Toward a working theory for Human-Data Interaction in the Built Environment

Li, Jennifer; Calis, Gulben; Calvetti, Diego; Demian, Peter; Kassem, Mohamad; Msawil, Mahir; Shayboun, May; Soman, Ranjith

DOI
10.35490/EC3.2024.282

Publication date
2024

Document Version
Final published version

Published in
Proceedings of the 2024 European Conference on Computing in Construction

TOWARD A WORKING THEORY FOR HUMAN-DATA INTERACTION IN THE BUILT ENVIRONMENT

Jennifer Li [1], Gulben Calis[2], Diego Calvetti[3], Peter Demian[4], Mohamad Kassem[5], Mahir Msawil[6], May Shayboun[7] and Ranjith Soman[8]
[1]Northumbria University, Amsterdam, The Netherlands; [2]Ege University, Bornova, Turkey; [3]CONSTRUCT/Gequaltec, Porto University, Porto, Portugal; [4]Loughborough University, Loughborough, United Kingdom; [5]Newcastle University, Newcastle, United Kingdom; [6]Northumbria University, Newcastle, United Kingdom; [7]Halmstad University, Halmstad, Sweden; [8]TU Delft, Delft, The Netherlands.

Abstract

Human-Data Interaction is central to the dynamic nature of the construction industry and its growing reliance on substantial data sets. In this paper, we evaluate literature on the topic of Human-Data Interaction across several technologies and concepts within the built environment with the aim of identifying underlying aspects that can lead to an underpinning theory of Human-Data Interaction in the built environment to support advancement of the research in the field. Those aspects were identified as trust, game theory, empowerment of humans, human control, safety, accessibility, enhancing understanding, and the three pillars of Human-Data Interaction of agency, legibility, and negotiability.

Introduction

The evolution from specialised to ubiquitous and interconnected computing has instigated a profound transformation within the construction sector impacting every facet of the design, construction and operation of buildings and infrastructure. During the design phase, the adoption of tools such as Building Information Modelling (BIM) has revolutionised approaches by enabling collaborative planning and precise 3D modelling in real time. This technological advancement continues into the construction phase with the use of devices like smartphones and tablets along with devices and sensors that monitor progress and site conditions. These innovations contribute to decision making and efficient resource management. In the operations phase, integrating building technologies and data analytics plays a crucial role in optimising maintenance, energy usage and overall facility management. This promotes sustainability and cost effectiveness.

The widespread use of various devices and networks, along with their technological interactions, results in the creation of extensive data trails with considerable consequences. Data are under constant revision and extension or transformation and do not only concern the individual who provided the data or about whom the data is, but also other stakeholders that might have different interpretations of the data (Hornung *et al.*, 2015).

Considering the dynamic nature of the construction sector and its growing reliance on substantial data sets, it becomes crucial to delve into the study of Human-Data Interaction (HDI), particularly in the construction informatics domain. The aim of this paper is to make progress towards establishing a theory of HDI to underpin subsequent research into construction informatics.

HDI described as interdisciplinary encompassing "*data visualization, user interface design, and interaction as well as psychology, behavioral science, and human cognition*" (Widjojo *et al.*, 2017, p. 3). In this paper, we uncover some of these aspects through investigations into HDI across the built environment (BE). In the next section, HDI is positioned in the context of its evolution from HCI and offer an understanding of the role of data in the construction sector. The methodology describes how the study was conducted and is followed by the findings of a literature review on HDI across several technologies and concepts in the BE. Analysis of the identifies potential elements that could contribute to development of a theory of HDI for the BE leading to a roadmap to support development of such a theory. The paper end with a discussion and conclusions.

From Human-Computer Interaction (HCI) to Human-Data Interaction (HDI)

Human-Computer Interaction (HCI) has been a cornerstone in our interaction with technology, focusing on the design and use of computer technology alongside the interactions between humans and computers (Dix, 2017). HCI's evolution has shifted from viewing computers merely as devices to a complex understanding of these interactions (Grudin, 1990), encompassing psychology (Card, 2018), hardware, software, interface, and even deeper organisational aspects. This expanded view transcends the simple operator-hardware relationship, leading to a more in-depth exploration of user-computer system interrelationships, particularly as systems become more integrated within organisations (Bowers and Rodden, 1993).

An emerging dynamic ecosystem, marked by both collaborative and competitive interactions, revolves around data creation and usage by individuals (Brown, 2013). This dynamic ecosystem along with advanced machine learning (ML) and artificial intelligence (AI) has initiated a transition towards HDI (Crabtree and Mortier, 2015). Unlike HCI, HDI focuses on the interaction between humans and data, moving beyond traditional interfaces. HDI's realm extends to the subtle and often passive engagement of individuals with complex infrastructures that are typically misunderstood or ignored (Haddadi *et al.*, 2013). Although the concept of *data* is nebulous (as will be discussed below), it is arguably less important than the elements of *human interaction*. HDI views data as a boundary object, open to diverse interpretations and relevant to a broad spectrum of stakeholders, including those it concerns, collectors, legal custodians, and users (Mortier *et al.*, 2015). The challenge of HDI lies not just in usable ("user-friendly") interfaces but in making complex data comprehensible, actionable, and ethically managed. Its core themes include *Legibility, Agency*, and *Negotiability. Legibility* aims to make data and algorithms transparent and understandable, balancing intellectual property protection. *Agency* empowers individuals to manage their data, including opting in or out, without necessitating constant engagement but offering the choice for those interested or concerned. *Negotiability* addresses the evolving perceptions and relationships around data, societal responses, social norms, legal frameworks, and the ambiguity in data's subject and meaning (Mortier *et al.*, 2015).

As the era of big data in the construction sector progresses, HDI's themes become more pertinent, signifying a shift from mere interaction with technologies (e.g., BIM, IoT, Digital Twins, etc.) to a deeper engagement with data and its implications.

What is data and where is it produced in the construction sector?

According to the Mirriam-Webster (2023) dictionary, data is "*factual information (such as measurements or statistics) used as a basis for reasoning, discussion, or calculation*", "*information in digital form that can be transmitted or processed",* and *"information output by a sensing device or organ that includes both useful and irrelevant or redundant information and must be processed to be meaningful*".

The definition in data science might be extended beyond the use of data, as a data application acquires its value from data and produces more data as a result (Loukides, 2011). The meaning is that data do not only function as input in data driven applications or services but also outputs data that need to be processed or transmitted. On the other hand, big data can be defined as an overall term for a "*complex and substantial collection of data, which needs advanced engineering strategies and analytics systems to process, store and manage*" (Li *et al.*, 2023). Big data includes a data transformation flow and is produced in the construction sector from sites and procedures within the entire construction process, consisting of modelling, designing, planning, scheduling, and management. Big data in construction, therefore, has different sources such as BIM as a main source of data supporting planning, design, building and operation (Li *et al.*, 2023), in addition to sensors and embedded devices, operational and maintenance data including text, audio, video and more (Bilal *et al.*, 2016; Li *et al.*, 2023). In this context, there is a role of project participants in data input, acquisition, modification, and integration of information leading to information exchange as well as integration (Li *et al.*, 2023). Furthermore, construction digital twin is distinguished from BIM because of its connection to the physical twin and updates when the physical one changes throughout the process (Sacks *et al.*, 2020). The digital twin would allow for 'data centric' construction management with a flow of information between the physical and virtual models in a cycle of Plan-Do-Check-Act where 'check' marks the significant difference with current construction control if the data can be capably interpreted and automatically produce accurate and comprehensive information (Sacks *et al.*, 2020). So, data function in a process of input, processing, output in a flow of various types of information in constant interaction with the human at any point of any process.

Methodology

This inductive study conducted a literature review of HDI across several themes related to the BE and the focus of the HDI Committee of the European Council on Computing in Construction (EC3). The framework under which the literature review was conducted stemmed from the mission of the HDI Committee (EC3, n.d.) across the three research perspectives in Figure 1. This is devised to avoid the risk of developing an unordered and incomplete list of research directions omitting important areas.

HDI Research Perspective	Description
i) Understanding and evaluation (end user-oriented)	• Users and their behaviours. • Individual and societal implications of emerging interactions. • Novel/emerging uses of ICT and AI-related applications.
ii) Development (technology-oriented)	• Platforms, architectures or component technologies. • Interactions between human actors and the data systems within projects and the built environment. • Testing methods for existing technologies and design methodologies.
iii) Foundation (theory building-oriented, supported by or supporting i) and ii))	• Theories, methodologies and methods involved in emerging HDI. • Paradigms and frameworks for analysis, design and application of HDI. • Ethics and Regulations.

Figure 1: Factors of the HDI Committee Mission

Literature specific to the concept of HDI and the six themes of blockchain; machine learning (ML) and deep learning (DL); robotics for industrialised construction; sensored construction sites; smart buildings; and virtual reality (VR) and augmented reality (AR) were reviewed. These themes follow the topics of the HDI Committee's seminal white paper (Kassem and Kifokeris, forthcoming) as key considerations for the BE. We acknowledge these themes are not exhaustive to the concept of HDI in the BE, however, they provide a springboard from which to commence discussions on development a theory of HDI *for* the BE. The purpose of the review was to identify similarities across these themes that could begin to form a working theory for HDI in the BE and specifically computer informatics. Inclusion criteria included papers considering the factors above with a particular focus on users, their behaviour, their interaction with data, and the testing and/or development of technology that integrates both user data and humans. The findings of the study were consolidated to propose key elements for consideration when developing a theory of HDI in the BE and are supported by a roadmap to support achieving that theory.

Literature review

This section presents the literature reviewed across the six themes and is followed by Figure 2 showing the HDI research perspectives across each theme. The themes are ordered alphabetically with no one theme having more importance over another. Literature on each of these

References	i) Understanding and evaluation (end user-oriented)	ii) Development (technology-oriented)	iii) Foundation (theory building-oriented, supported by or supporting 'i' and 'ii')
Blockchain			
Becherer et al. (2020)	Human-centred design to generate trust between the system and humans. Focuses on operational risk management.	Public-key infrastructure (PKI) that gives control to an individual over their data is centrally managed with blockchain.	Humans put trust into the blockchain-based system.
Lockwood (2021a; 2021b)	Increasing agency of human actors over their personal data. Focuses on privacy and its importance in society, with a particular focus on surveillance.	SSI and Web3 gives agency and control to the user over personal data. Blockchain can facilitate persistence, portability, consent and interoperability of data.	The three pillars of HDI: agency, legibility, negotiability.
Calvetti et al. (2021)	Facilitation of contractual arrangements for sensored construction sites.	Considers relationships between workers, sensors, systems and methodologies, performance, human factors, and on-site conditions where individual data are collected.	Increased transparency for users of the system.
Hunhevicz et al. (2021); Wang et al. (2022)	Changes in ideas of ownership where it is shifted from the human (or organisation) to the asset.	no1s1 – a prototype for self-owning assets facilitated by a decentralised autonomous organisation (DAO).	Removal of profit-seeking intermediaries returning some agency to the individual.
Li et al. (2019)	A framework emphasising the need for societal input into development and deployment of blockchain-based applications in construction.	Blockchain and smart contracts as integrating technologies for construction sector applications alongside BIM, IOT, GIS etc.	A sociotechnical implementation framework, considering the dimensions of technology, process, society, and policy.
Machine Learning (ML) & Deep Learning (DL)			
Urquhart et al. (2019)	Framing regulatory, interactional and technical challenges of Adaptive Architecture.	Adaptive buildings and integrated IoT devices need to be ensured by legal rights informed by how occupants and their data interact with the buildings.	Physical and information security; establishing responsibility; understanding flows, collection, use and control of personal data; sensitive personal data and monitoring routine activities.
Schia et al. (2019)	AI trust is conditioned by knowing the AI is based on the right data and how and why the output was made.	A framework including technology, process and culture to determine the success of implementation.	The control of humans on the input creates a sense of trust. The more automated the process, the less humans have control.
Mosqueira-Rey et al. (2023)	Humans act as teachers of an ML system depending on the level of involvement of the human in the learning process.	Humans can also be at the end of the process trying to understand how and why ML decisions were made.	HITL, explainability, usable and useful AI.
Hanafy (2023)	User iterates multiple queries for optimal output whether by choosing best version or changing terms.	General purpose generative text-to-image AI used in architecture interior and exterior design.	Augmenting creativity and visualisation in architecture. Iterative trial and error and produce various choices.
Robotics for Industrialised Construction			
Sun et al. (2023)	Investigating and evaluating the impact of robots on psychological and physical safety of human agents.	Robotics technology requires a complicated integration of multidisciplinary technologies and their safety needs to be considered in regulations.	Physical risks imposed on human agents. Visual and cognitive distraction of involved human agents. Cognitive load and negative emotional state.
Rodrigues et al. (2023)	Classifying human traits, identifying human roles, understanding the main factors that impact HRI. Proposing safety mechanisms for HRI that involve the human side.	Classifying robot traits, identifying robot types, and categorising level of robot autonomy (LoRA). Proposing safety mechanisms for HRI that involve the robotic side.	Human-related factors in respect to safety: (i) mental workload, (ii) situation awareness, (iii) trust in automation, and (iv) (physical) ergonomics.
Liu et al. (2021)	Proposing a human-centred collaborative framework that enables robots to interpret human agents' cognitive loads to improve HRC.	Enabling robotic agents to capture and predict human psychological behaviour through developing machine learning models.	Acknowledging the need for reducing the cognitive and mental load imposed on human agents involved in working with robotic agents.
Fu et al. (2024)	Understanding the impacts of HITL collaboration in off-site construction that employ robots.	Alluding to the need for training various robot types in actual work environments rather than in experimental environments.	Recognising the need for respecting the value of human skills and strengths to empower human agents over robotic agents.
Zhang et al. (2023)	Understanding risks imposed on human agents in HRC in on-site construction (e.g., physical and psychology-related risks).	Understanding risks imposed on robotic agents in HRC in on-site construction (e.g., program failure and maintenance-related risks).	focus is needed on safety and management issues of HRC in on-site construction with an implicit emphasis on empowering the human side.
Sensored Construction Sites (SCS)			
Tang et al. (2020)	Evaluates worker PPE compliance using computer vision.	Develops a new vision-based system for PPE compliance.	N/A
Park et al. (2016)	Evaluates detection of unsafe conditions and collects and analyses the trajectories of workers with respect to potential safety hazards.	Develops a Bluetooth and BIM system for safety monitoring.	N/A
Kim et al. (2021)	Evaluates compliance levels and informs site managers about non-compliance.	Develops an autonomous detection method based on sensor technology for hard hat non-use.	N/A
Jebelli et al. (2018)	Evaluates early detection of workers' stress feasibility.	Develops a wearable EEG-based field stress recognition procedure.	N/A
Aryal et al. (2017)	Evaluates workers' real-time fatigue.	Develops a multi-sensor approach using heart rate monitor, infrared temperature sensors and an EEG sensor.	N/A
Shanti et al. (2022)	Evaluates workers' safety compliance at-height activities.	Develops a novel system integrating DL and drones to monitor workers in real-time Personal Fall Arrest System components compliance such as safety harness, lifeline, and helmet.	N/A
Smart Buildings			
Konstantakopoulos et al. (2019)	Incorporates HITL modelling by creating an interface to allow building managers to interact with occupants and potentially incentivize energy efficient behaviour.	Develops a gamification application to motivate humans toward energy efficient behaviour.	Introduces a smart building social game concept based on a friendly competition between occupants. Game theory concepts are used to learn models of players' decision-making in residential buildings.
Kim et al. (2022; 2023)	Incorporates user's baseline thermostat usage and allows users to set up thermostat schedules.	Develops a cloud-based eco-feedback and gaming platform (MySmartE app) that aims to promote energy conserving thermostat adjustment behaviours in multi-unit residential buildings.	Introduces a mathematical framework through a new methodology for personalized eco-feedback design integrated with a collaborative social game to assist residents in enhancing thermostat use while promoting community-level energy-savings.
Jenkins et al. (2019)	Enables occupants to take an active role in monitoring and managing plug loads via a mobile application.	Develops web and mobile applications for building administrators and users for plug load management.	Incorporates occupant engagement, gamification, automation, and machine learning for enabling occupants to take an active role in monitoring and managing their plug loads.
Virtual Reality (VR) & Augmented Reality (AR)			
Widjojo et al. (2017)	Human perception, cognition and decision-making are deeply subjective. Systems should aim to enhance humans' understanding of big datasets through enhanced visualisation.	Systems must be scalable in terms of quantity of data and dimensionality/complexity. The focus of a technological system must be in simple accessibility of data to the human.	Systems for Visual Analytics need to exploit the respective strengths of humans (perception and reasoning) and computers (computation and visualisation) and consider meaningful ways for humans to understand data through visualisation.
Alhakamy et al. (2021)	Humans can potentially interact with data using their whole bodies.	Tested two styles of interacting with data: using gestures and body movements in a large (semi-immersive) environments versus a joypad.	The mode of interacting with data affects the easing and conclusions drawn from the data.
Schiavi et al. (2022)	N/A	N/A	It is important to develop data flows between the systems used in construction (BIM, CDE, Digital Twins) and VR/AR systems.
Grübel et al. (2022)	N/A	Situated analytics allows the anchoring of data in its spatial context, which is particularly important in built environment applications of AR/VR.	N/A

Figure 2: Literature considered against the framework of HDI research perspectives

themes with a specific focus on HDI is limited within and outwith the BE. That which is present has been carefully selected and reviewed to highlight the components that could be relevant to overarching HDI theory in the BE.

Blockchain

Blockchain was established as a technology to challenge capitalist power structures that have existed for many years (Ekblaw *et al.*, 2016). Research has demonstrated its potential to challenge many economic structures, business models, operations and much more. For the BE, and particularly the construction sector, the focus of research generally is heavily rooted in data and how they can be leveraged to make efficiencies and increase productivity whilst still maintaining an advantage over competitors (Li and Kassem, 2021). Such capitalist systems have resulted in challenges including predatory business models (e.g., underbidding for projects, using projects funds for cashflow), the justification of competitive advantage to limit information sharing, and a general lack of trust between contracting parties (Li and Kassem, 2021). Combined with new insights gained from [big] data, a new level of exploitation of individual actors from such actions drove enactment of the General Data Protection Regulation (GDPR) and highlighted the need to rethink companies' collection and use of data. Blockchain is presented as a way to support this rethinking, particularly from a Web3 perspective where focus is on giving *agency* back to the individual.

Lockwood's (2021a; 2021b) research focuses on self-sovereign identify (SSI) as an inevitable aspect of Web3. Privacy and its importance in society is central to the discussion and understanding that centralisation removes agency of personal data. The solution focuses on *agency*, *legibility* and *negotiability*, and the principles for an SSI technology: existence of the user; control and access by the user; transparency within the system; persistence of identities; portability of information and services about the data; interoperability to ensure wide usability; consent from users to use their identity; minimalization of disclosure of claims; and protection of users' rights. From a trust perspective, Becherer *et al.* (2020) present an exploration of human data engineering (HDE) for operational risk management with blockchain. Human-centred design of a blockchain-based system generates trust between it and humans. Public key infrastructure (PKI) is used to manage data centrally. While this goes against the decentralised nature of blockchain, its key HDI considerations put trust and human-centricity at the centre of the big data-designed system.

Specific to construction, Calvetti *et al.* (2021) address HDI for sensored construction sites (SCS) proposing the use of blockchain to facilitate data in contractual arrangements by investigating the perceptions of HDI and SCS. A use case enhances the relationships between workers, sensing technologies, performance, systems and methodologies, human factors, and on-site conditions where data are collected about the individual. Standards of transparency between workers and companies of data collection is raised as a result of implementing their proposed HDI information process. Li *et al.* (2019) proposed a socio-technical framework encompassing dimensions of technology, process, policy and society that gives consideration to the users of the system and the interactions they have with the technology to achieve the desired outputs and benefits. Finally, Hunhevicz *et al.* (2021) and Wang *et al.* (2022) present an alternative to the current model of ownership of assets shifting it away from the human (or organisation) to the asset. *no1s1* (no one's one) is a prototype for self-owning assets facilitated by a decentralised autonomous organisation (DAO) that increase cooperation and coordination between transacting parties whilst removing profit-seeking intermediaries returning some agency to the individual.

Machine Learning (ML) and Deep Learning (DL)

Much of AI subfields function by studying data and making decisions such as ML, computer vision, optimisation, and knowledge-based systems (Abioye *et al.*, 2021). AI has been gradually advancing in the construction sector with different aspects including health and safety, BIM, supply chain management and other (Abioye *et al.*, 2021). The subfields of AI can vary but they share one characteristic of being dependent on input data and result in output data whether this output is a prediction, a recommendation, or an automatic procedure (Allen, 2020). Human-in-the-loop ML (HITL-ML) is a concept situating the human in the process of ML in terms of the human's role in developing an AI system which is determined by the level of control humans and machines have in the learning process, in explaining the outcome to humans, and in the usability and usefulness of AI systems (Mosqueira-Rey *et al.* 2023). An HDI theory should explicit the relationship of its principles between the human and the data at any point where interaction takes place.

Urquhart *et al.* (2019) focused on how inhabitants interact with their buildings that adapt with the environment and its inhabitants through a series of IoT devices integrated with ML or DL technologies. The analysis discusses the challenges of adaptive architecture in dimensions of physical and information security, establishing responsibility, rights over personal data, sensitivity of visible emotions and bodies, and continuous monitoring. From a user perspective, Schia *et al.* (2019) presented an AI scheduling system at an early stage of implementation and conducted interviews about the system and compared it to other well-implemented digital tools. Interviewees with low AI knowledge working in construction projects indicated that trust in AI is conditioned by knowing the AI is based on the right dataset of BIM, and visualisation of why and how the output was made will help understanding of causality and trust. Moreover, AI technology and other digitalised applications are successfully used when they fulfil maturity in terms of technology, process, and culture. Hanafy (2023) investigated the use of general-purpose technology text-to-image generative AI images was used for architecture. The interaction of the user with the AI system showed that there is a behaviour of refinement as the user did not opt to create the perfect query from the start but built the best one by choosing the best version of adding extra terms. The study found the most suitable aspect of such a tool to be in the early design phase. These tools can be seen as augmenting creativity and visualisation within architecture. In interior design, the tool has potential to allow the user flexibility in design to explore different choices which indicates that the users interact with the output in a form of judgement and learning.

Robotics for Industrialised Construction

In the field of robotics research, Bock (2015) and Sun *et al.* (2023) predicted that the future of construction projects will rely on both robotic and human agents working together to execute construction-related activities. Since the functional operation of robots pivots on embedded sets of data within them (i.e., data are their lifeblood), human-robot interaction (HRI) can be considered a specific instance of HDI. Despite the plethora of research efforts on robotics in construction, only a limited number of studies could be analysed with the goal of distilling a set of specific themes that underlie HRI focusing on the human (Fu *et al.*, 2024).

The psychological and physical safety of human agents in HRI has been the focus of three studies. In their conceptual study, Sun *et al.* (2023) reveal HRI could result in negative psychological impacts on involved human agents including (i) anxiety and acute stress, (ii) cognitive load, (iii) negative emotional state, and (iv) visual and cognitive distraction. These impacts tend to cause increased likelihood of physical accidents, thereby impacting the physical safety of human agents. In the same vein, Rodrigues *et al.* (2023) underline the importance of understanding HRI to protect health and safety of human agents as well as to create an enhanced task environment. Resonating with Sun *et al.* (2023), the authors identify human-related factors that can impact the effectiveness of HRI: (i) mental workload, (ii) situation awareness, (iii) trust in automation, and (iv) (physical) ergonomics. The materialisation of any factor highlighted by Sun *et al.* (2023) and Rodrigues *et al.* (2023) may undermine the perceived benefits of employing robots on construction sites. To this effect, Liu *et al.* (2021) make an attempt at reducing the cognitive and mental load imposed on human agents involved in HRI. They proposed a solution that enables robots to recognize psychological signals of human brains with the goal of adjusting their robotic performance. Such an adjustment is argued to enhance the physical and psychological safety of the involved human agents, thereby improving the efficiency of human-robot collaboration (HRC).

Scholarly suggestions to empower human agents over robotic agents are evident in two studies. Within the context of HRC for modular construction manufacturing (MCM), Fu *et al.*, (2024) place emphasis on the impact of 'human-in-the-loop collaboration' concluding that involved human agents can take on a range of collaborative roles and interact with robotic agents at varying levels. These roles and levels depend on a combination of factors: requirement of the MCM tasks, capacity of robotic agents, and preferences of the human agents. The latter alludes to the value of humans in HRC. Zhang *et al.*, (2023) illuminate this in a recent review on HRC for on-site construction where human agents are tasked with working with robots. They suggest designing human-led systems for HRI to empower humans to control and manipulate robotic agents.

These studies draw our attention to the importance of empowering the human side of the HRI as a proactive measure to enhance interaction efficiency. It can be inferred that the optimization of HRI as a specific instance of HDI relies on human-centric system designs that need to take account of four induced factors: psychological health, physical safety, explicit human value, and empowerment of human agents. These factors can be translated into the three HDI principles as: *legibility*: empowerment of human agents; *agency*: human value and empowerment of human agents; *negotiability*: psychological health and physical safety.

Sensored Construction Sites (SCS)

Information communication technologies (ICT) such as wireless sensors networks, wearables, CCTV, robotics and mobile BIM technologies are increasingly adopted on construction sites (Rossi *et al.*, 2019). Construction sites are challenging environments due to their inherently dynamic, complex, and dangerous nature involving complex interactions between various parties, each with human actors, equipment, and products.

HDI-relevant studies exist across several areas such as safety (e.g., monitoring, compliance, inspection), worker ergonomics (e.g., stress, fatigue), productivity measurement/monitoring, and activity monitoring/ recognition. Common themes include the development and evaluation of technologies like computer vision, sensor technology, and machine learning for safety and compliance monitoring on construction sites. There is a notable gap in addressing human-centred aspects. Many studies do not document participant feedback, assumptions about data ownership, transparency, or algorithm bias. Where participant involvement is mentioned, it often lacks depth in their perspectives or considerations about data matters. Another notable gap is the lack of theories produced or used to support the studies within this domain. This highlights a significant gap in the current research, emphasising the need for more human-centred approaches than human interactions (e.g., consequences/implications/feelings from or when using the developed systems, user acceptance), and data matters (e.g., ethical use, bias, consent, confidentiality, compliance with regulations, etc.).

While the list of papers reviewed is not exhaustive, the trend of insufficient attention to human-centred aspects in safety-related research is indicative of a broader issue in the field. Indeed, safety is a domain where human-centricity is expected to be prominent, and the lack of depth in addressing HDI aspects in these studies suggests a likely widespread gap in the field at large. This trend likely reflects a general oversight in the digitalisation of construction sites, underscoring the need for more comprehensive research that considers HDI, particularly in areas as critical as safety.

Developments in this field will often collect data actively or passively, including workforce location, movements, gestures, physiological levels, and physical efforts (Joshua and Varghese, 2014; Aryal *et al.*, 2017; Park *et al.*, 2017; Jebelli *et al.*, 2018; Tang *et al.*, 2020; Kim *et al.*, 2021; Shanti *et al.*, 2022). However, distinguishing between task analysis and the analysis of the individual performing the task is a delicate matter. It is challenging to determine which data pertain to the project tasks and which data truly reflect the physical and motor abilities of the workers themselves (Calvetti *et al.*, 2021). This complexity underscores the fine line between performance metrics and personal data.

Smart Buildings

As technology continues to advance, a growing interest is observed for integrating cutting-edge technologies and interconnected systems into buildings to enhance their efficiency, functionality, and sustainability. These *smart buildings* leverage various sensors, devices, and automation to create intelligent and responsive spaces that optimize resources, improve occupant comfort, and reduce environmental impact. Subsequently, smart connected devices and advancements in sensing, actuation, and communication bring new modes of interaction within the BE across different contexts and scales. The most common robust examples that illustrate this interaction in smart buildings are interactive systems, which require IoT devices to enable occupants to interact with the building environment actively. Among these systems, automatic climate controllers autonomously regulate indoor air temperature, relative humidity, and air quality in a designated BE (Favero *et al.*, 2022; Lee *et al.*, 2019). Lighting control systems fine-tune the luminous environment and ensure the desired illumination level (Rossi *et al.*, 2015; Cho *et al.* 2020) whereas automatic window systems respond dynamically by automatically opening and closing windows based on fluctuations in indoor air temperature, relative humidity, CO_2 levels, VOC concentrations, as well as PM2.5 and PM10 particles (Cheng *et al.*, 2016; Kim *et al.*, 2019). Most of these systems aim at improving occupant satisfaction, perceived service quality and indoor air quality in buildings.

Recent advancements in this domain evolved around user experience and user-centric features, such as intuitive or guided wayfinding systems, interactive kiosks, and personalised environmental controls. For example, Target retail stores in the United States installed LED lighting systems integrated with Visible Light Communication capabilities. Through the coupling of these LEDs with Visible Light Positioning and the incorporation of image sensors on smartphones, in-store shoppers can navigate in the stores via maps and find the locations of specific products via the directions sent by the retailer's app (Halper, 2019). The EDGE building in Amsterdam exemplifies advanced utilisation of human-building interaction (HBI) data to enhance the user experience via a seamless integration of individual preferences and real-time environmental adaptability. For example, occupants at EDGE find themselves effortlessly connected through an integrated system, with the application aligning with their daily agendas and informing the building of their arrival. The building's advanced system promptly detects incoming vehicles, guiding occupants to designated parking areas. Users have access to a diverse range of work environments and the system, finely tuned to individual preferences, adjusts lighting and temperature within these environments to ensure maximum comfort and efficiency (Randall, 2015).

The current state of smart buildings shows that incorporating human knowledge, behaviour, needs, and preferences into the operational phase is crucial for achieving optimal performance in buildings. Even in automated buildings where users lack direct control over systems, integrating the human dimension through feedback mechanisms can enhance perceived control, leading to higher satisfaction levels. Therefore, HBI emerges as a viable alternative, introducing consensus-based decision-making for building operations to potentially minimize conflicts. It is noteworthy, however, that despite endeavours to create 'smart buildings', this objective remains unattainable without 'smart users'. Users need to be active participants in the building ecosystem rather than passive recipients, engaging with the data available to them in the most effective manner.

Virtual Reality (VR) and Augmented Reality (AR)

With its focus on 3D space on a large scale, the BE is a natural application domain for VR and AR, arguably comparable to entertainment and tourism. Literature on VR/AR within and outwith the BE disciplines presents applications to visualise the 3D BE or abstract data which is somehow related to the design/construction/ operation of buildings but with no natural 3D mapping. Data visualisation is often broadly classified into *information visualisation* and *scientific visualisation*; the former focuses on abstract data, whereas the latter deals with datasets mapped to space and time.

In their position paper, Widjojo *et al.* (2017) define HDI in slightly narrower terms than elsewhere in the literature, drawing an analogy to HCI, but focusing specifically on data. They coin the term *Visual Analytics (VA)* as the integration of human efforts and computation in the analysis of data. They argue that VR technology is particularly suited to VA as it overcomes the challenge of limited display real estate. Coupled with innovative interaction design, VR can facilitate the visual analysis of complex data, such as high-dimensional and multivariate data. The authors highlight the importance of understanding human cognition when exploiting it for VA. They go on to cite various interaction patterns that can inform the design of VR systems for HDI, such as selection, manipulation and viewpoint control. In the case of 3D spatial data, clearly of interest in the BE, Widjojo *et al.* highlight three components of the immersion enabled by VR that are particularly beneficial: head tracking, field of regard, and stereoscopic rendering. These help enable a sense of presence and spatial awareness. The paper culminates with a research agenda framed in an abstract way: research on the required understanding of a specific data set, and on the VR visualisation/interaction functionality that would enable this understanding.

Outside the BE, Alhakamy *et al.*, (2021) study how users perceive correlation and causation in numerical data. They compare two styles of interacting with the data: using gestures and body movements in a large (semi-immersive) environment versus a joypad, similar to the controller used in games. It is reported that participants tended to agree more with statements that portray correlation and causation in data after using the semi-immersive system. The conclusions are presented in the light of theories of embodied learning/cognition.

Two applications of VR/AR in the BE seem implicitly to address the interaction of humans with data, but without explicitly using the term HDI: data for construction safety (Schiavi *et al.*, 2022) and data fused from digital and physical twins to provide spatial context in a smart city (Grübel *et al.*, 2022). The general themes emerging from these are that both VR and AR offer benefits for BE applications in particular because of the 3D orientation of most digital content; VR is used preferentially for office uses whereas AR is more often used with tablets and mobile devices 'in the field'; data flow between the systems used in design, construction and operation of the BE and VR/AR systems remains an issue.

Potential elements for a theory of HDI in the built environment (BE)

Following a review of the literature, the papers were mapped to the framework in Figure 1, as shown in Figure

2. *Understanding and evaluation* focused on what the studies are trying to achieve from the technologies and their potential impact on the end-user in the context of HDI. *Development* focused on any development of the technologies to achieve improvements in HDI for the end-user. *Foundation* considered any proposed frameworks or theories specific to HDI that could lead to an underpinning theory of HDI for the BE that may be applicable across all technologies and concepts.

Trust appears to be a central facet of a system when considering interaction of humans and data. This is applicable in terms of human-centred systems as well as humans knowing the correct data are being used when data-driven decisions are made *about* or *for* them without human input, especially when related to wellbeing. ***Game theory*** is considered as a way to encourage engagement with data. This relates to ***empowerment of humans*** when collaborating with technologies (e.g., robots, surveillance systems) and moving toward automation that (partially) removes ***human control***. Humans are often wary of new technologies and can result in an unwillingness to collaborate. This links to potential issues of ***safety*** of the human from both physical and psychosocial perspectives when humans work with hardware (e.g., robots, wearables) and/or software where concerns may be less detectable. ***Accessibility*** considers how humans interact with data and affects how they make sense of correlation and causation of those data. ***Enhancing understanding*** of data via visualisation can increase accessibility and willingness to engage. The three pillars of ***agency***, ***legibility***, and ***negotiability*** appear across much of the literature, though agency appears more than the other two. Acknowledgement and inclusion of these pillars could be as a result of research offered by Mortier *et al.* (2015) representing seminal work in the field and offering empirical contributions to the term.

Discussion and conclusions

With the pervasive nature of technology and the increasing volume of data produced and then used to drive decision-making in the BE, consideration is required in terms of how humans interact with data, at which levels this takes place, and for what purposes those data are used. Analysis of literature on the theme of HDI sees the emergence of key issues requiring further attention before a theory of HDI for the BE can be conceived. Of the technological themes and concepts evaluated (blockchain; ML and DL; robotics for industrialised construction; sensored construction sites; smart buildings; and VR and AR), there are some interesting elements that can begin to form the basis of such underpinning theory including trust, human empowerment and control, safety, accessibility and understanding of data, and of course agency, legibility and negotiability. However, we are at the beginning of the journey to propose a robust theory for HDI as highlighted by the literature and the many gaps that exist when analysed against a framework of HDI research perspectives spanning i) understanding and evaluation (end user-oriented), ii) development (technology-oriented), and iii) foundation (theory building-oriented). It seems that research has a grasp of understanding and evaluation in terms of the HDI benefits a system or concept is attempting to do for humans, but developments in technology that specifically consider and integrate HDI concepts and foundations to support theory building are somewhat lacking. Also missing across the literature is acknowledgement of education, policy and standardisation that would have implications for HDI, and consideration of the Digital Services Act.

Encapsulating the induced human-driven considerations in a theory will likely extend our understanding of HDI within the context of the analysed technologies and concepts. As an initial step toward building such a theory, future work is proposed as follows: 1) broaden the analysis of literature that considers elements of HDI, whether or not so-called, to further understand the challenges of HDI and the extent to which current technological systems and concepts address issues of HDI within and outwith the BE; 2) establish the array of purposes for which data are collected and/or used by technologies in the BE, for example, to make decisions on behalf of humans or to influence human behaviour, that will enhance our understanding of the true implications of humans' interaction with data; 3) survey individuals and organisations about production, collection, processing and storage of data to achieve such purposes as identified in 2) to identify the benefits, challenges and areas requiring attention in future studies; 4) consider the education, policy and standardisation aspects surrounding HDI; and 5) evaluate any emerging theory against technologies and concepts such as those in this paper and beyond (e.g., BIM, IoT, GIS) to establish its robustness and suitability for computer informatics in the BE.

References

Abioye, S.O. *et al*. (2021) 'Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges', *Journal of Building Engineering*, 44, 103299.

Alhakamy, A.A. *et al.* (2021) 'Show Me How You Interact, I Will Tell You What You Think: Exploring the Effect of the Interaction Style on Users' Sensemaking about Correlation and Causation in Data', in *Designing Interactive Systems Conference 2021*, pp. 564-575.

Allen, G. (2020) *Understanding AI technology*. Joint Artificial Intelligence Center (JAIC), The Pentagon United States.

Aryal, A. *et al.* (2017) 'Monitoring fatigue in construction workers using physiological measurements', *Automation in Construction*, 82, pp. 154–165.

Becherer, M. *et al*. (2020) 'Engineering a Trustworthy Private Blockchain for Operational Risk Management: A Rapid Human Data Engineering Approach Based on Human Systems Engineering', In Handley, H. and Tolk, A. (eds.) A Framework of Human Systems Engineering: Applications and Case Studies, Wiley Online Library, pp. 205-223.

Bilal, M. *et al*. (2016) 'Big Data in the construction industry: A review of present status, opportunities, and future trends', *Advanced engineering informatics*, 30(3), pp. 500-521.

Bock, T. (2015) 'The future of construction automation: Technological disruption and the upcoming ubiquity of robotics', *Automation in Construction*, 59, pp. 113–121.

Bowers, J., and Rodden, T. (1993) 'Exploding the interface: Experiences of a CSCW network', *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, pp. 255–262.

Brown, I. (2013) 'The Economics of Privacy, Data Protection and Surveillance', *SSRN*, Scholarly Paper 2358392. https://papers.ssrn.com/abstract=2358392.

Calvetti, D. *et al*. (2021) 'Human Data Interaction in Sensored Sites, Challenges of the Craft Workforce Dimension', *EC3 Conference 2021 (Online)*, 26-28 July, pp. 173-180.

Card, S. K. (2018) *The Psychology of Human-Computer Interaction*. CRC Press.

Cheng, Z. *et al.* (2016) 'Satisfaction based Q-learning for integrated lighting and blind control', *Energy and Buildings*, 127, pp. 43–55.

Cho, Y. *et al.* (2020) 'Platform design for lifelog-based smart lighting control', *Building and Environment*, 185, p. 107267.

Chowdhury, S.N. and Dhawan, S. (2016) 'HDI based data ownership model for smart cities', in *2016 International Conference on Recent Trends in Information Technology (ICRTIT)*, 8-9 April, Chennai, India: IEEE, pp. 1–5.

Crabtree, A., and Mortier, R. (2015) 'Human Data Interaction: Historical Lessons from Social Studies and CSCW', in N. Boulus-Rødje, G. *et al*. (Eds.), *ECSCW 2015: Proceedings of the 14th European Conference on Computer Supported Cooperative Work*, 19-23 September, Oslo, Norway, pp. 3–21.

Dix, A. (2017) 'Human–computer interaction, foundations and new paradigms', *Journal of Visual Languages & Computing*, 42, pp. 122–134.

EC3 (n.d.) *Human-Data Interaction Committee*. Available at: https://ec-3.org/governance/technical-committees/human-data-interaction-committee/ (accessed: 9 January 2024).

Ekblaw, A. et al. (2016) 'Bitcoin and the myth of decentralization: Socio-technical proposals for restoring network integrity', *2016 IEEE 1st International Workshops on Foundations and Applications of Self* Systems (FAS*W)*, Augsburg, Germany, pp. 18–23.

Favero, M. *et al.* (2022) 'Human-in-the-loop methods for occupant-centric building design and operation', *Applied Energy*, 325, p. 119803.

Fernandez, L. (2021) 'Teaching students how to frame human-computer interactions using instrumentalism, technological determinism, and a quadrant learning activity', *Frontiers in Computer Science*, 3, p. 771731.

Fu, Y. *et al.* (2024) 'Human-robot collaboration for modular construction manufacturing: Review of academic research', *Automation in Construction*, 158, p. 105196.

Grudin, J. (1990) 'The computer reaches out: The historical continuity of interface design', *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 261–268.

Grübel, J. *et al*. (2022) 'The hitchhiker's guide to fused twins: A review of access to digital twins in situ in smart cities', *Remote Sensing,* 14(13), 3095.

Haddadi, H. *et al*. (2013) *Human-data interaction*. (UCAM-CL-TR-837). University of Cambridge.

Halper, M. (2019) 'Target finally confirms that Acuity provides the IoT lighting', *LEDs Magazine*. Available at: https://www.ledsmagazine.com/leds-ssl-design/networks-controls/article/16699076/target-finally-confirms-that-acuity-provides-the-iot-lighting (accessed: 8 January 2024).

Hanafy, N.O. (2023) 'Artificial intelligence's effects on design process creativity: A study on used AI Text-to-Image in architecture', *Journal of Building Engineering*, 80, p. 107999.

Hunhevicz, J.J. *et al*. (2021) 'no1s1 – a blockchain-based DAO prototype for autonomous space', *Proceedings of the 2021 European Conference on Computing in Construction [Online]*, 26-28 July, pp. 27–33.

Hornung, H. *et al*. (2015) 'Challenges for Human-Data Interaction – A Semiotic Perspective', in: Kurosu, M. (eds) *Human-Computer Interaction: Design and Evaluation*. HCI 2015. Lecture Notes in Computer Science, 9169. Springer, Cham.

Jebelli, H. *et al.* (2018) 'EEG-based workers' stress recognition at construction sites', *Automation in Construction*, 93, pp. 315–324.

Jenkins, C. *et al.* (2019) 'Effective management of plug loads in commercial buildings with occupant engagement and centralized controls', *Energy and Buildings*, 201, pp. 194-201.

Joshua, L. and Varghese, K. (2014) 'Automated recognition of construction labour activity using accelerometers in field situations', *International Journal of Productivity and Performance Management*, 63(7), pp. 841–862.

The remaining references in this paper are available at this link: https://drive.google.com/file/d/1U30mDhy6BluXQPXuyopZ1VoPsSErsxGa/view?usp=drive_link.