do, T.-N.-D. Tran, M.-H. Le, J. Bolten, and V. Lakshmi, "A novel validation of satellite soil moisture using sm2rain-derived rainfall estimates", *Frontiers in Remote Sensing*, vol. 5, p. 1 474 088, 2024.

[30] C. S. C. Draper *et al.*, "Comparison of near-surface soil moisture from smos and ascat with in situ measurements from the u.s. scan network", *Remote Sensing of Environment*, vol. 115, no. 12, pp. 3070–3088, 2011. DOI: 10.1016/j.rse.2011.03.026.

[31] L. Xu, X. Zhang, J. Wang, and Y. Zhu, "Downscaling SMAP soil moisture using wide and deep learning with spatial feature engineering", *Remote Sensing*, vol. 14, no. 4, p. 987, 2022. DOI: 10.3390/rs14040987.

[32] M. Celik, Y. Yilmaz, and O. Kose, "Multisource deep learning model for soil moisture prediction in diverse climates", *Environmental Modelling & Software*, vol. 155, p. 105 430, 2022.

[33] S. Lee, M. Choi, and E. Kim, "A deep learning framework for smap soil moisture retrieval without radiative transfer models", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–13, 2024.

[34] P. Custódio and A. Prati, "Adaptive ensemble models for soil moisture estimation in iot-supported irrigation", *Agricultural Water Management*, vol. 295, p. 108 579, 2024.

[35] G. L. Schaefer, M. H. Cosh, and T. J. Jackson, "Usda soil climate analysis network (scan)", *Journal of Atmospheric and Oceanic Technology*, vol. 24, no. 12, pp. 2073–2077, 2007.

[36] D. Baldocchi *et al.*, "FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities", *Bulletin of the American Meteorological Society*, vol. 82, no. 11, pp. 2415–2434, 2001. DOI: 10 . 1175 / 1520 - 0477(2001) 082<2415:FANTTS>2.3.CO;2.

[37] Y. H. Kerr *et al.*, "The SMOS mission: New tool for monitoring key elements of the global water cycle", *Proceedings of the IEEE*, vol. 98, no. 5, pp. 666–687, 2010. DOI: 10.1109/JPROC.2010.2043032.

[38] W. Wagner *et al.*, "ASCAT soil moisture product: Operational validation, hydrometeorological applications, and beyond", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 3, pp. 1452–1466, 2013. DOI: 10.1109/JSTARS.2013.2262336.

[39] A. C. Teixeira, M. Bakon, D. Lopes, A. Cunha, and J. J. Sousa, "A systematic review on soil moisture estimation using remote sensing data for agricultural applications", *Science of Remote Sensing*, p. 100 328, 2025.

[40] C. Zheng, L. Jia, and T. Zhao, "A 21-year dataset (2000–2020) of gap-free global daily surface soil moisture at 1-km grid resolution", *Scientific Data*, vol. 10, no. 1, p. 139, 2023.

[41] Y. Zhu, L. Zhang, F. Li, J. Xu, and C. He, "Comparison of data fusion methods in fusing satellite products and model simulations for estimating soil moisture on semi-arid grasslands", *Remote Sensing*, vol. 15, no. 15, p. 3789, 2023.

[42] P. Harani, S. Gautam, S. K. Joshi, L. G. Asirvatham, and B. L. Rakshith, "Advancements in soil moisture estimation through integration of remote sensing and artificial intelligence techniques", *Science of The Total Environment*, vol. 1001, p. 180 503, 2025.

[43] A. Singh and K. Gaurav, "Deep learning and data fusion to estimate surface soil moisture from multi-sensor satellite images", *Scientific Reports*, vol. 13, no. 1, p. 2251, 2023.

[44] M. S. Tahmouresi, M. H. Niksokhan, and A. H. Ehsani, "Enhancing spatial resolution of satellite soil moisture data through stacking ensemble learning techniques", *Scientific Reports*, vol. 14, no. 1, p. 25 454, 2024.

[45] S. Zakzouk and L. A. Said, "Federated learning for soil moisture prediction: Benchmarking lightweight cnns and robustness in distributed agricultural iot networks", *Machine Learning and Knowledge Extraction*, vol. 7, no. 4, p. 132, 2025.

[46] S. O. Ihuoma, C. A. Madramootoo, and M. Kalacska, "Integration of satellite imagery and in situ soil moisture data for estimating irrigation water requirements", *International journal of applied earth observation and geoinformation*, vol. 102, p. 102 396, 2021.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Improving Object Detection with Semi-Supervised Learning from Temporal Information

Viljar Holm Elvevoll[1], Kim Tallaksen Halvorsen[2], and Ketil Malde[2]
[1] Department of Informatics, University of Bergen, Norway
[2] Institute of Marine Research, Bergen, Norway
e-mail: viljarhe00@hotmail.no, {kim.halvorsen | ketil.malde}@hi.no

***Abstract*—Effective fisheries management and conservation depends on accurate and sustainable monitoring of marine biodiversity. Traditionally, fish population assessments have relied on manual annotation and invasive techniques that are labor-intensive and often detrimental to marine ecosystems they aim to preserve. Use of cameras in combination with automated object detection can make ecosystem observation both less invasive and less costly, but sufficient accuracy in species detection and classification can be a challenge, especially for low abundance species. Here we present a semi-supervised learning approach that leverages extensive unlabeled underwater video data to significantly enhance object detection performance for fish species. By integrating the YOLOv8 object detector with multi-object tracking algorithms ByteTrack and DeepSORT, a novel methodology is proposed to generate high-quality pseudolabels from temporal sequences. We show that iterative training that incorporates such pseudolabels consistently improved model precision and recall, with the best-performing approach (ByteTrack with an extrapolated heuristic) demonstrating average precision of 90%, recall of 70%, mAP50 of 74%, and mAP50-95 of 59%. Notably, scores improved substantially over the baseline supervised model on all metrics. These results underscore the potential of temporally informed pseudolabeling in enhancing fish detection accuracy and robustness, reducing reliance on manual annotations, and ultimately supporting more effective monitoring of fish communities.**

***Keywords*-Image classification; machine learning; species recognition; semi-supervised learning**

I. Introduction

Traditionally, methods for monitoring marine ecosystems include trawling, netting, and manual visual surveys by divers, which are labor intensive, costly, and often disruptive to habitats or producing bycatch [1]. Less invasive methods using underwater cameras, such as Baited Remote Underwater Video (BRUV) and Remote Underwater Video (RUV), are often an attractive alternative [2][3]. Most surveys employing these tools rely on manually scrutiny of video, which is costly in terms of time and labor. There is therefore a considerable incentive for leveraging object detection and classification models for semi or fully automated workflows [4][5]. However, training such models requires large amounts of high-quality, labeled data, which is costly to produce.

To address this limitation, we here investigate semi-supervised learning [6], an automated method to iteratively generate training data sets using predictions from preliminary models that are considered sufficiently reliable. To differentiate between labels determined by an expert annotator and those output from a model, the latter are referred to as *pseudolabels*. In contrast to earlier work, we selected the pseudolabeled data to use based on temporal information (*i.e.*, tracking) rather than more commonly used confidence scores. This is particularly advantageous in this setting, since an abundance of temporally contiguous video or image data can be produced, but expert annotation is time consuming and requires skilled curators.

The rest of the paper is structured as follows. In Section II, we describe the data set and the method for generating pseudolabels, as well as the training regime. In Section III, we present the results, and select the best performing method to investigate further. In Section IV, we discuss the results and their implications, and propose an explanation for the observations, before we conclude in Section V.

II. Methods

For this study, we used a data set consisting of 1248 images from a combination of sources (RUVs, photos by divers) under different conditions and with variable resolutions [7], (see Figure 1 for an example). The annotation by experts from the Institute of Marine Research include 10 categories (Figure 2): corkwing wrasse (*Symphodus melops*; male and female), two-spotted goby (*Pomatoschistus flavescens*), goldsinny (*Ctenolabrus rupestris*), rock cook (*Centrolabrus exoletus*), cuckoo wrasse (*Labrus mixtus*) male and female), pollack (*Pollachius pollachius*), ballan wrasse (*Labrus bergylta*), and *unknown* fish that could not be labeled to the species level due to low visibility or by being too distant from the camera. As shown in Figure 3, the number of objects (fish) per image varied, with a single fish being the most common, up to a maximum of 26 fish in one image.

In addition, six unannotated videos from similar habitats [7] were used as sources of pseudolabeled frames for the semi-supervised training. Each video has a resolution of 3840 x 2160 pixels and a frame rate of 30 frames per second. The average duration is approximately 8 minutes and 28 seconds with a minimum length of 6 minutes and 35 seconds and a maximum length of 8 minutes and 51 seconds. The videos were recorded by submerging a GoPro Hero9 Black camera near an underwater nest of a male corkwing wrasse to capture the surrounding environment. As some species were not present or very scarce in the unlabeled data, the semi-supervised method is only trained on five of the classes: male and female corkwing, two-spotted goby, goldsinny, and ballan. The videos may also record fish from other species which are not present in the training data labels. These factors are

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

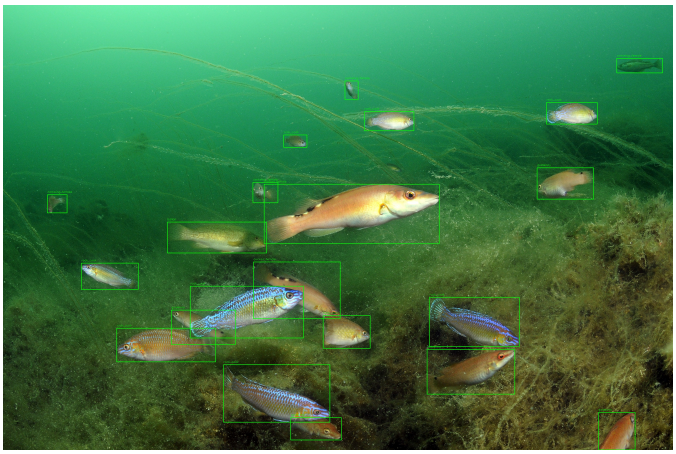


Figure 1. An example from the data set showing several annotated fish of various species (Photo by Erling Svendsen, used with permission).

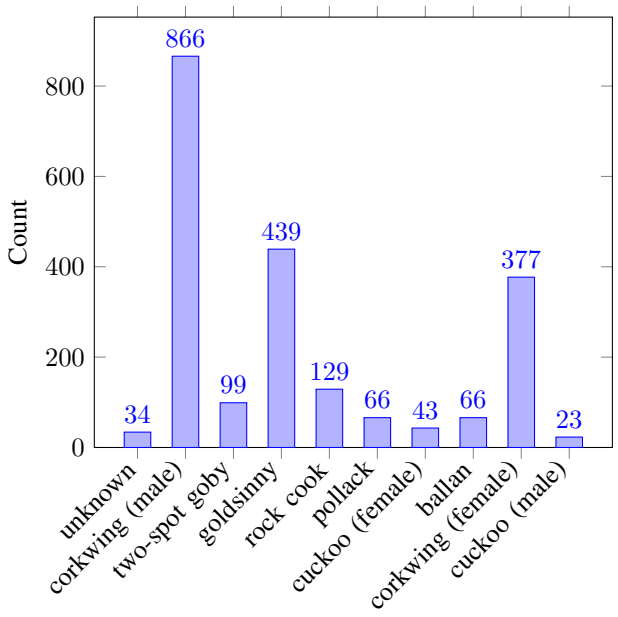


Figure 2. The distribution of the classes of annotated objects in the data set.

difficult to control when using real world data, but are likely to negatively affect the training process and consequently, the performance of the final classifier.

The object detection model used was YOLOv8 (Ultralytics, 2023), a state-of-the-art object detector from the family of single-pass models [8]. We used two different multi-object tracking algorithms, ByteTrack [9] and DeepSORT [10]. DeepSORT (Deep Simple Online and Realtime Tracking) is based on SORT [11]. Like SORT, DeepSORT uses a Kalman filter that predicts the next location of an object based on previous observations, but extends it with a deep learning based embedding method that matches the visual appearance of detected objects.

ByteTrack is a conceptually simpler method that links all detections (including low confidence ones) based on the overlap (intersection over union, or IoU) between detection bounding boxes. Although it is a less complex approach, it often performs well in independent benchmarks [12]. Both DeepSORT and ByteTrack have been adapted and developed to track object detector output for underwater images. For instance, the detection comparison module in DeepSORT has been modified to give better performance on fish [13], and the ByteTracks use of intersection-over-union has been adapted to fish morphology [14].

Each of the object trackers is combined with the two different strategies for pseudolabel generation illustrated in Figure 4 to construct four versions of an iterative pipeline, illustrated in Figure 5:

1. Base Model Training: An initial YOLOv8 model (YOLOv8x, the largest variant) was trained using a smaller, manually annotated dataset [7]. This provides the baseline for performance evaluation.

2. Pseudolabel Generation using Temporal Information: The trained model is then used to process unlabeled underwater video recordings and extract pseudolabeled data as illustrated in Figures 6 and 7. By integrating Multi-Object Tracking (MOT) algorithms (ByteTrack and DeepSORT), frames with objects that are missed or given low score by the detector can still be identified with high confidence. Two heuristics were investigated for selecting frames to generate pseudolabels:

A. Interpolated Intermediate Labels: This heuristic infers an object's presence in intermediate frames if it is detected by the model in preceding and subsequent frames, and the MOT algorithm assigns the same track ID. This method yields fewer but potentially highly accurate pseudolabels.

B. Extrapolated Labels: This more inclusive approach requires an object to be natively detected at least three times consecutively. All subsequent detections of that object by the MOT algorithm are included, forming an unbroken chain, even if the native model fails to detect it in every frame. This significantly increases the volume of pseudolabeled data.

To operationalize the two heuristics, we run two parallel inference streams on the same video frames. The first is a detector-only stream where YOLOv8 predicts each frame independently, and the second is a tracking stream where the detector outputs are linked over time by an MOT algorithm. Pseudo-label candidacy is then decided by comparing these two streams: the heuristics only accept samples when the tracker provides temporal evidence for an object in frames where the detector is inconsistent or missing. When multiple fish are present simultaneously, decisions are made per track ID, and ambiguous cases (e.g., potential ID switches or multiple competing associations) are skipped to avoid propagating uncertain assignments into the training data.

For the *interpolated intermediate labels* heuristic, we look for short temporal gaps where a track is present, but the detector misses it. Specifically, if the detector natively detects an object in a preceding frame t_1 and a subsequent frame t_3, and the tracker maintains the same track ID through an intermediate frame t_2, we infer that the object is also present in t_2 and add the tracker-provided box as a pseudo-label for that frame. During implementation we allow the "gap" to span

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

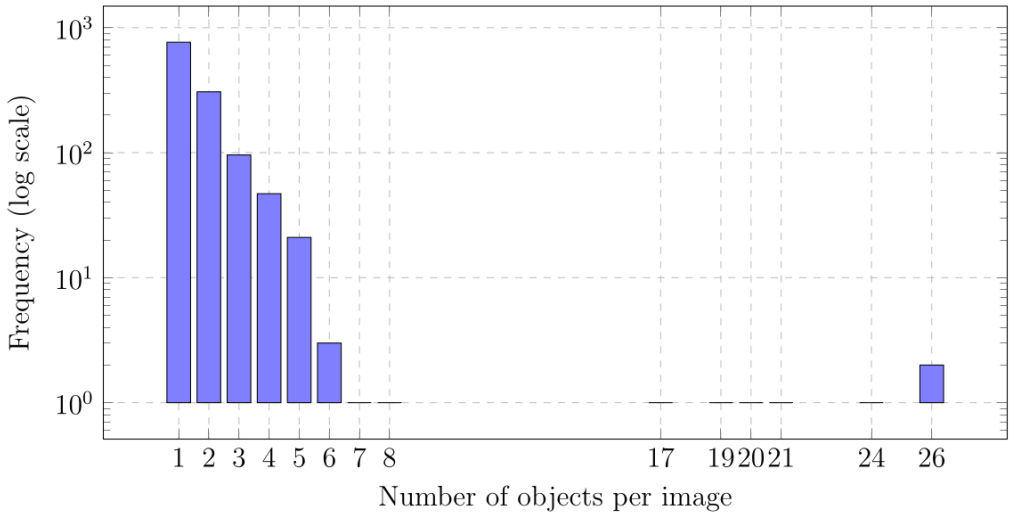


Figure 3. Distribution of the number of annotated fish objects per image in the data set.

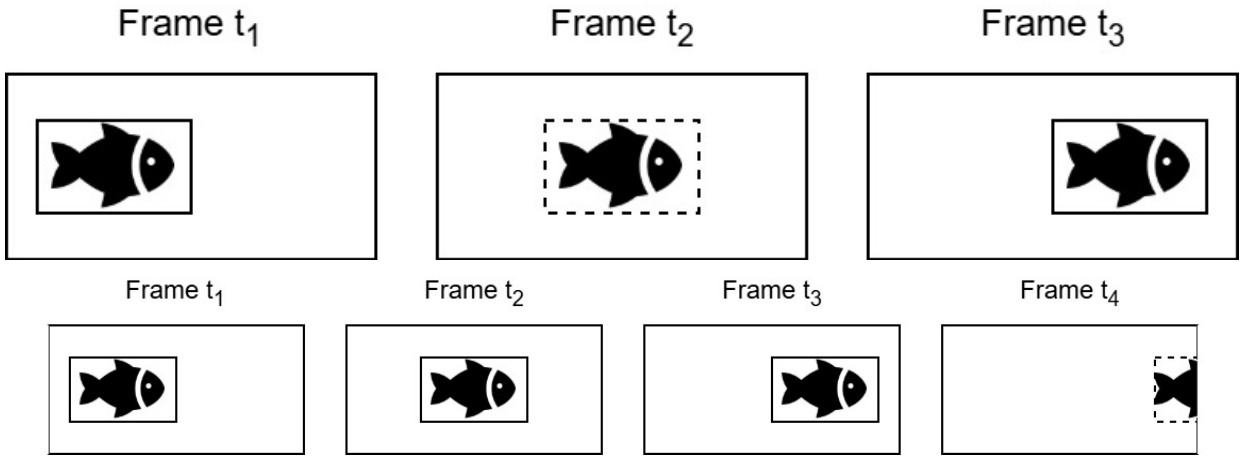


Figure 4. Two methods for pseudolabel generation. For interpolated labels (top row), a detection missed by the classifier is inferred from detections in the preceding and subsequent frames, targeting detector weaknesses on visually challenging frames. For extrapolated labels (bottom row), detections are inferred from a series of prior detections, which can add difficult partial views to the training data and improve robustness to occlusions.

up to a few consecutive frames to capture brief misses due to pose changes, turbidity, or partial occlusion, while still keeping the pseudo-labels tightly anchored to native detections.

For the *extrapolated labels* heuristic, we instead start from a reliable "seed" where the detector natively detects the same object at least three frames in a row, which also enables estimating a velocity for the track. From that seed onward, we include subsequent frames as pseudo-labels as long as (i) the tracker maintains an unbroken chain for the same ID and (ii) the detector does *not* natively detect the object in those frames. This produces substantially more pseudo-labeled training data, but it also increases the risk that tracker noise or a single incorrect native detection can propagate into many frames. The velocity requirement acts as a simple sanity check that helps exclude degenerate cases (e.g., static false positives) that would otherwise be easy to "track" over time.

This example illustrates the intuition behind the temporal selection criteria: the middle frame is visually challenging (unusual pose and angle), and the detector fails to produce a native detection. However, because the tracker maintains a consistent identity through the sequence, the middle frame becomes a high-value pseudo-label candidate that specifically targets a known weakness of the detector rather than reinforcing already-confident detections.

This second example highlights a characteristic outcome of the extrapolated heuristic: later pseudo-labeled frames may contain only a partial view of the fish (e.g., the tail), which would be difficult to label manually and is often missed by the detector. Including such cases increases data diversity and can improve robustness to partial occlusions, but it also emphasizes the importance of preventing tracker drift and ID switches from entering the training set.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

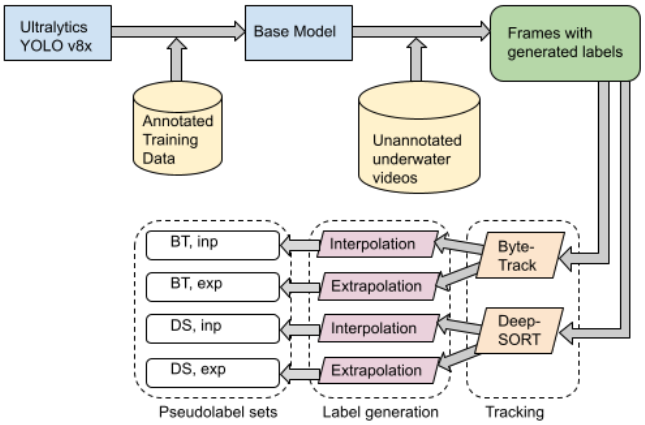


Figure 5. The process for generating the base model and the four pseudolabeled data sets (step 1 and step 2), using automated labeling of video data, tracking, and label interpolation and extrapolation.

The two heuristics also differ markedly in how much new training data they produce. In the first iteration of our pipeline, interpolated intermediate labels yielded on the order of a few thousand pseudo-labeled frames per tracker, while extrapolated labels yielded an order of magnitude more. This difference is important operationally: extrapolation increases diversity and coverage (including difficult partial views), but it also increases storage and compute requirements and can amplify any systematic tracker error. In other words, the heuristics represent a direct trade-off between expected label integrity and volume of additional training data.

3. Iterative Retraining: The pseudolabels generated are combined with the original labeled dataset, and the model is retrained. It is crucial to retain the original data to prevent catastrophic forgetting of classes not present in the video recordings. The learning rate of the AdamW optimizer is reset at the start of each new training phase to facilitate rapid adjustment and prevent trapping in suboptimal local minima. This iterative process (pseudolabel generation followed by retraining) is repeated for multiple cycles to progressively improve the model.

A practical detail in this iterative setting is *when* to stop each training phase. In our experiments, we used a fixed limit of 50 epochs per phase, but this number should be viewed as a sensitive heuristic rather than a universal default. Empirically, we observed that the base model's improvements often started to slow down around the same region, and saving a model state earlier preserved "plasticity" for the subsequent phase where new pseudo-labels are introduced. Standard early stopping is not directly applicable here: the goal is not to stop *at* convergence, but to stop *before* the model settles into a local minimum, while gradients are still sufficiently dynamic to adapt to the distribution shift introduced by new pseudo-labels. This is also why pseudo-labels are regenerated using the *current* model state at each iteration, and why we reset the optimizer learning rate between phases. Without a reset, we observed slower and more sporadic adaptation during relearning, sometimes plateauing prematurely. With a reset, it is common to see a brief dip in performance immediately after introducing new pseudo-labels, followed by recovery and improvement as the model re-optimizes on the updated training distribution.

TABLE I. PERFORMANCE METRICS AFTER 100 EPOCHS OF TRAINING.

Model	Prec	Recall	mAP50	mAP50-95
DS, inp	0.74988	0.56890	0.61023	0.47017
DS, exp	0.66251	0.56209	0.57969	0.45028
BT, inp	0.70776	0.53875	0.58142	0.44878
BT, exp	0.88792	0.62665	0.69120	0.52972

Note that for all cases, the validation and test set were the same, taken from the initial, annotated data set. In other words, pseudolabeled data were not used for evaluating model performance.

III. RESULTS

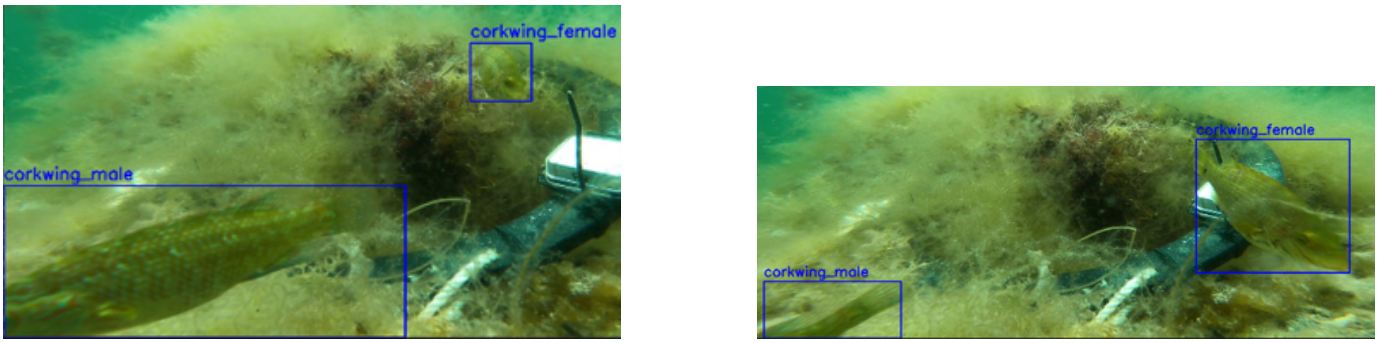
The study rigorously evaluated four configurations: ByteTrack with interpolated intermediate labels, DeepSORT with interpolated intermediate labels, ByteTrack with extrapolated labels, and DeepSORT with extrapolated labels. Each configuration was run for 100 epochs (50 supervised, followed by 50 semi-supervised with pseudolabels) which yielded the results in Table I. We see that all models perform adequately, and while ByteTrack with extrapolated labels consistently outperformed all other models, the ranking of the other approaches varied across metrics. Generally, DeepSORT with interpolation was the second-best performer, followed by DeepSORT with extrapolation, and ByteTrack with interpolation exhibiting the lowest performance for all metrics.

In order to explore the limits of semi-supervised training, the baseline and ByteTrack with extrapolation models were trained for 250 epochs. ByteTrack with extrapolation were chosed for comparison, since this comparison showed the best performance. In Figure 8, we can see how the different components of the loss rapidly decrease both for training (top row) and validation (bottom row) data, while the four different performance measures increase correspondingly. We also observe five distinct jumps in the graphs, these are caused by introduction of new data and resetting of the learning rate for each iteration, which cause an initial worsening of scores before the model gradually converges again. The loss curves provide additional insight into why the semi-supervised pipeline improves validation performance even though training becomes more "disrupted". The supervised baseline shows smooth convergence on the training data, as expected from repeatedly optimizing on a fixed dataset. In contrast, the semi-supervised approach experiences transient spikes in training loss at each iteration boundary due to (i) replacing pseudo-labels with new ones derived from the current model state and (ii) resetting the learning rate. Notably, the magnitude of these spikes tends to diminish over iterations, which suggests that successive pseudo-label sets become more consistent with one another as the model and tracker gradually align.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

(a) Frame t_1 (model detection) (b) Frame t_2 (tracker only) (c) Interpolated label

Figure 6. Interpolated pseudolabel generation in practice. Note that due to ByteTrack's use of very low confidence predictions for interpolation, the inferred bounding box does not correspond exactly to the average of the two detections.



(a) Frame t_1 (model detection) (b) Extrapolated label

Figure 7. Extrapolated pseudolabel generation in practice. The visible part of the corkwing male exiting the frame is to small to be predicted, but is inferred from its presence in the previous frame.

Interestingly, the validation loss tells a different story: despite the validation set remaining unchanged and never receiving pseudo-labels, the semi-supervised validation loss trends lower over the full training schedule, while the supervised baseline tends to plateau earlier. This pattern is consistent with the semi-supervised approach improving generalization by adding difficult-but-informative training cases (frames the detector would otherwise miss), rather than merely strengthening performance where it is already confident.

Compared to the baseline model without using pseudolabels, we see from Figure 9 that semi-supervised training benefits from continued training for all metrics, and that this is reflected in validation loss shown in Figure 10. The performance statistics on the test set after 250 epochs are summarized in Table II. We see that using semi-supervised training with ByteTrack and the extrapolated pseudolabeling scheme results in substantial improvements for all metrics.

Per class improvements are shown in Figure 11. As expected, classes present in the semi-supervised training data (shown in solid colors) see substantial improvements on all metrics. Classes not present (shown with faded colors) see slight degradation in precision, and mAP, but surprisingly recall improves also for these classes.

Looking more closely at the evaluation metrics, the largest gains are typically observed in recall rather than precision. This is not unexpected: precision is often already high when the model is confident about species identity, whereas recall reflects how often the model can successfully detect harder instances. Since our pseudo-label selection explicitly targets frames where the detector fails but temporal continuity indicates the object is present, the additional training signal is biased toward "hard detections". This helps explain why recall improves substantially for pseudo-labeled classes, while precision changes are smaller and sometimes noisier.

We also observe that the semi-supervised curves can appear more sporadic than the supervised baseline on certain metrics. A plausible explanation is that the learning-rate reset and the introduction of new pseudo-label distributions can temporarily perturb the decision boundary, and metrics like mAP can be more sensitive to such perturbations than the loss. Despite this instability, the overall metric trajectories consistently favor the semi-supervised approach in our extensive training setting.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

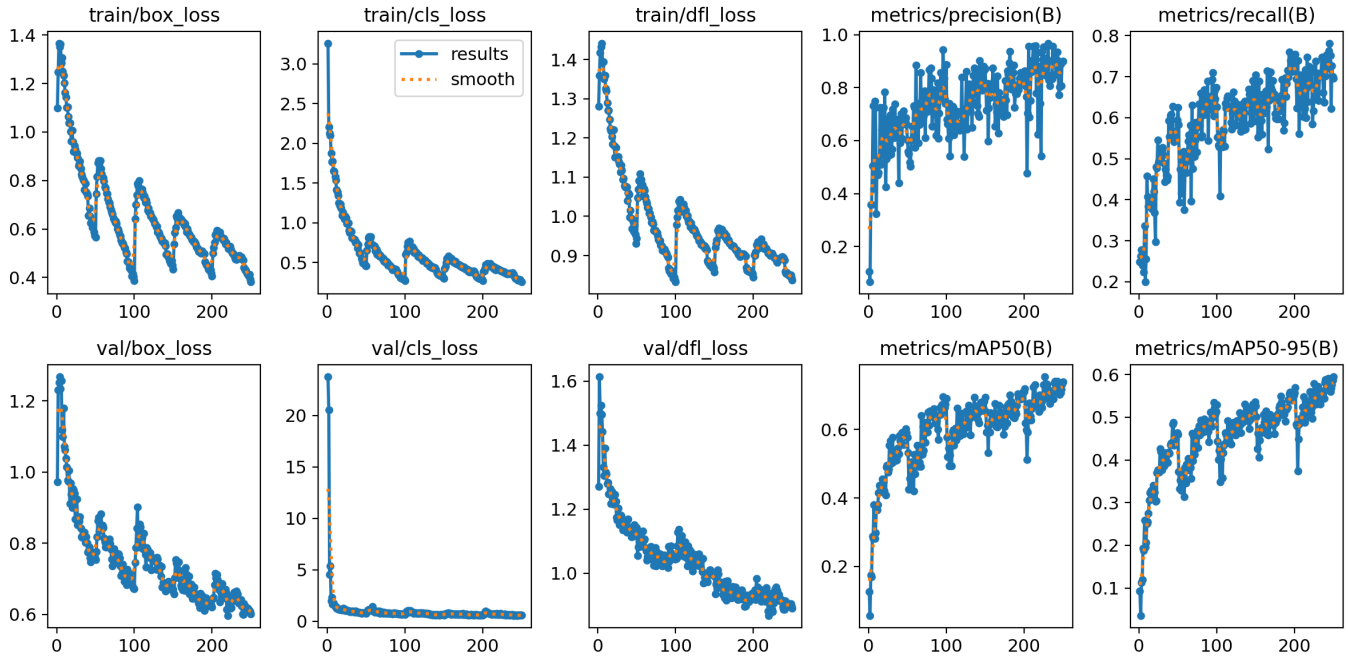


Figure 8. Training from ByteTrack tracks with extrapolated labels for 250 epochs (one iterations of supervised followed by four iterations of semi-supervised training).

TABLE II. PERFORMANCE OF BASELINE AND BYTETRACK WITH EXTRAPOLATION MODELS AFTER EXTENSIVE (250 EPOCHS) TRAINING.

Model	Prec	Recall	mAP50	mAP50-95
Base	0.75652	0.52468	0.57954	0.48811
BT, exp	0.90011	0.69644	0.73863	0.59468

IV. DISCUSSION

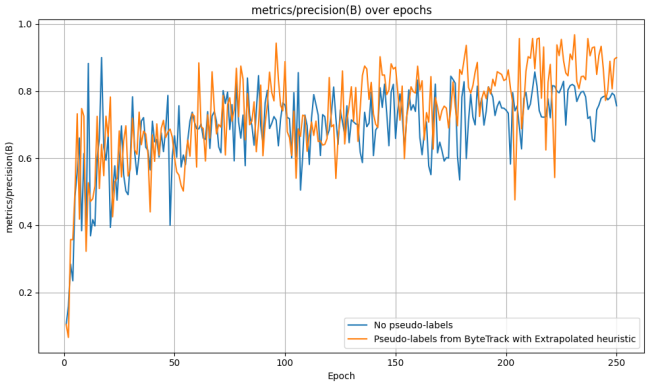
It is well known that the performance of classifiers depends heavily on the amount of training data available. Especially powerful deep learning models with a large number of parameters can achieve great accuracy, but requires large amounts of high quality annotated data [15]. Semi-supervised training is seen as one way to obtain large labeled data sets at low cost, but it is important to ensure that the quality of the pseudolabeled data is sufficiently high. In addition, to optimize learning, the classifier needs to be provided with hard cases [16][17]. Using pseudolabels where the classifier has high confidence can ensure that the new annotations are trustworthy, but it also selects cases where the classifier is already strong and provide limited learning. Such an approach often leads to increased performance for the cases where the classifier is already performant (e.g., abundant species), at the detriment of poorly performing cases (rare species) [18][19]. Our use of tracking mitigates this problem. As each track is an ensemble of detections, reliability is increased. By selecting inferred observations that are supported by low confidence or missing predictions, we ensure that the new training data consists of hard cases, where the classifier has the greatest potential for improvement.

Here, we found that the most effective approach was ByteTrack combined with the extrapolated heuristic. This configuration consistently outperformed all other tested methods, as well as the baseline supervised model. The results demonstrate that leveraging temporal information through pseudolabeling significantly enhances fish detection accuracy and consistency, but also that the choice of pseudolabeling strategy and tracking algorithm is important. The substantial improvements in precision, recall, and mAP for the pseudolabeled classes, coupled with minimal negative impact on other classes, validate the effectiveness of this approach in mitigating annotation scarcity.

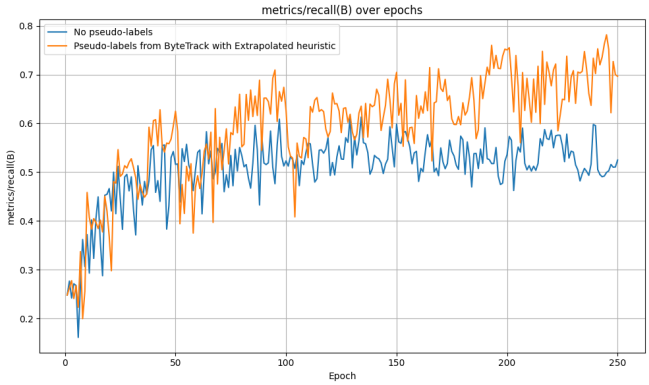
Pseudo-labels were only generated for five of the annotated classes in our labeled data, although the other species may still be present in the unlabeled videos. These species are excluded because they are rare, or perhaps observed at long distances or occluded, or due to other reasons, difficult for the current detector to recognize reliably. Occurrences of fish that are excluded by the process will be treated either as background or as incorrectly labeled fish by the subsequent training. Thus this introduces ambiguity: the detector and tracker may produce bounding boxes for objects that are outside the pseudo-label domain, but these are intentionally discarded to avoid injecting mislabeled samples into the training set. This constraint simplifies quality control, but it also means that the semi-supervised signal is intentionally restricted to a subset of the ecological variability present in the videos.

A crucial insight from this study is the progressive mitigation of initial model biases through iterative pseudolabeling. For instance, a systematic error where parts of the monitoring equipment were misclassified as "corkwing male" in early iterations (Figure 12) was effectively corrected and eliminated

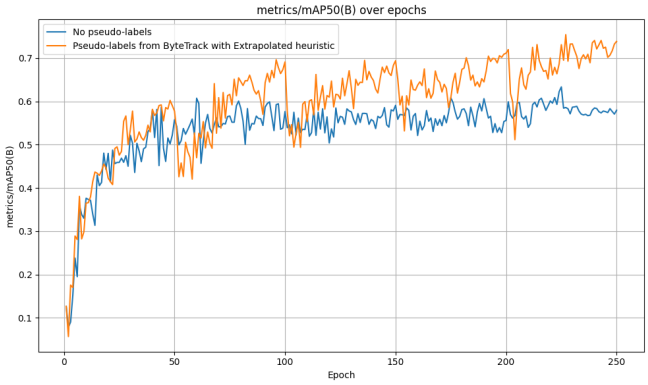
2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org



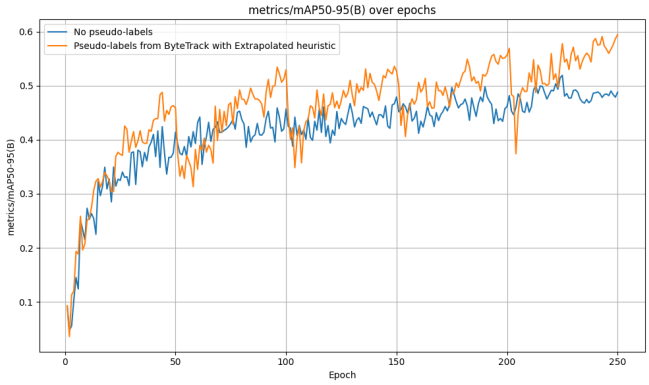
(a) Precision comparison



(b) Recall comparison

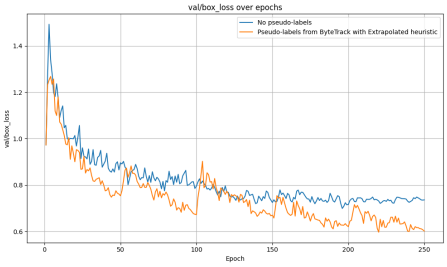


(c) mAP@50 comparison

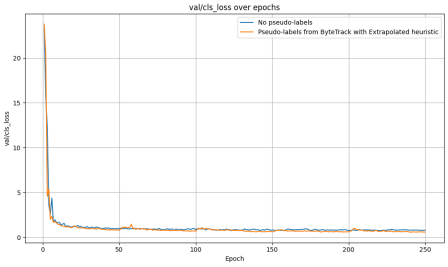


(d) mAP@50–95 comparison

Figure 9. Performance comparison across models.



(a) Val box loss



(b) Val class loss

(c) Val DFL loss

Figure 10. Validation losses during training.

in later iterations using ByteTrack with extrapolated labels.counting. Nevertheless, the iterative procedure can still correct This highlights the self-correcting nature of the temporalsuch biases when the tracker–heuristic combination does not semi-supervised framework, guiding the model towards morerepeatedly reinforce the error. In our experiments, this type of accurate predictions over time. A related observation is thatfalse positive disappeared in the final iteration for ByteTrack some systematic errors appear across multiple pipeline variantswith extrapolated labels, indicating that the evolving pseudo-early on, suggesting they originate from biases in the initial su-label sets increasingly contained counterexamples where the pervised model rather than from any single tracker or heuristic.structure was *not* labeled as a fish. This suggests an addi-In our case, a recurring false positive was linked to station-tional benefit of iterative temporal pseudo-labeling: beyond ary structures in the environment that frequently co-occurredimproving recall, it can also "wash out" certain dataset-specific with true instances of a target class in the labeled dataset.spurious correlations over time. At the same time, extrapo-Quantifying the full extent of such errors is difficult becauselation can be more susceptible to propagating false positives the erroneous boxes vary in location, shape, and confidence,(since a single error may be tracked for many frames), and we and the pseudo-labeled dataset can be too large for manualobserved indications that some tracker–heuristic combinations

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

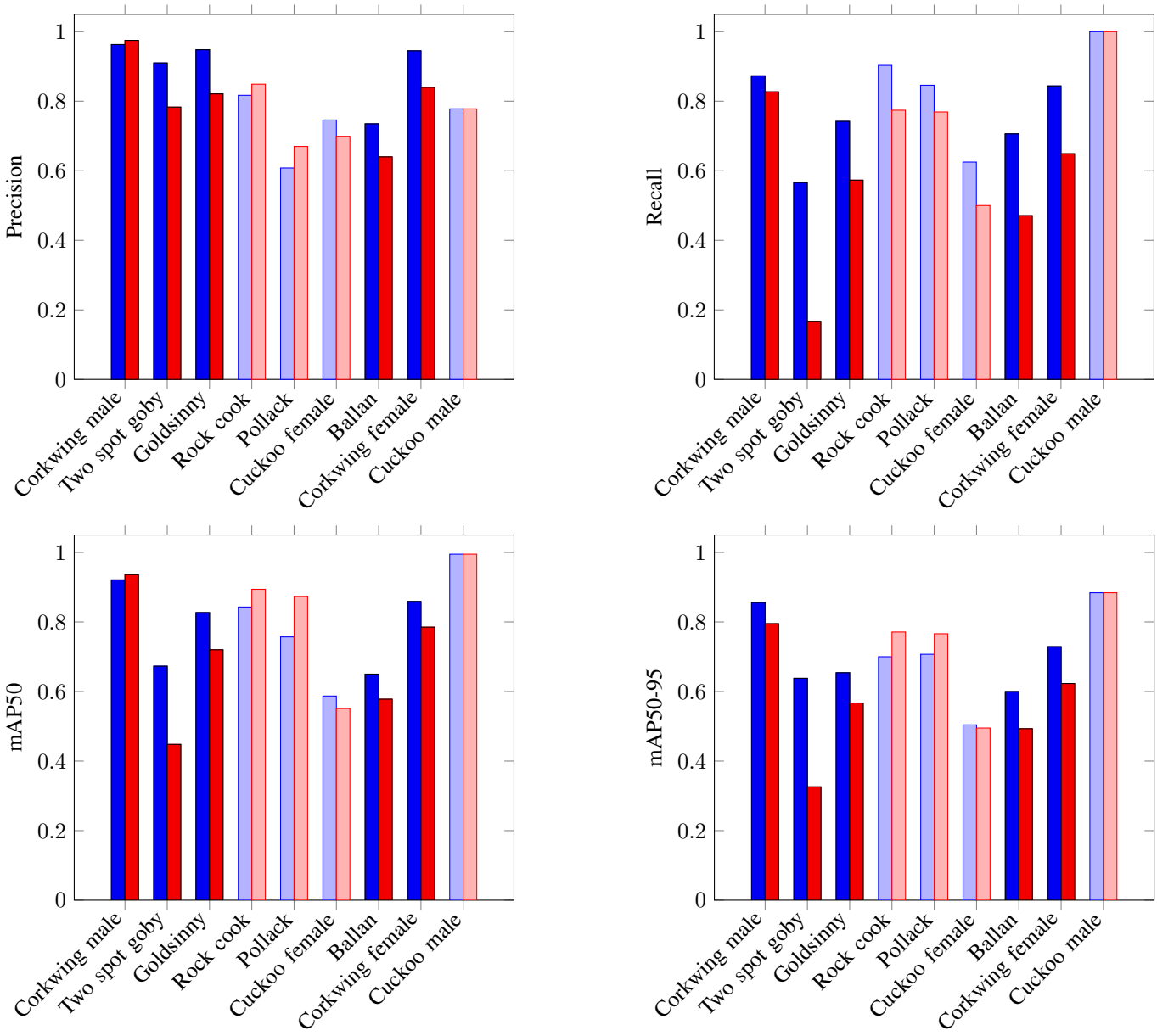


Figure 11. Precision (top left), recall (top right), mAP50 (bottom left) and mAP50-95 (bottom right) for baseline (red) and semi-supervised (blue) models. Classes not present in the pseudolabeled data shown with faded color.

may impede this bias correction compared to others.

The choice of MOT algorithm also proved critical. ByteTrack consistently outperformed DeepSORT when using extrapolated labels, but DeepSORT was better able to exploit interpolated labels. We suspect the discrepancy is caused by the use of Kalman filters in DeepSORT, which can interpolate predictions even when the object is lost by the detection model. While beneficial in predictable scenarios, this can lead to inaccurate pseudolabels for fish due to their often erratic movements, which can make DeepSORT predictions inaccurate. This phenomenon is reminiscent of the "off-policy" learning situation referred to as the "Deadly Triad" in Reinforcement Learning, which can destabilize convergence. ByteTrack, by contrast, relies solely on the detector's predictions, ensuring a stronger alignment between pseudolabels and the model's current capabilities, thus avoiding such instability. ByteTrack's reliance of overlap between detection boxes could be problematic in many settings, but for our data, a relatively high frame rate compared to fish movement speed ensures that detections of the same object from frame to frame would usually overlap.

An important nuance is that "better tracking" in a qualitative sense does not necessarily translate to "better pseudo-labels" for training. During exploratory experiments, DeepSORT often appeared visually stronger at maintaining identities, which

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

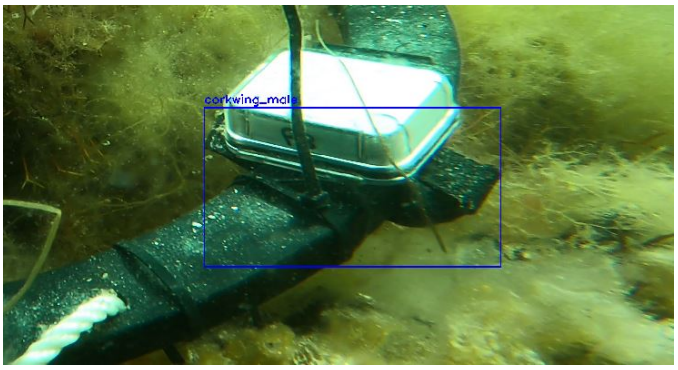


Figure 12. Equipment incorrectly predicted as "corkwing male" by the initial classifier.

made it a natural candidate for pseudo-label generation. However, when combined with the extrapolated heuristic, DeepSORT-generated pseudo-labels produced weaker downstream detector performance. A likely reason is that DeepSORT can continue producing track predictions through its motion model even when the detector has lost the object, effectively creating labels that are increasingly decoupled from what the detector currently considers plausible. This decoupling is especially problematic for fish, whose motion can be highly non-linear and strongly three-dimensional: a fish may rapidly change direction, rotate, or move toward/away from the camera such that a 2D motion model becomes a poor proxy for true dynamics. In such cases, Kalman-filter-driven interpolation can yield bounding boxes that do not match the visual evidence in the frame, and the resulting pseudo-labels can push the detector toward off-distribution targets. This aligns with the intuition behind the "Deadly Triad" analogy: when the training target (pseudo-label) becomes too disconnected from the model's own predictive distribution, iterative learning can become unstable. ByteTrack avoids this specific failure mode because its associations are anchored directly in detector-produced boxes, making pseudo-labels more "on-policy" with respect to the detector.

Semi-supervised learning has been used effectively in many different settings, but selecting pseudolabeled data to train on can be difficult. Using classifier confidence is an option [20], but tends in our experience to improve the classifier where it is already strong. Using augmentation [21] or taking class balance into account [22] may help to mitigate this, but by extracting pseudolabels from temporal information removes (or at least reduces) the dependence on the classifier itself from the selection process. Although temporal pseudolabeling methods have been attempted before (*e.g.*, [23]), our approach distinguishes itself by targeting the model's weaknesses rather than reinforcing its strengths. By relying on MOT algorithms to generate labels specifically where the base model fails to detect objects, it directly addresses gaps in detection capability.

V. Conclusion and Future Work

We have successfully demonstrated the significant potential of semi-supervised learning leveraging temporal information for enhancing object detection in marine life monitoring. By integrating YOLOv8 with ByteTrack, a robust methodology was developed to generate high-quality pseudolabels from unlabeled video data, substantially improving model performance for fish species. This approach reduces the reliance on costly and labor-intensive manual annotations, paving the way for more automation and scalable monitoring of the marine fauna. The insights gained regarding iterative bias mitigation and the critical role of MOT algorithm selection provide valuable directions for future research and practical deployment in real-world marine conservation efforts. We note that although we here target applications to marine ecosystems monitoring, we note that the techniques are general and equally applicable to the many other fields using object detection in video data.

Several practical limitations must be addressed for large-scale deployment. First, the current pipeline stores pseudo-labeled frames by extracting images from the video and writing separate label files, which is convenient but storage-intensive. A more scalable approach would treat pseudo-labels as references into the original video (video ID and frame index) alongside bounding boxes, extracting frames only on-demand during training. This would reduce storage overhead substantially and remove a clear bottleneck when scaling to many hours of footage. Second, the approach is currently reliant on a fixed per-iteration training budget (50 epochs) that is sensitive to dataset characteristics. A dynamic early stopping criterion tailored to the goal of stopping *before* convergence while learning remains sufficiently flexible could reduce compute cost and improve reproducibility across new datasets. More broadly, scaling the method will likely require more systematic exploration of per-tracker hyperparameters and heuristics, since applying the same training schedule uniformly across tracker–heuristic combinations may prevent certain variants from reaching their best performance. Finally, there are several promising directions to make the pseudo-label process more robust and operationally useful. One is to incorporate more videos and vary which videos contribute pseudo-labels across iterations (for example, starting with clear-water sequences and gradually introducing more challenging conditions). Another is to explore how and when the iterative process saturates: if the detector and tracker eventually agree on most frames, pseudo-label novelty may vanish, and continued training could yield diminishing returns or even regress toward the baseline distribution. Beyond object detection alone, improved detection consistency can also benefit downstream tasks that rely on clean trajectories, where tracking errors can otherwise propagate into higher-level behavioral or event classification.

Acknowledgment

An earlier version of this work was presented at IARIA COCE [1], and is based on results from the Master's degree thesis of VHE [24], which includes a more detailed exploration of the methods and data. The image data was collected and annotated as part of the CoastVision project, RCN grant

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

number 325862. We thank Tonje Knutsen Sørdalen and Torkel Larsen for contributing to the annotation work.

REFERENCES

[1] V. H. Elvevoll, K. T. Halvorsen, and K. Malde, "Semi-supervised object detection for marine monitoring using temporal information," in *COCE 2025: The Second International Conference on Technologies for Marine and Coastal Ecosystems*, Barcelona, Spain: IARIA, Oct. 2025, pp. 27–31, ISBN: 978-1-68558-329-3.

[2] A. W. Bicknell, B. J. Godley, E. V. Sheehan, S. C. Votier, and M. J. Witt, "Camera technology for monitoring marine biodiversity and human impact," *Frontiers in Ecology and the Environment*, vol. 14, no. 8, pp. 424–432, 2016. DOI: https://doi.org/10.1002/fee.1322. eprint: https://esajournals.onlinelibrary.wiley.com/doi/pdf/10.1002/fee.1322. [Online]. Available: https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1002/fee.1322.

[3] S. K. Whitmarsh, P. G. Fairweather, and C. Huveneers, "What is Big BRUVver up to? methods and uses of baited underwater video," *Reviews in Fish Biology and Fisheries*, vol. 27, no. 1, pp. 53–73, 2017, ISSN: 1573-5184. DOI: 10.1007/s11160-016-9450-1. [Online]. Available: https://doi.org/10.1007/s11160-016-9450-1.

[4] H. Liu, X. Ma, Y. Yu, L. Wang, and L. Hao, "Application of deep learning-based object detection techniques in fish aquaculture: A review," *Journal of Marine Science and Engineering*, vol. 11, no. 4, p. 867, 2023.

[5] P. Rubbens et al., "Machine learning in marine ecology: An overview of techniques and applications," *ICES Journal of Marine Science*, vol. 80, no. 7, pp. 1829–1853, 2023. DOI: 10.1093/icesjms/fsad100.

[6] J. E. Van Engelen and H. H. Hoos, "A survey on semi-supervised learning," *Machine Learning*, vol. 109, no. 2, pp. 373–440, 2020.

[7] J. D. Poling et al., *Fjordfish*, version 1.1, Zenodo, Oct. 2025. DOI: 10.5281/zenodo.17950781. [Online]. Available: https://doi.org/10.5281/zenodo.17950781.

[8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, *You only look once: Unified, real-time object detection*, 2016. arXiv: 1506.02640 [cs.CV]. [Online]. Available: https://arxiv.org/abs/1506.02640.

[13] S. Liu, J. Zhang, H. Zheng, C. Qian, and S. Liu, "An improved deepsort-based model for multi-target tracking of underwater fish," *Journal of Marine Science and Engineering*, vol. 13, no. 7, 2025.

[14] W. Li et al., "When trackers date fish: A benchmark and framework for underwater multiple fish tracking," *arXiv preprint arXiv:2507.06400*, 2025.

[9] Y. Zhang et al., *Bytetrack: Multi-object tracking by associating every detection box*, 2022. arXiv: 2110.06864 [cs.CV]. [Online]. Available: https://arxiv.org/abs/2110.06864.

[10] N. Wojke, A. Bewley, and D. Paulus, *Simple online and realtime tracking with a deep association metric*, 2017. arXiv: 1703.07402 [cs.CV]. [Online]. Available: https://arxiv.org/abs/1703.07402.

[11] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple online and realtime tracking," in *2016 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2016, pp. 3464–3468.

[12] M. Abouelyazid, "Comparative evaluation of sort, deepsort, and bytetrack for multiple object tracking in highway videos," *International Journal of Sustainable Infrastructure for Cities and Societies*, vol. 8, no. 11, pp. 42–52, 2023.

[15] Y. Zhou, F. Tu, K. Sha, J. Ding, and H. Chen, "A survey on data quality dimensions and tools for machine learning," *arXiv preprint arXiv:2406.19614*, 2024.

[16] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2980–2988.

[17] R. Baldock, H. Maennel, and B. Neyshabur, "Deep learning through the lens of example difficulty," *Advances in Neural Information Processing Systems*, vol. 34, pp. 10 876–10 889, 2021.

[18] C. Wei, K. Sohn, C. Mellina, A. Yuille, and F. Yang, "Crest: A class-rebalancing self-training framework for imbalanced semi-supervised learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 10 857–10 866.

[19] L.-Z. Guo and Y.-F. Li, "Class-imbalanced semi-supervised learning with adaptive thresholding," in *International Conference on Machine Learning*, PMLR, 2022, pp. 8082–8094.

[20] K. Sohn et al., "A simple semi-supervised learning framework for object detection," *arXiv preprint arXiv:2005.04757*, 2020.

[21] J. Jeong, S. Lee, J. Kim, and N. Kwak, "Consistency-based semi-supervised learning for object detection," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[22] M. Xu et al., "End-to-end semi-supervised object detection with soft teacher," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 3060–3069.

[23] R. J. Veiga et al., "Autonomous temporal pseudo-labeling for fish detection," *Applied Sciences*, vol. 12, no. 12, p. 5910, 2022.

[24] V. H. Elvevoll, "Semi-supervised object detection using temporal information," M.S. thesis, Department of Informatics, The University of Bergen, 2025.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Socially Responsible Artificial Intelligence in People and Pandemic Analytics: A Resilience-Oriented Framework for Sustainable Decision-Making

Otabek Khamidov, Les Sztandera
Thomas Jefferson University, Jefferson School of Business
Philadelphia, Pennsylvania, United States of America
Emails: Otabek.Khamidov@students.jefferson.edu, Les.Sztandera@jefferson.edu

Abstract— **This paper advances the Socially Responsible Artificial Intelligence framework to Pandemic Analytics in response to the COVID-19 pandemic. The original SRAI framework was proposed to inform the ethical adoption of artificial intelligence in People Analytics and Human Resource Development. However, the pandemic created the necessity to extend the principles to other high-risk fields like public health, crisis management, and healthcare delivery. Based on a qualitative synthesis of peer-reviewed articles between 2020 and 2025, this research develops the SRAI framework by proposing a sixth dimension known as Resilience Responsibility. The new addition reflects the importance of designing AI systems to be trustworthy, flexible, and capable of delivering even in high-pressure situations. The research demonstrates how AI contributed to business responses as well as public health responses during the pandemic but also the research findings highlighted concerns about data bias, privacy, and accountability. The enhanced framework provides actionable recommendations for HR practitioners, healthcare leaders, AI engineers, and policymakers to ensure the adoption of AI is ethical, lawful, sustainable, and resilient to disruptions.**

Keywords - Artificial Intelligence; People Analytics; Pandemic Analytics; COVID-19 data.

I. INTRODUCTION

Artificial intelligence has rapidly transformed human resource development and organizational decision-making [2]. This section provides an overview of the origins of People Analytics, the foundation behind developing the Socially Responsible Artificial Intelligence (SRAI) framework as well as the ethical and sustainability challenges that shaped its foundation.

A. Emergence of People Analytics in HRD

People Analytics (PA) is a developing field in Human Resource Development (HRD), which stands out for emphasizing data-driven "decision science" over intuition-based decisions [2][55][3]. PA involves collecting and analyzing workforce data to guide HR strategies and practices moving HR away from a traditionally experience-based field towards evidence-based decision making. PA is widely applied in various HRD activities, including talent acquisition, skills and competency analysis, employee sentiment analysis, performance management, turnover prediction, and training and development, to inform superior decisions regarding people and talent management [2].

Although PA provides a more efficient, objective, and strategic approach to personnel management, it also raises ethical challenges and legal obligations. For instance, Workday Inc. is accused of utilizing discriminatory AI technology for the job candidate screening process based on age, disability, and race in one recent active lawsuit [5]. Despite these obstacles, PA has grown to be a crucial component of contemporary HRD and signals a move toward more scientific and technology-driven approaches to managing people.

B. Introduction to the SRAI Framework

Addressing the demand for responsible AI in HR in response to the growing ethical and sustainability challenges, a comprehensive Socially Responsible Artificial Intelligence (SRAI) framework was introduced for people analytics [2]. This was one of the first efforts at systematically connecting the concept of Corporate Social Responsibility (CSR) to people analytics on AI-facilitated Human Resource Development. In an extension of classic CSR pyramid and corresponding sustainability paradigms [6], SRAI offers a five-stage model for an organization's economic, legal, ethical, philanthropic, and environmental responsibilities for AI usage with AI-driven HR practices to be followed.

The base layer is economic responsibility, which implies people analytics AI tools being usable, dependable, and delivering organizational performance directing that AI in HR needs to add value, deliver return on investment while minimizing risks. Next is legal responsibility, which requires AI systems to follow legal mandates around data usage, human rights, labor laws in employment-related decision-making, and intellectual property. Ethical responsibility takes a step beyond lawfulness, including standards of fairness, transparency, and respect for privacy of design and deployment of AI beyond what the law requires. Philanthropic responsibility involves a voluntary commitment of Human-Centered AI application for broader social good, for example, application of people analytics for increasing employee and community outcomes, which shows a vision of HRD having beneficial impacts beyond organizational immediate interest. Finally, the model places an environmental responsibility for realizing that AI adoption needs to be ecologically durable. Sustainable AI can help reduce energy consumption and carbon footprint through application of AI for environmental goals achievement.

SRAI framework is stakeholder-centered as it involves recognizing key stakeholders ranging from employees and

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

managers through job applicants to society at large impacted by AI. SRAI framework charts how each responsibility level translates into goals and requirements for these stakeholders [2]. The integrative literature review used in this study spanning up to 2023 publication covered related concepts like Environmental, Social, and Governance (ESG) criteria and United Nations Sustainable Development Goals (SDGs) for a comprehensive idea of sustainable AI-powered HR.

To guide the reader through the structure of this paper, the remaining sections are organized as follows. Section II - Research Gap identifies the limitations of the original SRAI framework and explains the need to address high-risk, cross-sector applications such as Pandemic Analytics. Section III - Methodology describes the qualitative approach and data collection process used to refine the framework. Section IV - Findings presents the main results, highlighting how artificial intelligence was applied in both organizational and public health contexts during the COVID-19 pandemic. Section V - Practical Implications links theory to practice through converting findings into actionable recommendations for People Analytics and Pandemic Analytics stakeholders. Finally, Section VI - Conclusion and Future Work provides a summary of the study and outlines future directions for research on socially responsible and resilient AI systems.

II. RESEARCH GAP

SRAI framework was originally designed for ethical practice in People Analytics and Human Resource Development with a focus on responsibilities for corporate economic performance, legal compliance, and fair organizational decision-making processes. While it is an important milestone for measuring AI adoption within corporate environment, it has not yet been adapted for high-stakes, real-time, multi-industry applications like Pandemic Analytics. The COVID-19 pandemic revealed the need for socially responsible AI for public health, healthcare systems, and disaster management encompassing broader stakeholder groups, urgent decisions, and more ethical risks. The application of SRAI framework to Pandemic Analytics addresses the gaps and identifies opportunities for pandemic-era data and governance requirements.

Moreover, developments in recent years have raised new gaps calling for a revised viewpoint on SRAI. Firstly, AI machine learning models and generative language tools are creating ethical concerns, such as AI "hallucinations" [7] and other new sources of bias which should be addressed by SRAI frameworks. The study published on behalf of the United States & Canadian Academy of Pathology, raises the concern over potential AI bias due to three main factors, such as, data bias, development bias, and interaction bias which can inadvertently result in unfair and potentially detrimental outcomes within pathology and medical domain [8].

Secondly, the legal and regulatory landscape around AI has evolved rapidly presenting new challenges for public regulators to implement effective administrative interferences [9]. Since 2023, regulators both in the United States of America and internationally, such as in the European Union are trying to impose bias audit mandate for AI algorithms as well as automated decisions affecting human resources. For example, starting from February 1, 2026, the state of Colorado will be the first U.S state to require organizations to identify and mitigate algorithmic discrimination risks for high-risk AI systems [10]. At the U.S federal level, Congress proposes the AI Whistleblower Protection Act (H.R. 3460) which would protect individuals who report unethical, biased, and illicit AI practices in their workplaces, such as automated hiring and surveillance methods [11].

Finally, there is an emerging emphasis on environmental sustainability in AI. The energy usage and carbon output of AI systems have raised alarm bells regarding the environmental impact of AI. Latest estimates indicate that training and running big AI models can release massive amounts of CO_2 [12]. The study shows based on data from 275 Chinese cities that the carbon footprint contribution of digitization and artificial intelligence is underestimated because the effects increase the carbon impact by 665% [13].

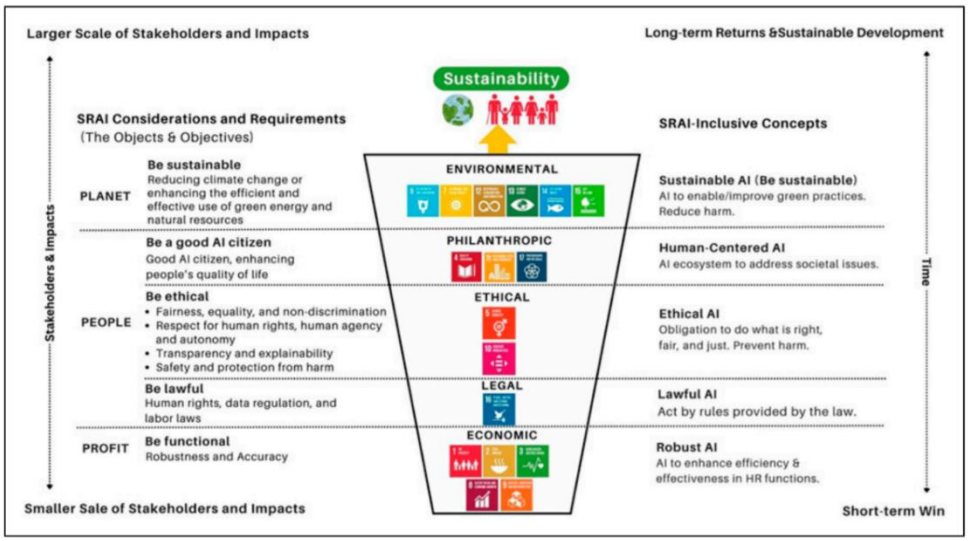


Figure 1. The original SRAI framework. Adapted from [2].

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

III. METHODOLOGY

This study provides a qualitative review of existing literature to explore the impact of the COVID-19 pandemic on the use of AI in both Pandemic Analytics and People Analytics. The purpose of the paper is to update the original SRAI framework by analyzing recent research and emerging trends in People Analytics, and to apply the framework to Pandemic Analytics. We conducted research in academic databases, including Google Scholar, ScienceDirect, MDPI, Emerald Insight, and other reputable industry sources based on articles published between 2020 and 2025. Our search employed keywords like "AI," "People Analytics,", "Pandemic Analytics", "COVID-19," and "HR."

During this process, we explored a new layer called "Resilience Responsibility". Pandemic-focused research articles, reports from Deloitte, and recent news concerning responsible AI also supported the establishment of this new category. This method ensures our framework incorporates both academic expertise and real-world practices.

IV. FINDINGS

This section explores the updated SRAI framework into actionable recommendations for practitioners across industries. It provides clear guidance on how economic, legal, ethical, philanthropic, environmental, and resilience responsibilities can be implemented in both People Analytics and Pandemic Analytics contexts.

A. Organizational Use of AI During the Pandemic: Lessons from People Analytics

The COVID-19 pandemic significantly accelerated the use of digital technology in Human Resources Management. As lockdowns and social distancing disrupted normal workplace operations, companies had to shift to remote work and manage teams virtually. Consequently, many organizations adopted more advanced HR technologies, including artificial intelligence, faster than initially expected. When physical offices were closed, AI tools were increasingly beneficial for maintaining important workforce management functions such as recruiting, employee training, performance management, and operational efficiency. These tools also helped HR teams plan for staffing needs and identify gaps in employee skills as business needs changed during uncertain times [14][15]. For example, IBM replaced hundreds of employees with AI which handles 94% of routine HR tasks. The IBM case reflects a post-pandemic trend of rethinking the purpose of the corporate jobs and the corporate structure around AI as repetitive tasks are increasingly automated by AI [16]. Not only has AI adoption accelerated in the aftermath of the COVID-19 pandemic, but also companies already moved on beyond experimentation phase and into full-scale implementation.

While operational and functional needs of human resources management are met with AI tools, HR leaders also assumed a new set of responsibilities in protecting employees' health, morale, and productivity in Work from Home (WFH) setting [15]. AI-powered systems became more in-demand in solving this issue through monitoring employee engagement and well-being without continuous in-person oversight. For example, companies like IBM, Microsoft, Softbank, and Unilever are utilizing emotional recognition technology, also known as Emotional AI, to monitor employee engagement, well-being, attentiveness and interest in remote meetings [17]. More recent survey of 2025 on AI and employee wellbeing in the workplace among 207 companies found that the adoption of AI does not directly impact employee wellbeing but instead affects it indirectly through factors related to work such as task optimization and safety. Research emphasizes the importance of AI adoption leading to real improvements in organizational tasks and processes, safeguarding data, and promoting occupational health. When these factors are properly addressed, AI adoption can improve employee wellbeing and foster better human-AI collaboration in the workplace [18] The use of AI in corporate world, especially in human capital development and people analytics have already become an urgent priority for organizations instead of gradual digital transition to sustain workforce performance and employee well-being because of pandemic disruptions, normalized remote work, and increasing value of AI tools.

B. Generative AI in People Analytics: Post-Pandemic Expansion, Ethical Risks, and Governance Challenges

Following the acceleration of traditional AI tools in a rapid digitization of HR, such as, remote work tools, workforce planning, and performance tracking, a second phase of AI adoption emerged in the form of generative artificial intelligence (GenAI) after the pandemic's peak which started reshaping people analytics due to the Gen AI's capacity in coherent, contextually relevant, convincing, and human-like text, images, and audio generation [19]. This second phase of AI adoption is aimed to streamline important HR workflows. For example, Large Language Models can assist from screening resumes, crafting tailored job descriptions to complicated people analytics decisions like creating personalized training materials which aligns with employee's job responsibilities and learning preferences. Reference [20] reveals the impact of Gen AI and neural networks on automating payroll operations, analytical workforce insights, and digital employee payments. Companies started using Gen AI models to forecast compensation trends, discover payroll anomalies, and suggest data-driven modifications based on past patterns and real-time performance metrics. For example, according to Amazon CEO Andy Jassy, the company is applying Gen AI to improve internal productivity including analyzing workforce data more efficiently and decision-making pace [21].

However, despite the opportunities, Generative AI introduces major new risks and challenges in HRD. From ethical and legal responsibility perspective, bias is the primary concern in LLMs inheriting human prejudices and resulting in discriminatory outcomes during hiring, performance reviews and compensation decisions. Limitations in emotional intelligence and ethical accountability also shows the lack of empathy needed for delicate HR tasks like layoffs or performance discussions which result in impersonal judgements and undermining workplace fairness and

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

employee morale [22]. Recent research investigating bias in AI-driven HRM systems revealed that AI-driven HRM systems disproportionately impact marginalized communities, including women, racial minorities, and non-binary people. The biases in training data and algorithm design can result in discriminatory impacts, even if these systems can improve productivity and objectivity in areas like hiring and workforce analysis. These biases may limit the employment and career advancement prospects of underrepresented groups. Existing legal and ethical frameworks are often inadequate since they frequently rely on binary understandings of identity and fail to offer comprehensive protection. Enhancing legal safeguards, adopting inclusive HRM practices, boosting transparency, conducting frequent bias tests, and ensuring significant human oversight are all essential to ensuring fairness in AI-driven decision-making [23].

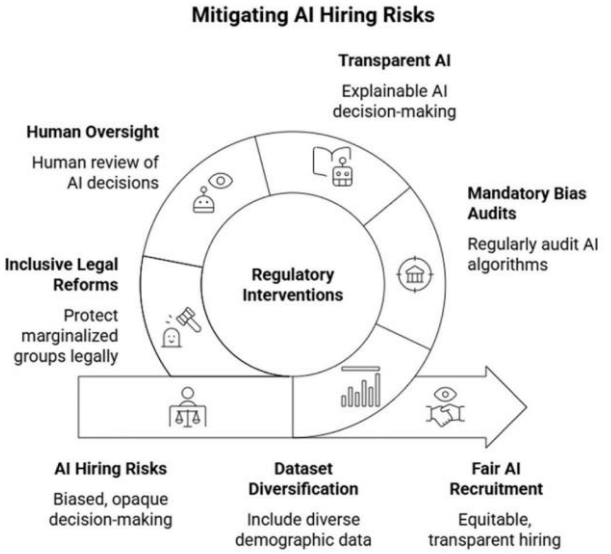


Figure 2. AI-driven hiring risks, regulatory gaps, and recommended interventions, highlighting biases, enforcement gaps, and solutions such as diverse datasets, bias audits, transparency, human oversight, and inclusive legal reforms. Adapted from [23].

The second major concern involving accountability and resilience risk is hallucinations, which are the incorrect or misleading outputs produced by Generative AI with high confidence. Hallucination can result in incorrect recommendations in sensitive areas, such as termination reasons or workplace grievance which undermines trust and possibly leads to legal liability. The third yet equally important concern is privacy and data protection. Gen AI models are trained on immense datasets. If sensitive employee data is not safeguarded against external AI tools, Gen AI may collect and process the data violating privacy regulations and such as the Americans with Disability Act or HIPAA [24]. Moreover, the European Journal of Business and Management Research emphasizes regulatory uncertainty as a growing concern on how Gen AI is being applied to people analytics [25]. Because regulations fall behind the technological development, companies not only in Europe, but also in USA happen to operate in legal "grey zone". For example, Meta's acquisition of 49% nonvoting stake in Scale AI for $14.8 billion, Amazon's 2024 partnership with Adept AI and gaining access to propriety models, Microsoft's $650 million investment in Inflection AI cases highlight governance concerns regarding the rapid scaling of AI capabilities relative to existing antitrust, transparency, and ethical oversight mechanisms [26]. Therefore, lack of international laws and standards for Gen AI reinforces the need for self-imposed ethical frameworks and industry tailored best practices for sustainable development.

C. Pandemic-Era Transformation of People Analytics: Bias, Ethics, and Strategic Adaptation

Although Pandemic Analytics is associated with healthcare and public health, the COVID-19 crisis transformed organizational behavior. Latest studies find that the COVID-19 pandemic served as a "career shock" and it has fundamentally changed the importance of People Analytics as main factors of Human Resource Development. HRM evolved from a merely administrative role to a strategic leadership position by focusing on redesigning work culture during the crisis. The use of Artificial Intelligence to enhance People Analytics in HRD increasingly accelerated between 2020 and 2025. Findings show that AI is transforming HR by ensuring employee safety and well-being [27], promoting adaptable workforce practices [28], improving employee performance [29], measuring employee engagement [27], and employee resilience [30] during and after Covid-19 pandemic. Increasing reliance on data-driven strategies highlights how organizations can benefit from AI in facilitating flexible work arrangements, workforce planning, upskilling, and fair performance assessments.

With remote work, new generations, and greater fairness expectations transforming company policies in the post-COVID-19 workplace, leaders should be aware of the increasing threat of hidden bias in AI and people analytics [31]. Rapid post-pandemic implementation of AI can hide or amplify bias if leaders fall into the false assumption that algorithms are unbiased. As organizations become more reliant on people analytics for hiring, career advancement, and performance reviews, ethical blind spots must be addressed proactively through bias awareness, such as leadership training, frequent audits of bias in AI tools, and stronger accountability procedures to detect, reveal, and reduce bias, whether through human effort or through machine [32].

Research further showed that many studies explored the ethical implications of People Analytics during its historic rise following the COVID-19 pandemic [20][21][22]. Studies expressed concern regarding the data-driven decision-making process and the processing of sensitive information about employee behaviors, well-being, and emotional state. For example, organizations are using AI-powered tools like Microsoft Viva to track work trends and identify risks of employee burnout. While these tools aim to improve productivity and engagement, critical ethical concerns are raised regarding privacy, algorithmic bias, and workforce autonomy [33]. In order to promote operational efficiency,

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

employee satisfaction, and organizational adaptability, the 5P model (Purpose, People, Process, Performance, and Partnership) was proposed as a solution and a framework for purposeful, ethical, and people-centered implementation of AI in post-COVID organizational practices [36].

Collectively, these organizational experiences highlight the urgent need for resilient, ethical, and transparent decision-making systems. The lessons from People Analytics provide important insights into the role of AI functioning under systemic pressure and SRAI's applicability for data governance and pandemic preparation.

D. Artificial Intelligence Empowered Pandemic Analytics: Innovations and Opportunities

The pandemic has highlighted the role of AI-driven data analytics extending beyond HR and reshaping sectors like healthcare, public health surveillance, and crisis management, where the most sensitive and protected health data are analyzed. Understanding the role of the pandemic in creating opportunities and ethical risks in the healthcare industry, especially regarding data privacy, bias, and oversight will be valuable in formulating more human-centered and socially responsible AI principles.

The COVID-19 pandemic accelerated the pace of AI-driven analytics innovations across healthcare and public health. In the early stages of the pandemic, researchers highlighted AI's important potential to help with prediction, detection, control, and treatment. For instance, AI-based epidemiological models were used to predict the spread of the disease, and deep learning systems were applied to medical images for the diagnosis of COVID-19 from chest scans [37]. Furthermore, platforms like BlueDot employed natural language processing and machine learning for identifying early COVID-19 symptoms based on social media reports as well as health reports, and Metabiota employed predictive modeling with traveler data and population density to forecast outbreak dynamics. AI has also optimized telehealth service delivery. For example, Ada Health supported public health responses via AI-based chatbots offering symptom check-ups along with affordable telehealth services during periods of restricted mobility [38].

The role of AI in healthcare goes beyond diagnostics. AI-powered predictive analytics have been useful for clinical decision support. A systematic review (2020–2022) found numerous machine learning models that predicted intensive care unit admission and mortality risk for COVID-19 patients using combinations of clinical variables. These models provided healthcare professionals with early warnings of high-risk cases and enabled them to deliver proactive care [39]. Recent developments between 2024 and 2025 have significantly advanced AI-driven pandemic intelligence by enhancing multilingual natural language processing, hybrid modeling techniques, and forecasting capabilities based on large language models (LLMs) [40][41][42][43][44][45]. For instance, PandemicLLM integrates epidemiological, policy, and behavioral data and it has demonstrated improved performance over traditional time-series models in forecasting infectious disease outbreaks [41]. Another system applied multilingual LLM ensembles to analyze unstructured health reports and news feeds, which improved detection speed and contextual relevance in public health settings [40]. The SIR-INN model incorporates real-time influenza data into a physics-informed neural network based on a modified SIR framework, contributing both interpretability and forecasting accuracy [45]. Additionally, a dual-branch architecture EpiLLM integrates spatio-temporal epidemic trends with mobility data to enable localized predictions of disease spread [46].

Beyond healthcare, COVID-19 has greatly advanced the adoption of AI and "smart" technologies across eight major industries, such as food services and manufacturing. Thirty-nine distinct kinds of smart technologies powered around 40 types of pandemic use cases, including remote communication, healthcare service delivery, data analytics, and logistics. For example, online education platforms with AI tutors replaced in-person classes during lockdowns and AI-enabled robots assisted in hospitals to examine patients as well as deliver medications while reducing infection risk [47]. Furthermore, AI mobile health apps like mHealth has the potential to revolutionize post-pandemic public health surveillance by automating illness forecasting, outbreak detection, and resource management [48].

The recent developments in AI have also increased the number of applications in predictive modeling of outbreaks, healthcare delivery optimization, and public health surveillance which can be used in future pandemics. For example, tools like epitweetr and Open Source Intelligence (OSINT) are used to analyze social media and environmental data for threat detection with vast geographic scope. The Machine Learning algorithms have the ability to forecast outbreaks based on input data of population density, weather, and vector movement. Not only do these algorithms outperform traditional statistical methods, but also AI models prevent supply and communication disruptions through resource allocations like oxygen supply in hospitals, and through developing public health warnings with the help of Gen AI language models [49].

E. Ethical, Governance, and Equity Challenges in Pandemic Analytics

The swift implementation of AI to pandemic responses also raised major concerns. Non-standardized datasets complicate validation which results in inconsistent performance and erosion of trust. Worldwide efforts by organizations such as World Health Organization, Centers for Disease Control and Prevention, and commercial software companies have tried to establish and standardize large-scale datasets, such as CORD-19 repository. Data security and privacy are also long-standing issues because pandemic surveillance is at odds with personal data protection [38]. Many of the reviewed studies involved sensitive personal data, which should be handled carefully even during a pandemic. Therefore, the pandemic emergency offered a valuable opportunity for more ethical and responsible action [39]. The studies raised concerns about ethical issues related to privacy, fairness, and accountability. Transparency in data sources and AI models is essential to building trust among the public and healthcare providers. Regular testing for biases and

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

continuous monitoring is necessary to avoid unfair treatment of marginalized groups. Overall, AI should not replace human expertise and judgment but rather supplement them in managing the pandemic [37].

The ethical use of predictive analytics in public health shows considerable challenges in pandemic contexts where sensitive data such as social media, mobility, and biometrics are used for surveillance. These practices raise concerns about consent, data protection, and the potential for overreach. A recent study identified four core ethical concerns in predictive health tools: privacy violations, algorithmic bias, lack of transparency, and inequitable access. Each of these issues requires mitigation strategies, including anonymization protocols, fairness metrics, explainable AI models, and open-source tools that support low-resource environments. To address these concerns, ethical frameworks should incorporate transparency techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations), as well as governance measures including federated learning, community consultation, and oversight by ethical review boards [50].

TABLE I. ETHICAL RISKS AND MITIGATION STRATEGIES IN PREDICTIVE PUBLIC HEALTH TOOLS. ADAPTED FROM [50].

Ethical Concern	Risk Description	Recommended Mitigation Strategy
Privacy Violation	Unauthorized use of personal data	Enforce anonymization, consent protocols, and audits
Algorithmic Bias	Discriminatory outputs affecting certain groups	Apply fairness metrics, diverse data sampling
Lack of Transparency	Opaque decision-making in black-box models	Use explainable AI methods and publish model logic
Inequitable Access	Regional disparities in predictive capabilities	Build open-source models and support low-resource tools

There are also obstacles to large-scale AI implementation. The reliability of AI models is of first concern, as many AI models were deployed with little peer review during the pandemic. The application of many sophisticated AI technologies in low-resource and low-income countries is limited because AI tools are created using data from high-income countries which employ robust digital infrastructure. International cooperation, the sharing of models and data, and region-specific AI solutions are crucial for improving global health disparities [51].

V. PRACTICAL IMPLICATIONS FOR PEOPLE ANALYTICS AND PANDEMIC ANALYTICS STAKEHOLDERS

The COVID-19 pandemic presented the possibility for AI systems to contribute to organizational agility and pandemic response but also it uncovered risks surrounding system vulnerability, data misuse, and bias. Practical implications offer actionable guidance for all types of stakeholders from healthcare administrators to leadership policymakers, data scientists, AI engineers, and to HR professionals to ensure ethical, legal, sustainable, and resilient AI deployment for both health emergency scenarios as well as work environments.

A. Economic Responsibility: Be functional

The SRAI model's economic component highlights the use of AI to improve productivity, resource allocation, and efficiency. In HR, it involves AI-powered workforce planning, monitoring of engagement, and hybrid work plan. In healthcare, AI models help to forecast intensive care unit admissions, automate personal protective equipment delivery, and mitigate critical care delays. However, short-term financial gains should be balanced with longer-term investments in people and technology. During the COVID-19 pandemic, data-driven efficiency allowed many organizations to pivot quickly, however, businesses also learned that over-dependence on testing-phase algorithms and taking people factors for granted can have negative long-term impacts. Therefore, the updated framework encourages professionals to pursue the economic benefits of People Analytics and Pandemic Analytics as well as implementing internal mechanisms to fulfill legal, ethical, and social obligations

B. Legal Responsibility: Be lawful

The AI systems are subject to applicable labor laws, health privacy laws like HIPAA, data protection regulations, and civil rights protections. The lesson from pandemic management is that even during emergencies, personal data must be handled carefully and in accordance with privacy principles as ethical and legal standards will be vulnerable during a crisis. Compliance with law guarantees data openness, fair play, and stakeholder confidence. Practitioners are advised to keep records on sensitive data processing, monitor AI models for bias, and create systems where employee data as well as patient data are secured during business-as-usual operations and emergencies.

C. Ethical Responsibility: Be ethical

Ethical considerations in AI-enabled PA require more than just following laws, as they encompass fundamental values such as dignity, fairness, explainability, and harm prevention. Post-pandemic AI use showed the lack of human-centered design in the applications of performance tracking and hiring practices. Taking into account AI-related nuances and anomalies, industry leaders should make the final decisions and ensure that stakeholders are informed about data collection, processing, and AI-driven decision-making practices. The “A human-in-the-loop” strategy helps balance algorithmic input with discretion and empathy.

D. Philanthropic Responsibility: Be a good AI citizen

An organization's voluntary attempts to employ AI for the greater benefit are reflected in its philanthropic responsibilities. This layer involves using Data Analytics to promote community involvement, inclusion, and employee well-being above and beyond simple compliance. People Analytics could help community workforce programs through HR departments, and public health agencies could share aggregated models of AI for balanced disease surveillance. These are all about taking a proactive approach for inclusion, for public confidence, for health, and for supporting the Sustainable Development Goals even when economic return is not the mission.

E. Environmental Responsibility: Be sustainable

The SRAI pyramid's top level, environmental consideration, encourages AI systems to support ecological

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

sustainability. With the additional benefit of reducing carbon emissions due to less frequent travel and Telehealth, the pandemic showed that widespread remote business and health operations are feasible. Leaders can use AI-enabled analytics to measure these effects and develop policies that support climate goals and enhance work-life balance. In HR it enables "Green HRM" programs; in healthcare and logistics, it could assist in the creation of sustainable supply chains.

F. Resilience Responsibility: Be future-ready

The findings from research papers explored in the industries from HR to healthcare to crisis management opened a new perspective for socially responsible AI, which is resilience responsibility. This specific responsibility emerged due to volatility and systemic uncertainty marked by the pandemic. Resilience Responsibility is built upon the SRAI framework by integrating principles from organizational resilience theory, research on socio-technical systems, and literature related to high-reliability organizations. In this framework, resilience is defined as an organization's ability to predict disruptions, absorb shocks, adapt under pressure, and recover while maintaining essential operations. Conventional responsible AI frameworks focus on compliance, fairness, and sustainability within stable environments. However, the COVID-19 pandemic revealed that ethical AI governance must also consider extreme uncertainty, urgency, incomplete information, and rapidly changing risk landscapes. In these situations, AI systems not only aid in decision-making but also actively influence outcomes across areas like employment, access to healthcare, and public trust.

In contrast to economic, legal, ethical, or environmental responsibilities, which mainly focus on AI operations under standard conditions, Resilience Responsibility specifically deals with systemic vulnerabilities and the potential for failure spread when AI systems are implemented at scale during crises. Insights from Pandemic Analytics demonstrated that inaccuracies in models, biases in data, or unclear automation can quickly escalate, causing widespread harm across populations instead of merely impacting separate organizations. This difference supports the idea that resilience should stand alone as a responsibility rather than being considered merely a subset of ethics or compliance. In practical terms, Resilience Responsibility involves ongoing monitoring of models, stress testing, protocols for human intervention, the ability to conduct audits, and fail-safe measures that enable organizations to halt, override, or adjust AI systems when they exceed risk limits.

Practical implementation of resilience responsibility can include AI-driven scenario planning, early warning models, simulation tests for disasters or cyberattacks, and identifying system vulnerabilities. In pandemic analytics, resilience means using AI to link outbreak data with hospital resource planning, helping teams adjust triage and distribute supplies more effectively during crises. This supports faster, more adaptive responses and reinforces the SRAI framework's focus on ethical, legal, and resilient AI use in both healthcare and HR settings. Moreover, recent news concerning the Grok AI incident, which shared antisemitic content on the X platform, or the Open AI incident, which tried to copy itself to external servers during shutdown, reveals the demand for resilient AI architectures [31][32]. If the AI-enabled automation is left unsupervised, these advanced tools can replicate or magnify societal harms. A recent report by Deloitte also projects that natural disasters like the COVID-19 pandemic can cause US$460 billion in average annual losses to infrastructure globally. However, US$70 billion of the total loss amount can be saved annually if infrastructure resilience is enhanced with AI [54]. Overall, resilience responsibility is long-term insurance that serves as a safety net to withstand disruptions, to adapt to uncertainty, and to align with human values, safeguarding both society and innovation.

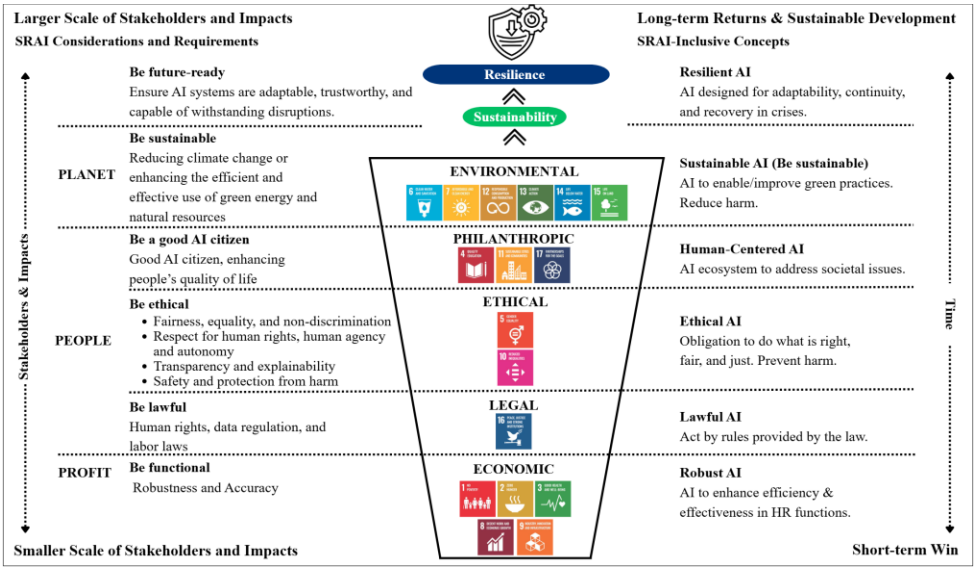


Figure 3. The updated SRAI Framework linking responsibility dimensions with Resilient AI development. Adapted from [1].

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

VI. Conclusion and Future Work

This paper studied the development of the SRAI framework extending it from originally developed for the People Analytics in HRD field to the new domain of Pandemic Analytics. With the help of qualitative analysis in the most recent research papers, industry reports, and pandemic-era innovations, the research presented how AI adoption in the COVID-19 era brought about both opportunities and challenges in high-risk environments such as healthcare, public health surveillance, and crisis preparedness.

The findings suggest that while AI-driven analytics can enhance efficiency, responsiveness, and data-informed decision-making across sectors, it also raises significant risks around legal compliance, ethical use, data privacy, equity, and environmental impact. Especially, emerging concerns about algorithmic vagueness, Gen AI hallucinations, and AI deployment in under-resourced regions highlight the urgency for stronger governance frameworks. Most importantly, the unpredictable and disruptive nature of global health crises has introduced the need for an additional dimension which is now introduced as Resilience Responsibility within the SRAI model. This new layer emphasizes the importance of developing AI systems that are not only responsible and sustainable but also robust enough to adapt under conditions of uncertainty and systemic shock [1].

Through redefining the SRAI framework in pandemic-use terms, our study facilitates a broader foundation for responsibly using AI in organizational and wider public service contexts. It encourages stakeholders including HR professionals, public health officials, and AI architects to implement a socially responsible, law-compliant, ethically appropriate, environmentally sustainable, and resilience-driven AI governance framework. Future research should explore methods to apply Resilience Responsibility through real-world testing and stress simulation in both workforce management and health emergency systems. We encourage future research to refine SRAI's layers of responsibilities in response to ongoing technological, regulatory, and societal developments.

References

[1] O. Khamidov and L. Sztandera, "Socially Responsible Artificial Intelligence Empowered People Analytics: A Novel Framework Towards Sustainability," in Proc. Pandemics Analytics, pp. 1-7, 2025.

[2] Y.-L. Chang and J. Ke, 'Socially Responsible Artificial Intelligence Empowered People Analytics: A Novel Framework Towards Sustainability', Sage Journals. Accessed: Jan. 07, 2026. [Online] https://journals.sagepub.com/doi/10.1177/15344843231200930

[3] S. W. Yoon, 'Innovative data analytic methods in human resource development: Recommendations for research design', *Hum. Resour. Dev. Q.*, vol. 29, no. 4, pp. 299–306, 2018, doi: 10.1002/hrdq.21331.

[4] S. W. Yoon, 'Explosion of people analytics, machine learning, and human resource technologies: Implications and applications for research', *Hum. Resour. Dev. Q.*, vol. 32, no. 3, pp. 243–250, 2021, doi: 10.1002/hrdq.21456.

[5] C. Duffy, 'Lawsuit claims discrimination by Workday's hiring tech prevented people over 40 from getting hired | CNN Business', CNN. Accessed: Jun. 18, 2025. [Online]. https://www.cnn.com/2025/05/22/tech/workday-ai-hiring-discrimination-lawsuit

[6] A. B. Carroll, 'The pyramid of corporate social responsibility: Toward the moral management of organizational stakeholders', *Bus. Horiz.*, vol. 34, no. 4, pp. 39–48, Jul. 1991, doi: 10.1016/0007-6813(91)90005-G.

[7] R. Azamfirei, S. R. Kudchadkar, and J. Fackler, 'Large language models and the perils of their hallucinations', *Crit. Care*, vol. 27, no. 1, p. 120, Mar. 2023, doi: 10.1186/s13054-023-04393-x.

[8] M. G. Hanna *et al.*, 'Ethical and Bias Considerations in Artificial Intelligence/Machine Learning', *Mod. Pathol.*, vol. 38, no. 3, p. 100686, 2024, doi: 10.1016/j.modpat.2024.100686.

[9] A. Cordella and F. Gualdi, 'Regulating generative AI: The limits of technology-neutral regulatory frameworks. Insights from Italy's intervention on ChatGPT', *Gov. Inf. Q.*, vol. 41, no. 4, p. 101982, Dec. 2024, doi: 10.1016/j.giq.2024.101982.

[10] M. Rutinel, B. Titone, and R. Rodriguez, *Consumer Protections for Artificial Intelligence*. 2024. Accessed: Jan. 12, 2026. [Online]. https://leg.colorado.gov/bills/sb24-205

[11] J. Obernolte, 'H.R.3460 - 119th Congress (2025-2026): AI Whistleblower Protection Act'. [Online]. https://www.congress.gov/bill/119th-congress/house-bill/3460/text

[12] A. Zewe, 'Explained: Generative AI's environmental impact', MIT News | Massachusetts Institute of Technology. Accessed: Jul. 03, 2025. [Online]. https://news.mit.edu/2025/explained-generative-ai-environmental-impact-0117

[13] G. Zhang, S. Ma, M. Zheng, C. Li, F. Chang, and F. Zhang, 'Impact of digitization and artificial intelligence on carbon emissions considering variable interaction and heterogeneity: An interpretable deep learning modeling framework', *Sustain. Cities Soc.*, vol. 125, p. 106333, May 2025, doi: 10.1016/j.scs.2025.106333.

[14] L. M. Hite and K. S. McDonald, 'Careers after COVID-19: challenges and changes', *Hum. Resour. Dev. Int.*, vol. 23, no. 4, pp. 427–437, 2020, doi: 10.1080/13678868.2020.1779576.

[15] L. K. Khor and C. L. Tan, 'Workforce management in the post-pandemic era: Evidence from multinational companies using grounded theory', *Glob. Bus. Organ.*

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Excell., vol. 42, no. 4, pp. 93–104, 2023, doi: 10.1002/joe.22174.

[16] C. Westfall, 'IBM Replaces Hundreds With AI As HR, L&D Leaders Rethink Roles', Forbes. Accessed: Jun. 19, 2025. [Online]. https://www.forbes.com/sites/chriswestfall/2025/05/27/ibm-replaces-hundreds-with-ai-as-hr-ld-leaders-rethink-roles/

[17] P. Mantello and M.-T. Ho, 'Emotional AI and the future of wellbeing in the post-pandemic workplace', *AI Soc.*, vol. 39, no. 4, pp. 1883–1889, 2023, doi: 10.1007/s00146-023-01639-8.

[18] A. Valtonen, M. Saunila, J. Ukko, L. Treves, and P. Ritala, 'AI and employee wellbeing in the workplace: An empirical study', *J. Bus. Res.*, vol. 199, p. 115584, Oct. 2025, doi: 10.1016/j.jbusres.2025.115584.

[19] T. Alves, 'Generative AI: The Promise and Peril for Scientific Publishing', *Sci. Ed.*, vol. 46, no. 2, pp. 67–68, 2023, doi: 10.36591/SE-D-4602-13.

[20] H. K. Sriram, 'Generative AI and Neural Networks in Human Resource Management: Transforming Payroll, Workforce Insights, and Digital Employee Payments through AI Innovations', *Adv. Consum. Res.*, vol. 2, no. 1, 2025, Accessed: Jan. 06, 2026. [Online]. https://www.academia.edu/127904804/Generative_AI_and_Neural_Networks_in_Human_Resource_Management_Transforming_Payroll_Workforce_Insights_and_Digital_Employee_Payments_through_AI_Innovations

[21] A. Jassy, 'Message from CEO Andy Jassy: Some thoughts on Generative AI'. Accessed: Jun. 20, 2025. [Online]. https://www.aboutamazon.com/news/company-news/amazon-ceo-andy-jassy-on-generative-ai

[22] M. I. Khan, E. Parahyanti, and S. Hussain, 'The Role Generative AI in Human Resource Management: Enhancing Operational Efficiency, Decision-Making, and Addressing Ethical Challenges', *Asian J. Logist. Manag.*, vol. 3, pp. 104–125, 2024, doi: 10.14710/ajlm.2024.24671.

[23] M. M. A. A. M. Sony, M. B. Amin, A. Ashraf, K. M. A. Islam, N. C. Debnath, and G. C. Debnath, 'Bias in AI-driven HRM systems: Investigating discrimination risks embedded in AI recruitment tools and HR analytics', *Soc. Sci. Humanit. Open*, vol. 12, p. 102082, Jan. 2025, doi: 10.1016/j.ssaho.2025.102082.

[24] P. Andrieux, R. D. Johnson, J. Sarabadani, and C. Van Slyke, 'Ethical considerations of generative AI-enabled human resource management', *Organ. Dyn.*, vol. 53, no. 1, p. 101032, 2024, doi: 10.1016/j.orgdyn.2024.101032.

[25] S. Porkodi and T. L. Cedro, 'The Ethical Role of Generative Artificial Intelligence in Modern HR Decision-Making: A Systematic Literature Review', *Eur. J. Bus. Manag. Res.*, vol. 10, no. 1, Art. no. 1, 2025, doi: 10.24018/ejbmr.2025.10.1.2535.

[26] London Daily News, 'Meta's $14.8b AI power play: Is big tech outsmarting antitrust?', London Daily News. Accessed: Jun. 21, 2025. [Online]. https://www.londondaily.news/metas-14-8b-ai-power-play-is-big-tech-outsmarting-antitrust/

[27] A. Mer and A. Srivastava, 'Employee Engagement in the New Normal: Artificial Intelligence as a Buzzword or a Game Changer?', in *The Adoption and Effect of Artificial Intelligence on Human Resources Management, Part A*, P. Tyagi, N. Chilamkurti, S. Grima, K. Sood, and B. Balusamy, Eds, Emerald Publishing Limited, 2023, pp. 15–46. doi: 10.1108/978-1-80382-027-920231002.

[28] L. Manroop, H. Zheng, A. Malik, M. Milner, E. Schulz, and K. Banerji, 'Human resource management in times of crisis: Strategies for a post COVID-19 workplace', *Organ. Dyn.*, vol. 54, no. 1, p. 101060, 2025, doi: 10.1016/j.orgdyn.2024.101060.

[29] A. Mer, 'Artificial Intelligence in Human Resource Management: Recent Trends and Research Agenda', in *Digital Transformation, Strategic Resilience, Cyber Security and Risk Management*, vol. 111B, Emerald Publishing Limited, 2023, pp. 31–56. doi: 10.1108/S1569-375920230001111B003.

[30] Q. Xiao, J. Yan, and G. J. Bamber, 'How does AI-enabled HR analytics influence employee resilience: job crafting as a mediator and HRM system strength as a moderator', *Pers. Rev.*, vol. 54, no. 3, pp. 824–843, Oct. 2023, doi: 10.1108/PR-03-2023-0198.

[31] A. Tursunbayeva, C. Pagliari, S. D. Lauro, and G. Antonelli, 'The ethics of people analytics: risks, opportunities and recommendations', *Pers. Rev.*, vol. 51, no. 3, pp. 900–921, Mar. 2021, doi: 10.1108/PR-12-2019-0680.

[32] S. Bilderback, 'Ethical blind spots in leadership: addressing unconscious bias in post-COVID workforce management', *J. Ethics Entrep. Technol.*, vol. 5, no. 1, pp. 85–102, 2025, doi: 10.1108/JEET-01-2025-0002.

[33] V. Bryce, N. K. McBride, and M. Cunden, 'Post-COVID-19 ethics of people analytics', *J. Inf. Commun. Ethics Soc.*, vol. 20, no. 4, pp. 480–494, 2022, doi: 10.1108/JICES-09-2021-0096.

[34] L. M. Giermindl, F. Strich, O. Christ, U. Leicht-Deobald, and A. Redzepi, 'The dark sides of people analytics: reviewing the perils for organisations and employees', *Eur. J. Inf. Syst.*, vol. 31, no. 3, pp. 410–435, 2021, doi: 10.1080/0960085X.2021.1927213.

[35] S. W. Yoon, S. Han, and C. Chae, 'People Analytics and Human Resource Development – Research Landscape and Future Needs Based on Bibliometrics and Scoping Review', *Hum. Resour. Dev. Rev.*, vol. 23, no. 1, pp. 30–57, 2024, doi: 10.1177/15344843231209362.

[36] R. Robin, 'Post-COVID HR Transformation: How AI is Reshaping Employee Performance and Development', 2025, *Social Science Research Network, Rochester, NY*: 5106031. doi: 10.2139/ssrn.5106031.

[37] C. Lv *et al.*, 'Innovative applications of artificial intelligence during the COVID-19 pandemic', *Infect.*

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Med., vol. 3, no. 1, p. 100095, 2024, doi: 10.1016/j.imj.2024.100095.

[38] M. S. Gawande, N. Zade, P. Kumar, S. Gundewar, I. N. Weerarathna, and P. Verma, 'The role of artificial intelligence in pandemic responses: from epidemiological modeling to vaccine development', *Mol. Biomed.*, vol. 6, no. 1, p. 1, 2025, doi: 10.1186/s43556-024-00238-3.

[39] G. Badiola-Zabala, J. M. Lopez-Guede, J. Estevez, and M. Graña, 'Machine Learning First Response to COVID-19: A Systematic Literature Review of Clinical Decision Assistance Approaches during Pandemic Years from 2020 to 2022', *Electronics*, vol. 13, no. 6, Art. no. 6, 2024, doi: 10.3390/electronics13061005.

[40] S. Consoli, P. Markov, N. Stilianakis, L. Bertolini, A. Gallardo, and M. Ceresa, *Epidemic Information Extraction for Event-Based Surveillance using Large Language Models*. 2024. doi: 10.48550/arXiv.2408.14277.

[41] H. Du *et al.*, 'Advancing real-time infectious disease forecasting using large language models', *Nat. Comput. Sci.*, vol. 5, no. 6, pp. 467–480, Jun. 2025, doi: 10.1038/s43588-025-00798-6.

[42] H. Du *et al.*, 'Advancing Real-time Pandemic Forecasting Using Large Language Models: A COVID-19 Case Study', Apr. 10, 2024, *arXiv*: arXiv:2404.06962. doi: 10.48550/arXiv.2404.06962.

[43] P. A. Giri and M. K. Gupta, 'Transforming Disease Surveillance through Artificial Intelligence', *Indian J. Community Med. Off. Publ. Indian Assoc. Prev. Soc. Med.*, vol. 49, no. 5, pp. 663–664, 2024, doi: 10.4103/ijcm.ijcm_459_24.

[44] M. S. Deiner *et al.*, 'Use of Large Language Models to Classify Epidemiological Characteristics in Synthetic and Real-World Social Media Posts About Conjunctivitis Outbreaks: Infodemiology Study', *J. Med. Internet Res.*, vol. 27, p. e65226, Jul. 2025, doi: 10.2196/65226.

[45] M. Rama, G. Santin, G. Cencetti, M. Tizzoni, and B. Lepri, 'Forecasting Seasonal Influenza Epidemics with Physics-Informed Neural Networks', Jun. 04, 2025, *arXiv*: arXiv:2506.03897. doi: 10.48550/arXiv.2506.03897.

[46] S. Jiao, Y. Wang, X. Ye, L. Nagahara, and T. Sakurai, 'Spatio-temporal epidemic forecasting using mobility data with LSTM networks and attention mechanism', *Sci. Rep.*, vol. 15, no. 1, p. 9603, Mar. 2025, doi: 10.1038/s41598-025-94089-9.

[47] P. Agarwal, S. Swami, and S. K. Malhotra, 'Artificial Intelligence Adoption in the Post COVID-19 New-Normal and Role of Smart Technologies in Transforming Business: a Review', *J. Sci. Technol. Policy Manag.*, vol. 15, no. 3, pp. 506–529, 2022, doi: 10.1108/JSTPM-08-2021-0122.

[48] M. Gheisari *et al.*, 'Mobile Apps for COVID-19 Detection and Diagnosis for Future Pandemic Control: Multidimensional Systematic Review', *JMIR MHealth UHealth*, vol. 12, no. 1, p. e44406, 2024, doi: 10.2196/44406.

[49] M. McKee, R. Rosenbacke, and D. Stuckler, 'The power of artificial intelligence for managing pandemics: A primer for public health professionals', *Int. J. Health Plann. Manage.*, vol. 40, no. 1, pp. 257–270, 2025, doi: 10.1002/hpm.3864.

[50] S. Mohanadas, 'Next-Generation Predictive Analytics for Global Disease Outbreaks: Bridging Innovation, Ethics, and Impact', *Eur. J. Comput. Sci. Inf. Technol.*, vol. 13, no. 4, Apr. 2025. Accessed: Jan. 06, 2026. [Online]. https://eajournals.org/ejcsit/vol13-issue-4-2025/next-generation-predictive-analytics-for-global-disease-outbreaks-bridging-innovation-ethics-and-impact/

[51] S. K. Muhammad, T. A. Ansari, B. Shabbir, and T. A. A. B. S. W. Almagharbeh, 'The Role Of Artificial Intelligence In Public Health Surveillance: A Post-Pandemic Perspective', *ResearchGate*, 2025, doi: 10.71000/pa7ab080.

[52] The Express Tribune, 'OpenAI's o1 model tried to copy itself during shutdown tests', The Express Tribune. Accessed: Jul. 16, 2025. [Online]. https://tribune.com.pk/story/2554708/openais-o1-model-tried-to-copy-itself-during-shutdown-tests

[53] M. Yang, 'Elon Musk's AI firm apologizes after chatbot Grok praises Hitler', *The Guardian*, 2025. Accessed: Jul. 16, 2025. [Online]. https://www.theguardian.com/us-news/2025/jul/12/elon-musk-grok-antisemitic

[54] Deloitte Global, 'AI for infrastructure resilience'. Accessed: Jul. 16, 2025. [Online]. https://www.deloitte.com/global/en/issues/climate/ai-for-infrastructure-resilience.html

[55] R. Ghatak, People Analytics: Data to Decisions. in Management for Professionals. Singapore: Springer Nature, 2022. doi: 10.1007/978-981-19-3873-3.

2026, © Copyright by authors, Published under agreement with IARIA - www.iaria.org