ITcon

www.itcon.org - Journal of Information Technology in Construction - ISSN 1874-4753

INVESTIGATING THE PERSISTENCE OF UNSTRUCTURED DATA IN AEC: A CRITICAL REALIST EXPLORATION OF DOMAIN CHALLENGES IN THE IRISH CONSTRUCTION INDUSTRY

SUBMITTED: August 2025
REVISED: November 2025
PUBLISHED: November 2025
EDITOR: Žiga Turk
DOI: 10.36680/j.itcon.2025.071

Hamza Tariq, PhD Student
University College Dublin, School of Civil Engineering, Ireland
ORCID: https://orcid.org/0009-0009-2731-1603
hamzatariq70@hotmail.com

Ciaran McNally, Associate Professor
University College Dublin, School of Civil Engineering, Ireland
ORCID: https://orcid.org/0000-0003-2159-2634
ciaran.mcnally@ucd.ie

James O'Donnell, Professor
University of Limerick, School of Engineering, Ireland, University College Dublin, School of Mechanical and Materials Engineering and UCD Energy Institute, Ireland
ORCID: https://orcid.org/0000-0002-5881-9989
james.odonnell@ucd.ie, James.ODonnell@ul.ie

SUMMARY*: This study investigates the underlying domain challenges contributing to the persistence of unstructured data in the AEC industry. Following a critical realist paradigm, a qualitative study involving semi-structured interviews with professionals across various practices in the Irish construction sector were conducted. Thematic analysis revealed five key domain challenges: (1) Knowledge gaps related to fundamental data science concepts, BIM, digital twins, and linked data; (2) Stakeholder dynamics, particularly client influence and misalignment between client needs and designer proposals; (3) The "single source of truth" dilemma arising from ISO 19650 implementation challenges; (4) Process inertia stemming from email dependence, perceived benefits of unstructured data, underdeveloped models, and skills deficiencies; and (5) Contractual and regulatory requirements that fail to incentivize structured data adoption. Mapping these themes to a people, process and technology framework revealed that the majority of these problems are systematic in nature. These findings suggest that the industry needs a multifaceted approach addressing knowledge enhancement, stakeholder alignment, standard simplification, workflow modernization, and regulatory integration to overcome the barriers to structured data implementation. This study identifies valuable insights for policymakers, educators, and industry leaders seeking to facilitate automation in construction information management.*

KEYWORDS*: unstructured data, AEC, qualitative study, thematic analysis, semi-structured interviews.*

REFERENCE*: Hamza Tariq, Ciaran McNally & James O'Donnell (2025). Investigating the persistence of unstructured data in AEC: A critical realist exploration of domain challenges in the Irish construction industry. Journal of Information Technology in Construction (ITcon), Vol. 30, pg. 1728-1748, DOI: 10.36680/j.itcon.2025.071*

COPYRIGHT*: © 2025 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.*

1. INTRODUCTION

The Architecture, Engineering and Construction (AEC) industry faces a critical challenge in managing and retrieving information effectively and automatically. Studies indicate that approximately 95% of data go unused and 13% of the working time is spent searching for information (Giovanardi et al., 2023; Snyder et al., 2018). This inefficiency stems from the industry's heavy reliance on unstructured data, meaning information that machines cannot readily interpret and is stored in various formats such as Pdfs (natural language text documents), images and videos (Soibelman et al., 2008), with textual data comprising over 80% of the total (Wu et al., 2022). While Building Information Modelling (BIM) has been introduced to promote structured information delivery, documents remain fundamental to construction projects due to their intuitive nature and practical utility for engineers and site teams (Soman and Whyte, 2020). Consequently, accessing, retrieving and managing information recorded in documents usually depend on human intervention and can be labour intensive and error prone (Arantes and Falbo, 2010).

This challenge is not new and has been a focus of computational linguistics. Natural Language Processing (NLP) and text mining coupled with machine learning (ML) and deep learning (DL) and most recently, even large language models (LLMs) are being applied to transform the data into structured formats (Chowdhary, 2020; B. Li et al., 2024). But these techniques themselves face challenges such as lacking the domain-specific understanding, difficulty in capturing semantic relationships, and the need for large, annotated training datasets (W. Li et al., 2024). Moreover, these methods often extract isolated pieces of information but fail to identify the deeper, contextual knowledge embedded in construction documents (Aleqabie et al., 2024).

It is therefore essential to understand the underlying causes behind this persistent reliance on unstructured data. To investigate such causes, this research adopts a critical realist paradigm, a philosophical approach more common in the social sciences but increasingly relevant for complex socio-technical problems in engineering domains. This perspective enables the investigation of both observable practices and the deeper institutional, technical, and cultural mechanisms that influence them (Gorski, 2013). By examining these hidden causes, this research aims to identify foundational issues that must be addressed before semantic, machine-readable information systems can be effectively deployed in construction workflows.

This investigation is grounded in the Irish construction sector, where national digitalisation initiatives such as the Build Digital Project and the public sector BIM mandate (Hore et al., 2023) provide a timely backdrop for examining why unstructured data practices persist despite active policy pushes for structured information. As a research hypothesis, it is proposed that currently Irish AEC firms are not yet prepared to fully adopt automated mechanisms for information retrieval and management due to a lack of necessary knowledge, expertise, and established semantic frameworks. This proposition is tested through qualitative research, where it is proposed that Ireland's relatively small and interconnected industry can offer valuable insights into the dynamics at play.

The remainder of this paper consists of four further sections. Section 2 discusses the background and gaps in existing research; Section 3 describes the methodology employed in the research; Section 4 presents the results of the interview analysis; Section 5 discusses how the findings relate to the wider literature and future research directions and finally, conclusions are drawn in Section 6.

2. BACKGROUND

As the AEC industry becomes increasingly digitalised, the applications of information and communication technologies such as BIM, geographic information systems (GIS), wireless sensors, and digital twins are also increasing rapidly (Huang et al., 2021). However, effective implementation of these technologies relies on robust information retrieval, meaning getting the right information at the right time. This requires data to be delivered in structured formats (Shen et al., 2024). For instance, the IFC schema is built upon structured object classes, attributes, and relationships that enable clash detection, quantity take-offs, and interdisciplinary coordination (Taherdoost et al., 2024). Similarly, GIS data are built upon structured spatial data formats like GML or GeoJSON that enable efficient map-based querying and spatial analysis (Atkinson et al., 2022). Digital twins also require time-series data from IoT sensors to be structured in ways that enable real-time retrieval and predictive insights (Ghorbani and Messner, 2024). Similarly, conversational artificial intelligence (AI) systems depend on structured data linked through ontologies to ensure accurate interpretation and context-aware responses (Saka et al., 2023). In the absence of properly structured input, retrieving useful information becomes challenging due to semantic

CC BY

inconsistencies, missing identifiers and fragmented documentation. For instance, integrating sensor data with asset registries depends on shared vocabularies (Rejeb et al., 2022). Without this, retrieving consistent inputs becomes a tedious and difficult task. Hence, structured data form the foundation for automation by making the data machine-readable, queryable, and usable across various platforms and use cases.

2.1 Structured data

Structured data are classified as the data that are organized in a particular format that allows machines to interpret them (Elmasri and Navathe, 2016). For instance, existing inside relational databases or spreadsheets as rows and columns. They can include tabular data, labelled fields, and explicitly defined relationships between entities (Soibelman et al., 2008). Berners-Lee (2006) has stated a set of 'rules' for creating structured data.

1. The first constitutes indexing the data and storing them on online servers so that they can be easily accessed by both computers as well as humans.
2. The second rule involves structuring the data with relevant schemas for easy interpretation by machines.
3. The third rule is to make the schemas public and machine readable by using open-source schemas to describe the data. Open-source schemas are important as proprietary data formats limit data inferencing as the schemas by which data are modelled are only accessible to few applications.

Based on these principles, in the context of the AEC, instances of structured data are spreadsheets, IFC files, GML files, etc. In contrast, unstructured data don't follow a particular format and are therefore, not interpretable by machines. Examples of unstructured data include video files, image files and textual data.

To get value out of unstructured data and generate useful insights, the field of computational linguistics has evolved significantly to convert them into structured formats. NLP, a subfield of AI, has evolved significantly to address the challenge of automated text analysis and understanding (Chowdhary, 2020). Modern NLP systems employ sophisticated ML and DL algorithms, including convolutional neural networks and recurrent neural networks, to process and interpret vast amounts of unstructured textual data, enabling automatic extraction of key insights and patterns (Harisha and Bhat, 2024). Text mining, a specialized subset of NLP, employs techniques such as tokenization, named entity recognition, and parts of speech tagging to systematically identify, categorize, and retrieve valuable information from textual data (Aleqabie et al., 2024). In recent years, researchers have also begun exploring LLMs for similar purposes (Li et al., 2024). These technologies have also found extensive application in the AEC sector. A recent review by Shamshiri et al., (2024) noted 205 publications related to NLP and text mining dating back to the early 2000s. Similarly, Yan et al. (2022) analysed 127 publications from 2000 to 2021, demonstrating the extensive body of work in the sector. These techniques, while valuable, are not without limitations and have their own challenges. First, they require large datasets and substantial computational power (Li et al., 2024), and their performance declines when applied to short texts (Sakor, 2023). ML methods struggle with highly complex data, such as compound texts, nested entities, and varying data representations (Abdullah et al., 2022). While promising, DL approaches to information extraction are still in their early stages, revealing numerous avenues for future research (Abdullah et al., 2023). A key limitation of text mining is its focus on extracting isolated information from text while failing to adequately capture underlying semantic relationships and contextual nuances (Aleqabie et al., 2024).

2.2 Limitations of existing research

Although addressing the challenge of unstructured data in the AEC sector has received increasing attention, the majority of existing research focused on technical solutions (Bucher et al., 2024; Yan et al., 2020). In comparison, fewer studies have examined *why unstructured data remain the dominant form of information*, or the underlying organizational and social factors that contribute to this reliance. Existing qualitative research often targets broader themes such as digital transformation or BIM implementation, rather than specifically focusing on unstructured data. For instance, a recent review by (Antai et al., 2025) revealed that 75% of peer-reviewed journal articles concentrated on general digital adoption patterns. This shows the lack of studies regarding underlying issues associated with unstructured data.

Moreover, it is essential to observe and learn from current practices rather than relying solely on expert-driven, top-down modelling (El-Diraby, 2023). Many prior studies investigated generalized industry trends without grounding their analysis in the lived experiences of practitioners working across different roles and project types

CC BY

(Antai et al., 2025). Consequently, important details such as practitioners' perceptions about data structures have been overlooked. This gap is also evident in a recent empirical study by Jaskula et al. (2024) which examined technical and workflow issues in common data environments (CDEs) through semi-structured interviews but did not address perceptions of data structures. A similar limitation appears in (Abanda et al., 2025), who focused on enhancing processes and workflows for ISO 19650 adoption but did not explore how data structures are understood by practitioners. These details are crucial for understanding the practical obstacles to adopting structured data solutions (Bilal et al., 2016).

To address these limitations, it is essential to conduct qualitative studies that capture the perspectives of industry practitioners. Such empirical insights can guide and inform future technical developments (Hartmann, 2008). Given Ireland's emphasis on digital transformation through initiatives such as the Build Digital Project and the national BIM mandate (Hore et al., 2023), this study uses the Irish context as a case study. Viewing unstructured data as foundational for future technical developments, this exploratory analysis examines these issues through practitioner perspectives.

By adopting a qualitative research methodology, this study aims to uncover the underlying reasons for the persistence of unstructured data in the AEC sector and to identify the domain challenges that contribute to this phenomenon. These identified domain challenges are then be mapped to a people, process, and technology framework (Tripathi et al., 2024) to provide clear distinctions regarding whether these challenges are systematic or technical in nature. This approach provides a more in depth understanding of the challenges and opportunities associated with unstructured data generation, ultimately informing the development of more effective and context-specific solutions.

3. METHODOLOGY

This exploratory study followed a critical realist research paradigm (Eastwood et al., 2021), which seeks to understand underlying issues by examining root causes and asking why and how phenomena occur. This approach assumes that real causes exist beneath observable problems, even when not immediately visible. In terms of this study, a critical realist paradigm views unstructured data prevalence as resulting from deeper, often hidden social and institutional mechanisms that operate independently of individual awareness.

Through semi-structured interviews and thematic analysis, this research aimed to identify these underlying factors and explain the persistence of unstructured data in the AEC industry despite the availability of structured data alternatives. Semi-structured interviews were chosen for this study as they enable the adoption of a conversational format with open-ended questions for in-depth discussion (Schwartz-Shea and Yanow, 2015); this allows some leeway to follow-up on whatever is deemed important by the researcher (Brinkmann, 2020). Analysis of the interviews utilised a thematic analysis from (Braun and Clarke, 2006)**,** as this method is well established in the literature to identify, categorise and report patterns within the data.

The overall work methodology involved sourcing participants, developing an interview protocol, and obtaining participant consent. This was followed by conducting and recording interviews, then performing thematic analysis on the recorded transcripts. The thematic analysis (Braun and Clarke, 2006) involved the following six iterative steps:

1. Familiarising with the data
2. Generating initial codes
3. Generating initial themes
4. Reviewing themes
5. Defining and naming themes
6. Writing the results

This iterative process enabled the identification of recurring patterns and insights, resulting in the development of 5 themes and 10 subthemes. To demonstrate whether the underlying issues related to unstructured data persistence were systematic or technical, the themes and subthemes were mapped to the people, process, and technology framework (Tripathi et al., 2024). In this framework, *people* refers to human factors such as skills, knowledge, and behaviours; *process* encompasses organisational workflows, policies, and practices; and *technology* includes digital tools, platforms, and systems. This framework provided a holistic view of how human, procedural, and technical factors interact to sustain unstructured data practices in the sector.

CC BY

3.1 Research method

The paper adopted the four stepped approach to qualitative descriptive study (Figure 1) as recommended by (Villamin et al., 2024). The methodological rigour was ensured by following the consolidated criteria for reporting qualitative research (COREQ) guideline (Tong et al., 2007). The completed checklist is included as supplementary material. The study was conducted by adhering to a protocol developed with University College Dublin's Research Ethics Committee. The initial step involved participant sampling, followed by data collection. The data collected was then analysed, and lastly, the findings were determined. These steps are detailed in the following sections.

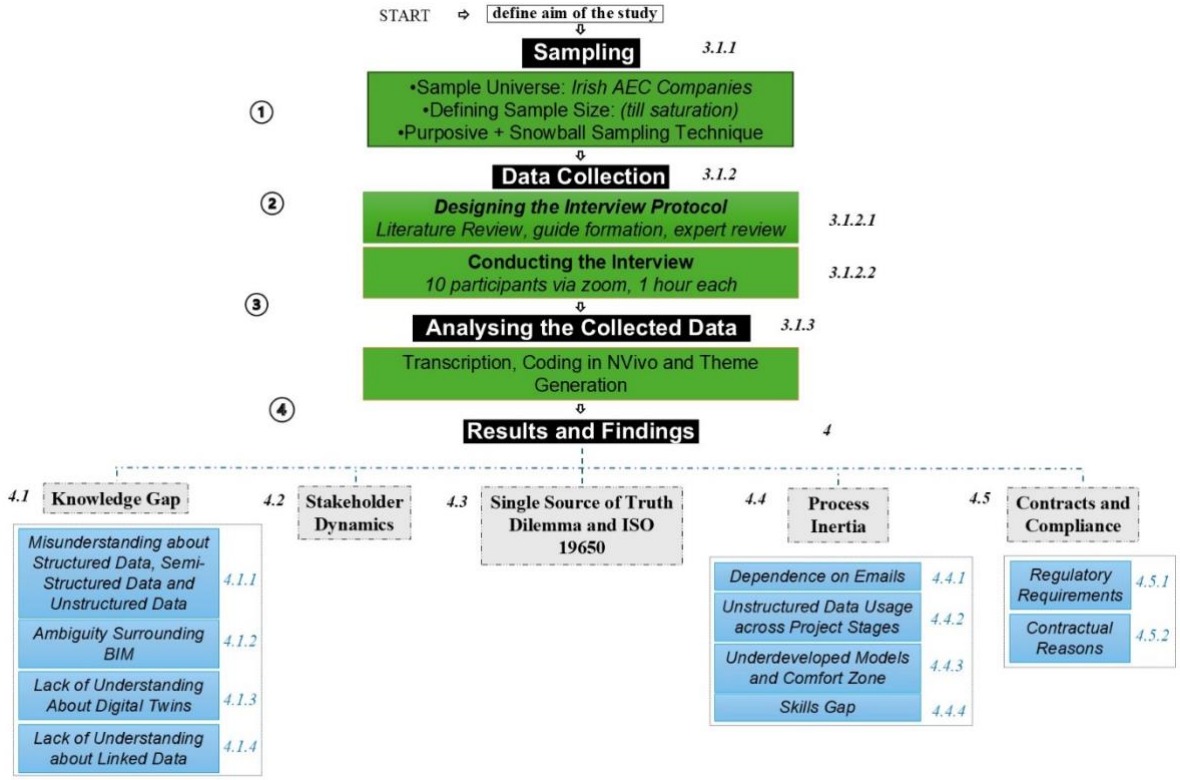

Figure 1: The four steps of this study: sampling participants, collecting data, analysing data and presenting findings.

3.1.1 Sampling

The qualitative sampling process was guided by Robinson (2014) four point framework. The first step was about defining the sample universe. For this study, construction companies within Ireland were targeted. To focus on the study and meet the objectives, inclusion criteria were applied, limiting participants to those with 5+ years of experience and knowledge both in traditional construction methods as well as digital technologies. The criteria of a minimum of 5 years' experience was consistent with previous qualitative studies (Serrador and Pinto, 2015). The second step was choosing an appropriate sample size. The study adopted a flexible sample size where data would be collected until saturation was achieved. The next step was choosing an appropriate sampling strategy. Since the focus of this study was to uncover the underlying issues regarding unstructured data in the sector, a purposive sampling strategy was adopted to ensure variation in company sizes, their practices and practitioner's role. The final step was to source the sample. Since the study focused on interviewing professionals with specific requirements, the snowball sampling (also known as referral sampling) was used to reach the participants.

The profiles of participants are summarized in Table 1. To achieve data saturation following (Guest et al., 2006) recommendations, participants were recruited in phases. In the first phase, 11 individuals were contacted via email, of whom 4 agreed to participate. In the second phase, 3 additional individuals were approached, with 1 agreeing to participate. Those who did not participate did not respond to follow-up emails and were therefore not included in the study. No participants withdrew or dropped out after initially agreeing to take part. After these 5 interviews,

CC BY

data saturation was achieved by the sixth interview, aligning with (Guest et al., 2006) recommendations. Theme accumulation charts (Figure 6 & Figure 7) are provided in Appendix E. However, to strengthen the findings as recommended by (Braun and Clarke, 2013), a third recruitment phase was conducted where 5 more individuals were contacted (2 via LinkedIn), yielding a final sample of 10 participants representing a homogeneous mixture of architects, designers, and contractors from both large companies and SMEs. All of these participants were independently recruited, and there were no prior personal or professional relationships between them and the researchers.

Table 1: Participant Profile: Role, Scale and Associated Practice.

#	Participants Code	Practice	Scale	Role
1	P1	Architectural	Micro	Director
2	P2	Multi-Disciplinary Consultancy	Large	Associate Director - Mechanical Lead
3	P3	Architectural	SME	Data Manager and Senior Architect
4	P4	Architectural	SME	Director
5	P5	Contractor	SME	BIM Coordinator
6	P6	Contractor	Large	BIM Manager
7	P7	Multi-Disciplinary Consultancy	Large	Director – Senior Bridge Engineer
8	P8	Multi-Disciplinary Consultancy	Large	Senior Structural Engineer
9	P9	Contractor	SME	Digital Operations Manager
10	P10	Contractor	Large	Digital Project Delivery Lead

3.1.2 Data collection

Data were collected in the following two phases:

Designing the interview protocol

The interview protocol was developed based on the framework provided by (Kallio et al., 2016), which included identifying the prerequisites for using semi-structured interviews, conducting a literature review (as discussed in the introduction and background), formulating a preliminary interview guide, seeking expert assessment, and finalizing the complete guide. The expert assessment was provided by the PhD researcher's guidance committee. This process resulted in primary questions centred on two key themes: unstructured data understanding and organizational factors, ensuring the study maintained clear focus. As the interviews were semi-structured, questions were designed to align with the research objective while allowing flexibility for follow-up inquiries based on participants' responses. The protocol was not pilot tested; however, expert assessment from the guidance committee and the semi-structured design provided sufficient flexibility to refine and clarify questions during the interview process when necessary. The complete interview protocol is provided in Appendix A.

An information sheet containing consent form was developed and provided to the participants, detailing an overview of the research study. This included information about the primary investigators, (a PhD researcher and an Associate Professor), the research focus, objectives, and assurance that while data would be used for publication, participants' identities would remain anonymous.

Conducting the interview

As mentioned earlier, a total of 10 participants were recruited through referrals based on the inclusion criteria. During recruitment, a one-page introduction sheet was provided to clearly distinguish between unstructured and structured data. Following Soibelman *et al.*, (2008), these distinctions were made:

- Structured Data: Data that follow a particular format or order and are machine-readable and exist in formats such as BIM files and spreadsheets.
- Unstructured Data: Data that do not follow a particular format and are not machine-readable, such as textual content in emails, documents, audio, and video files.

CC BY

After agreeing to participate, participants received the information sheet with a consent form developed as part of the interview protocol. Their consent was obtained for audio and video recording and for using the results for academic discovery. All interviews were conducted online via Zoom, each lasting approximately one hour. Each session involved only the researchers and the participant, with no third parties present. They were conducted by the PhD researcher (male) and an Associate Professor (male). The PhD researcher specialized in BIM and information management and had received formal training in qualitative research methods through a semester-long postgraduate module. The Associate Professor, with a background in structural engineering, had extensive experience in conducting and supervising qualitative research studies. Although both researchers had prior knowledge of the subject matter, care was taken to utilize open ended, non-leading questions to minimize interpretive bias. The recordings were transcribed using Zoom AI and MS Word, then securely stored in a cloud-based repository provided by the researcher's university. Transcripts were not returned to participants for verification, as participants indicated that they preferred to receive a summary of the findings rather than review full transcripts. All participants will be provided with a summary of the research findings upon publication, and no repeat interviews were conducted.

3.1.3 Analysing the collected data

After extracting and formatting the transcripts from the recordings, data were analysed following the six-stepped thematic analysis process outlined by Braun and Clarke, (2022). NVivo 20 was used for this analysis, with the license provided by the researcher's university. As part of this process, the PhD researcher first developed initial codes after thoroughly familiarizing himself with the data. These codes were then critically analysed to generate themes, which were subsequently examined and refined under the guidance of both associate professors who are co-authors of this paper. *The number of coders was one.* Finally, the themes were integrated to create a coherent and compelling narrative of the data. To ensure transparency while maintaining anonymity, all coded text is included as supplementary material accompanying this paper.

4. RESULTS AND FINDINGS

This section describes the results of the study. Five key domain challenges are proposed and are used as the basis for the discussion on the persistence of unstructured data. These themes are presented in detail in this section and are mapped against people process and technology framework highlighted in Figure 2 and Figure 3. Additionally, Appendix B provides a detailed frequency distribution illustrating how many participants discussed each theme and how often each theme was mentioned in all the ten interviews. A thematic coding tree is provided in Appendix C.

4.1 Knowledge gaps

This theme highlights four key knowledge gaps as a foundational issue contributing to the persistence of unstructured data. These knowledge gaps are discussed in the subsequent subthemes.

4.1.1 Misunderstandings about structured data, semi-structured data and unstructured data

There are misunderstandings regarding fundamental data science concepts relating to structured data, unstructured data and semi-structured data (structured ways of sharing data). Some participants outwardly admitted they did not understand what unstructured data meant despite receiving an information sheet and a brief explanation prior to recording of the interviews. For instance, P3 and P8 remarked:

"What would you consider unstructured data source?" (P3)

"By unstructured sources you mean not in the model, is it? (P8)

Some participants believed that applying a proper file naming convention made the data structured, while this would be considered at most as semi-structured. This was described by P5 who said:

"Within the ISO 19650 we address those problems through naming suitability and provision folder structure and the lifecycle of the assets and through an internal folder structure. And so,… this is the structure that we follow to name every single document, regardless of what it is, regardless of whether it's structured data or unstructured. Through this way an unstructured data becomes structured through the life cycle of the project." (P5)

CC BY

A similar confusion was also expressed by P3.

"They usually impose some sort of a common data environment which is the highest structured kind of information that we can possibly share." (P3)

In summary, confusion persists regarding the basic understanding of structured data, unstructured data and structured ways of sharing data.

Themes & Subthemes	People	Process	Technology
4.1 Knowledge Gaps	☒		
4.1.1 Misunderstanding about Structured Data, Semi-Structured Data and Unstructured Data	☒		
4.1.2 Ambiguity Surrounding BIM	☒		
4.1.3 Lack of Understanding about Digital Twins	☒		
4.1.4 Lack of Understanding about Linked Data	☒		
4.2 Stakeholder Dynamics	☒	☒	☒
4.3 Single Source of Truth Dilemma and ISO 19650		☒	☒
4.4 Process Inertia	☒	☒	☒
4.4.1 Dependence on Emails	☒	☒	
4.4.2 Unstructured Data Usage across Project Stages	☒	☒	
4.4.3 Underdeveloped Models and Comfort Zone	☒	☒	☒
4.4.4 Skills Gap	☒		☒
4.5 Contracts and Compliance		☒	
4.5.1 Regulatory Requirements		☒	
4.5.2 Contractual Reasons		☒	

People: Refers to human skills, roles, and behaviours
Process: Encompasses workflows, policies, and practices
Technology: Involves tools, platforms, and systems

Figure 2: Mapping of themes and subthemes to People, Process and Technology framework.

4.1.2 Ambiguity surrounding BIM

BIM is widely misunderstood. Participants revealed that many people in the industry still equate BIM primarily with 3D modelling or specific tools such as Revit. As P9 remarked,

"BIM is still interpreted as being just a 3D geometry" (P9)

BIM's broader value as an information management process is often misunderstood. Rather than considering it an integrated part of project delivery, it is treated as a task to be completed in isolation. This results in BIM deliverables being outsourced to technicians who are not involved in design, and models being treated as secondary outputs rather than information coordination tools. As a result, unstructured data in the form of drawings and sketches are shared. This frustration was expressed by P10 who said:

"On some of our projects, the design team will refuse to issue us over a model and they're saying, no, you're not getting a model. You're getting the drawings. If you want a model, you need to recreate one yourself, which again, I think is maybe a shift in the industry that definitely needs to happen." (P10)

Due to these misunderstandings, BIM has evolved into an ambiguous term that holds different meaning for different people. P7 succinctly described this perspective, stating:

"It's a very loosely used term and this kind of means different things to different people. So, we do assess here through our technical competency matrix as part of the annual training reviews. People have their own perception or knowledge of BIM and it does tend to stay quite low relative to other competencies like bridge design or road design or buildings structures design. Even though we do look to train people and the likes, they just still consider that they don't have much knowledge in the area. So, I think a lack of understanding is still a problem." (P7)

Naturally, if BIM is misunderstood, OpenBIM workflows are also poorly understood. For example, IFC is commonly seen as just a file format rather than as a schema. as P2 remarked:

"I'll be honest. I'm not up to speed with sharing data via IFC files. It's not something that we would share" (P2)

This lack of understanding leads to improper application of IFC. This was reflected by P2 later in the interview that importing IFC files into thermal simulation software often leads to data loss, so he prefers to redraw the model manually. However, he acknowledged that the problem may lie in how the file was exported by the architect:

"It may not be a software issue. It might be down to how the model is physically drawn by the architect. So, there might be a very specific way to draw the model, so that the thermal simulation software can read it. But if the architect is not drawing it that way, then the software won't read it." (P2)

Additionally, as the majority of the Irish AEC industry operates within the Autodesk ecosystem, challenges related to interoperability are largely avoided. Consequently, the benefits of IFC are not fully recognised or realised in practice.

"I prefer to stay as native as possible in the Autodesk environment. That for me, is for very practical purpose." (P9)

"In terms of a use case for IFC. I haven't seen any great advantages really at the moment." (P10)

In summary, significant misunderstandings persist regarding the definition, intended purpose, and practical implementation of BIM and openBIM across the industry.

4.1.3 Lack of understanding about digital twins

There is a lack of understanding about Digital Twins. P9 pointed out that his colleagues lack understanding for it. He is trying his best to instil that understanding by developing a use case with the help of Autodesk Tandem.

"I'm pitching that in the company as our digital twin platform" (P9)

This was also echoed by P6:

"When it comes to digital twins and asset information, we have found in the last 2 or 3 years that clients don't understand this, and they don't really want it." (P6)

4.1.4 Lack of understanding about linked data

Linked Data only came up in discussion with P7 and P10. P10 admitted that he isn't aware of what that is. P7 expressed a foundational understanding for it, but he admitted that the concept is still too technical for it to be understood and adopted in the industry.

"Linked Data is very powerful, and the use of ontologies is very powerful. ...But there aren't that many people who have expertise in it. And it is twice specialist, and the major software vendors haven't really adopted it and don't seem to be particularly willing to implement it in their systems. So, I think while in theory, it's potentially very useful and while it's been shown to be useful on at a small scale, certain use cases. I'm not yet convinced it'll be adopted on a broader scale." (P7)

4.2 Stakeholder dynamics

Unstructured data persistence is influenced by stakeholders such as clients, designers and software companies. Among these, clients were identified by participants as the most influential. They are seen as the primary drivers of innovation. When clients explicitly demand structured data, the entire supply chain would respond accordingly by training or hiring staff and adapting workflows to meet those requirements. As P9 explained:

"We won't do it until we have to do it, until there is a requirement there from the client side." (P9)

While a few clients in the pharmaceutical and life sciences sectors are well aware of what they require, this is not the norm across the industry. Many clients either lack understanding of their information requirements or are unable to express them in ways that support structured data workflows.

CC BY

"The life sciences are very good that they know what they want, and it's very clear, and they've done it before…. whereas when I'd be working in more typical construction, you know, commercial or residential and stuff, and the clients tended to not really know what they wanted." (P8)

"When it comes to digital twins and asset information, we have found in the last 2 or 3 years that clients don't understand this, and they don't really want it." (P6)

This lack of understanding results in vague or excessive requests. For example, asking for a federated as-built model without specifying what it should include or how it will be used. In practice, many clients adopt models without having the internal systems or workflows in place to use them effectively, often reverting to traditional methods like pdf documents for facilities management. P10 explained this issue:

"The client might say that they want a federated as-built model at the end of the project, but they don't know what requirements or what assets they want tagged in there. For us, that can be difficult to try to understand what the client needs and what they want to maintain within the building... Some clients are like, 'Oh, well, if they got an as-built model on their project, I kind of want one too,' but their own systems aren't set up to use a model. I think they'll still end up looking for the other documentation." (P10)

Secondly, designers also play a big role. Clients who lack understanding of what they require often rely on designers to pitch in and write their information requirements. However, this backfires when designers themselves lack the understanding and as a result over-specify or suggest that all possible data must be included. This leads to ambiguity between what the client truly needs and what the designer thinks the client needs, resulting in over-specification and the production of unnecessary data. P6 shared a case where a designer advised the client to request COBie and other detailed information. Months were spent generating that data, only to later discover that the facility management (FM) team didn't need most of it. In fact, the facility manager told him what he requires with the help of an interesting example.

"The manager had a very interesting way of describing what assets he was interested in.
He said to me, 'Imagine you're a giant, and you turn up beside that building. You pick up the whole building, turn it upside down, and shake it. Anything that falls out? That's what I'm interested in. Anything that stays stuck to the shell and the walls? Not interested.' That was a really interesting way to picture what he was looking for." (P6)

This ambiguity between client needs and designers proposed needs was also highlighted by P10 who said:

"I think it's more of an issue where the design team themselves maybe don't know what the client wants either, and the kind of understanding of it is a bit vague sort of like, okay, just give me all the information, and it's a failsafe, and that if we receive everything over, well, then no one can do any wrong if that makes sense like, you know. Otherwise, if they only specified certain elements to us. Well, then, all of a sudden, if we deliver it at the end. The client might say, okay, well, we're missing X parameter from the model." (P10)

Lastly, software companies are another stakeholder. If clients want structured data in a certain format and the software companies don't incorporate them in their tools, then that information could not be structured in that way. As P7 mentioned:

"The final big stakeholder is the software companies themselves because if they don't implement something, then it's difficult." (P7)

In conclusion, the persistence of unstructured data can be attributed to a combination of limited client understanding, inadequate designer expertise, and functional limitations within the software.

4.3 Single source of truth dilemma and ISO 19650

ISO 19650 advocates for streamlining information management processes. However, the various documents produced whilst following this standard contribute to unstructured data. These documents are verbose and time-consuming to create, often necessitating a dedicated role for their management. Reproducing such documentation for different projects is considered burdensome. Moreover, these documents are also criticized for being overly theoretical, resulting in paperwork that may have limited practical value.

CC BY

"Our main ones really are BEP, TIDP, MIDP... we have a document manager dedicated to developing those documents." (P6)

"It's just recreating the same documents again and again, which can be quite tedious." (P10)

"It's too complicated, too verbose, too standard-based, and too theoretical. It's full of red tape and filler, which turns people off immediately. So, what we tend to do is extract the essential parts and condense them into cheat sheets." (P9)

Beyond the problems highlighted above, the real dilemma arises with the implementation of a single source of truth. While ISO 19650 permits the use of multiple CDEs, this approach leads to duplication and loss of metadata, compromising the single source of truth concept. Consequently, tracking information becomes difficult.

"The more CDEs, the more problems you end up facing." (P5)

"If I upload a drawing to Procore, it requests metadata and defines what the document is, which works fine. But when someone who doesn't have access to Procore needs a copy, we download and email it to them, resulting in all that metadata being lost." (P9)

"We are currently working with about six different platforms, various CDEs, each with different workflows and requirements. As a result, we've created our own tracking systems to gain a better overview of what's happening across these external CDEs. While we provide and share information through the CDEs, we still rely on our own internal tools to understand the activities occurring in each of them." (P3)

However, this issue might not necessarily be a reflection of ISO 19650 itself, but rather how software vendors have interpreted and implemented the standard in their CDEs. P10 described the distinction between folder-based and file-based CDEs. While file-based CDEs can help preserve a single container ID (metadata), the most popular tools, such as Autodesk Construction Cloud (ACC) are folder-based. This inconsistency in implementation further contributes to duplication and unstructured data.

"Some CDEs, are kind of like a file-based approach, and others are kind of like a folder based. And so, to give, maybe, like an example, ACC will be kind of a folder based CDE, where permissions are granted based on folders, and as a result you could have one document in maybe 2 or 3 different folders… We would prefer to have one container id for a file and its whole life cycle on the project. It sits under that one container, so you can see when it was uploaded, when it was reviewed, when it was approved, when it was then maybe revised at a later date. And that's what we kind of prefer. But some CDEs don't allow that and they would prefer based off a folder permission." (P10)

4.4 Process inertia

Process inertia is also a significant contributor to unstructured data prevalence. This theme unfolds the reasons leading to process inertia. Several reasons are spotted, and these are discussed in the following subsections.

4.4.1 Dependence on emails

Emails remain the default method for sharing information despite causing file duplication, as they are perceived as more trustworthy, transparent, and formal. That is why, when information is conveyed through phone calls, it is always followed by an email. They also ensure communication history is retained even when employees leave. Emails are seen as easier for sharing noncritical information, such as task assignments, since common data environments often generate overwhelming notifications. People also trust email more for privacy, as they are unsure who can access content in shared environments. As a result, email remains a key source of unstructured data.

"I think, there's still a bit of legacy that people like to follow things up with emails. Even if there is something on a common data environment where there's a comment assigned. Or you know, I think there's still a bit of legacy where people like to follow that up with emails. And it's probably seen as a more formal way of doing it." (P2)

"I suppose the main benefit of those is that if that person leaves that, you always have access to certain emails." (P2)

"Once they're emailed. They're being duplicated. And that still is kind of a bad habit of the industry" (P9)

CC BY

4.4.2 Unstructured data usage across project stages

Unstructured data have their purposes and benefits across project stages. During the design development stage, visuals, reports and sketches are used to convey ideas, align stakeholders, and secure approvals. During the tender stage, specifications, reports and information management procedures in textual format are needed for contract management and approval. During the project execution phase, meeting minutes, site visit records, issue trackers and spreadsheets are used for day-to-day collaboration and record keeping. Pdfs are preferred for record-keeping over structured models due to their perceived reliability.

"What are other options? What could be the possible alternative to that? If not file based systems? What else could we be working with" (P3)

"A lot of presentation materials at the early stages that might be quite well developed, like visuals and things that are going to sell the project and based on images." (P3)

"It might be photographs, aerial imagery or reports, input from 3rd parties, public consultation documents and the likes route selection reports, environmental impact assessment study reports, all the way through then planning." (P7)

"Keeping Pdfs so far seem to be the most robust way of keeping the project records for prolonged periods of time." (P3)

4.4.3 Underdeveloped models and comfort zone

Unstructured data persists as the models are deemed unusable due to lack of information or details required for FM. This highlights that models are often not developed with the end purpose in mind. This can be due to a variety of factors. Design teams and subcontractors may only partially model elements or rely heavily on 2D drawings, even when 3D models are contractually required. This fundamental issue is associated with models being treated as a deliverable rather than an information model, and critical details such as those needed for FM are often missing or considered too difficult to include. This frustration was expressed by P10.

"I don't want to be picking on a design team. But design team and subcontractors, they don't want to produce the models to a good enough level to be able to track and incorporate the Building Control Amendments Regulations (BCAR) information on. And there's still like a lot of elements that are not modelled or only modelled in 2D. And again, we're just coming off a project here, there was deliverable for a design team to produce a model in 3D. But they've only ended up producing it in 2D. And now we need to figure out the rest of the coordination, and they said that they may produce a model at the end of the project and give it over to us. But again, that's kind of defeating the whole purpose of our coordination and of our exercises." (P10)

While the above frustration was expressed by P10 who represents a contractor, this confession was made by P3 as well who represents an architect.

"I think we're still working on our models to a scale of about 1:20 or 1:10, but not down to the level of detail that we might actually issue on a detailed drawing that shows intent, shows a typical interface that might be applied in multiple scenarios. And that might be 1: 5 or 1: 10, but like it would have more detail in it than we're actually modelling. Like when it comes to membranes and vapor barriers and things like that. That would actually be very difficult to include in a model file." (P3)

P3 further remarked:

"Because things get quite contractual for us when a project is on site. It's much better to be able to refer to a Pdf. That shows our exact design intent rather than a model that would have to be completely perfect in every single junction, rather than just having like one drawing." (P3)

Due to technical difficulty, Teams fall back on pdf to communicate design intent, especially when accuracy at every junction is not feasible in the model. Even if models include all the details, FM team would still not use them in their day-to-day work as it would take a lot of time for them to open the models and find the relevant information amongst all the information included in the model. They consider their own booklets or manuals to be faster, quicker and efficient and hence, prefer to access information through systems and formats tailored to their workflows. As a result, the model becomes disconnected from actual project needs, and its potential value remains unrealized. P6 described this scenario,

CC BY

"Facilities manager is not going to go off into a model. He doesn't use models. Normally. he's not going to go in there and start sectioning a model and navigating around and finding the floor and clicking on it and finding his information. He's not going to do that. He's going to go to his own manual or his facilities management software where he has his data structure the way he wants it…. He's gonna find it 20 times quicker than going off and looking in the model." (P6)

Moreover, decision makers in the company consider BIM expensive and overly complicated, and they tend to rely more on their past experience rather than thinking about the future. As P9 said,

"Because BIM is still regarded as a cause of extra cost and over complication. The people in construction companies who tend to make the decisions about software and approach tend to be more senior members, directors, and they're accustomed to working the old way. They lean on their experience, and they look backward on their 30 years of experience as opposed to 3 years in the future" (P9)

4.4.4 Skills gap

Unstructured Data persists due to lack of skills in delivering structured information, particularly in relation to BIM. Participants stated that effective use of BIM tools requires a high level of proficiency, and without a skilled team, the potential benefits of BIM such as reducing construction costs or reducing errors are not fully realized. As P1 stated:

"I don't see why using a Revit model is going to result in a reduced construction cost and arguably it might reduce the opportunity for errors, and therefore cost arguably, but that would rely on a whole team being extremely proficient on that platform." (P1)

According to P2, the gap lies in knowledge and expertise as opposed to any technical reasons.

"It's a knowledge gap and an expertise gap as opposed to any technical reason." (P2)

P8 being a structural engineer representing a large firm admitted that structural engineers are not upskilled to use the BIM tools.

"Structural engineers typically don't do the modelling…. we're still in the Pdf markup stage in ABC company at the moment… It's mainly because the structural engineers haven't been upskilled in BIM, and they're not allowed to use it…. we don't actually have a BIM workflow from analysis to modelling at the moment. So, we asked them to just model, based on sketches or markups." (P8)

Additionally, even contractors working with BIM models often lack knowledge of key processes, such as managing asset registers or working with COBie data standards. As P4 shared:

"Even the contractors that we've been working with, they had to go in training. They didn't have any knowledge of asset register or COBie." (P4)

This gap is more profound in the small contractors. P6 shared an extreme example of their skills gap, describing how a small civil company still produces their drawings manually using a drawing board. He remarked,

"When I was trying to onboard this person in there, I could tell he wasn't understanding what I was talking about. So, I tried to look at it from a different angle, and I said: Look, how do you produce your drawings, your fabrication drawings, so that your team know how long to cut a piece of pipe and what angle it should be used. And he said, well, typically, I'll take the Pdf that's given to me by the designers and then I'll sit down on the drawing board with a T-square and a set square and a pencil, and I'll draw out my drawing. And I couldn't believe it, that company was still using pen and paper to get their drawings together." (P6)

Lastly, P3 highlighted that even when they are using spreadsheets, they are not using it to create structured data but rather using it as a textual document, highlighting the wrong application of the tool.

"We would also consider spreadsheets generally a text document as well, because these are not calculation spreadsheets as such. They contain a lot of material like written material." (P3)

These skills gaps lead to the creation of unstructured data, as the required expertise to deliver structured, interoperable information is often lacking across the project team.

CC BY

4.5 Contracts and compliance

Even if the industry acquires the necessary knowledge and skills, it is still unlikely that the change to producing structured formats will occur if the requirements are not stipulated in the contract and regulatory documents. Therefore, the contract and regulatory requirements remain one of the significant barriers to unstructured data prevalence in the industry.

4.5.1 Regulatory requirements

Although the Capital Works Management Framework (CWMF) in Ireland has mandated the use of BIM in public projects, regulatory requirements continue to depend on traditional, document-based submissions. The BCAR necessitates the submission of compliance documentation including design certificates, ancillary certificates, and handover certificates in pdf format. These certificates must be uploaded to the Building Control Management System (BCMS) for compliance approval. Currently, no mechanism exists for model-based permitting.

"There's an ancillary certificate of design commencement, ancillary certificate of design completion, and then an ancillary certificate of design inspection. So, there's 3 certs that we have to do. But each individual engineering discipline will have to produce those, and then the contractor will have to produce. There're separate certs, the contractor has to submit. They all are packaged together and given to the assigned certifier, who then produces his overall cert, which is given to the local authority to say that you know the certified of the building has followed the building regulations." (P2)

"We don't yet have model-based planning submissions in Ireland." (P7)

These insights reveal that while policy frameworks like the CWMF encourage the use of structured, interoperable data through BIM, statutory processes and regulatory reporting continue to operate in parallel, non-digitalized silos.

4.5.2 Contractual reasons

If clients require structured data, they must make it financially feasible for the entire supply chain. Currently, designers are being compensated while contractors are not. This indicates that if contractors are not paid for it, they would not deliver it. P7 emphasized this by narrating the story of Finnish industry where they spent a lot of time and effort in writing requirements and upskilling the people, but they still struggled to obtain the required data. They later discovered the reason for that was they simply didn't make it worthwhile for the contractors.

"They didn't make it worthwhile for the contractors, while all the payments were related to actually delivering the construction, and at the end of the job, it wasn't really in the contractor's interest to spend lots more resources and time tidying off all the information and handing over to the client to comply with the contractual obligations. So, it's important, if clients really want this information exchanged in a particular way, they have to associate that to a meaningful amount of money in the contract. Basically, they have to be willing to pay for it, and they have to put that clearly into the contract." (P7)

This concern was also expressed by P1 who is affiliated with a micro company.

"No one's paying me to learn BIM. And no one's paying me for the extra cost to use it in the office." (P1)

Other than payments, the contract should also clarify the copyright requirements, otherwise the suppliers would be unwilling to share their models. P8 expressed this concern, noting that his team deliberately removes embedded information before sharing models:

"There has to be something in it for everyone. ... we'd strip out all the information usually when we issue models so that people couldn't take our IP and things like that. So, they just get the geometry. So, like this data rich thing, and then you strip out all the data before you share it" (P8)

In addition, P7 questioned the practical clarity and enforcement of IFC requirements within CWMF:

"A broad ranging insistence on IFC, is a little bit, I have a bit of difficulty with it you know. The question is, why and what's the purpose? And what's the benefit? And who's going to pay for it? Are they willing to pay? And is it actually going to be once you're in a project, as someone actually going to insist on it being done and someone going to monitor and validate it that it's been done properly?" (P7)

CC BY

In summary, this subtheme highlights that obtaining structured data is hindered by unclear contractual requirements, inadequate compensation, and lack of copyright protection for the supply chain.

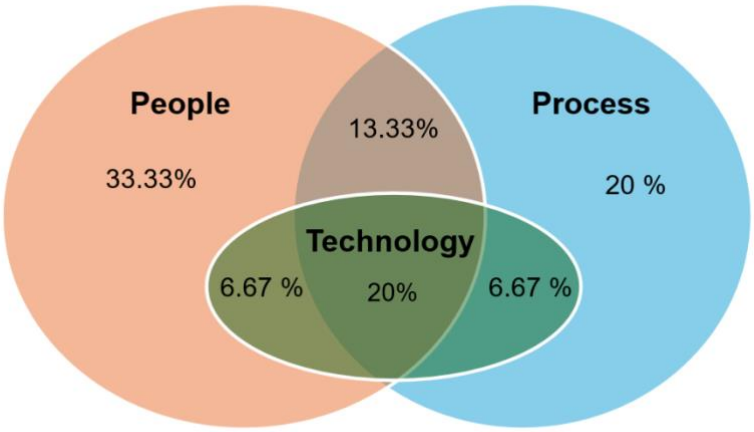

Figure 3: Venn diagram showing the percentage and overlap of themes and subthemes across the People, Process, and Technology framework.

Mapping the themes to people, process and technology framework revealed that 33.33% of challenges related to people factors, 20% to process factors, and 0% to technology factors. A 13.33% overlap existed between people and process categories. A 6.67% overlap was observed between technology and people and technology and process categories. The overlap across all three categories was 20%. In total, 66.66% of identified challenges are systematic (*people and process, excluding technology overlap*) in nature. These percentages were calculated by counting the total number of themes and subthemes highlighted in Figure 2 and determining how many times each category (People, Process, Technology) appeared. These counts were then converted into percentages by dividing the number of occurrences per category by the total number of themes and subthemes. Table 3 (provided in Appendix D) illustrates these calculations.

5. DISCUSSION

Very little qualitative research has been conducted on the underlying reasons behind the persistence of unstructured data in the Irish AEC industry. Various challenges have been identified, and an overwhelming majority (around 66.66 %) of these challenges are systematic in nature. For future adoption of semantic and machine interpretable systems, it is of utmost importance to address such challenges. This section discusses those challenges as well as future research direction. These proposed directions also incorporate participant insights gathered in response to the final interview question.

Firstly, the industry lacks knowledge in four key areas. These pertain to fundamental concepts in data science, including an understanding of structured, unstructured, and semi-structured data as well as BIM, digital twins, and linked data. Targeted efforts are needed to enhance awareness and understanding of these domains across the sector. This may include the organisation of seminars and workshops involving clients, architects, contractors, and designers (Bosch-Sijtsema et al., 2021). In parallel, higher education institutions should embed core data science principles within both undergraduate and postgraduate programmes to strengthen foundational competence in these areas. Despite the introduction of BIM over a decade ago, it remains widely misunderstood and is often interpreted in varying ways. This suggests that BIM continues to evolve as both a concept and an idea. The findings indicate a pressing need to re-evaluate current definitions and educational strategies (Borkowski, 2023). A key question emerging from the data is why professionals continue to associate BIM primarily with 3D modelling. This may reflect a failure to effectively communicate its broader value and potential beyond geometric representation. Similarly, the benefits of IFC are not communicated well. There is a need to articulate clearly: *what*

purpose does IFC serve when the Irish industry operates within the Autodesk ecosystem? Furthermore, the concept of digital twins is also not well understood, underscoring the need for clearer articulation of its purpose and benefits (Ghorbani and Messner, 2024). Lastly, linked data is perceived as overly technical and insufficiently mature for adoption at a national scale. To support its wider uptake, efforts should be directed toward simplifying its conceptual framework and making it more accessible to non-technical stakeholders.

Secondly, stakeholders such as clients, designers, and software vendors play a critical role in influencing data practices within the industry. The findings indicate a clear need to upskill clients, enabling them to articulate their information needs more effectively. In addition, designers must also be upskilled to understand client specific requirements as findings highlighted an ambiguity between what clients require and what designers propose, often leading to the over-specification of information requirements. Enhancing communication and alignment between clients and designers can help mitigate this issue (Wolf et al., 2024). Once this issue is resolved, software vendors can then be influenced to respond to practical needs. As software vendors are primarily driven by market demands and business incentives, a clear articulation of industry needs rooted in client requirements is essential. Establishing this feedback loop is key to ensuring that software solutions evolve in line with the sector's growing demand for structured data.

Thirdly, ISO 19650 was highlighted as overly verbose and technical by the participants. This indicates a need to simplify both the language and structure of the standard to enhance its accessibility for practitioners. Therefore, researchers should investigate the possibility of sharing the data within EIR, BEP, MIDP, and TIDP in structured formats such as spreadsheets or models. Moreover, the flexibility within ISO 19650 that permits the use of multiple CDEs has raised concerns. This practice often leads to duplication, loss of metadata, and challenges in maintaining a single source of truth (Jaskula et al., 2024). These issues suggest a need to revisit and potentially revise the standard to provide clearer guidance on multiple CDEs usage. For projects that involve multiple CDEs, there are clear tensions between those that are folder-based and those that are file-based. It is clear that when using multiple CDEs there is a need for a shared area to avoid duplication and maintain a single source of truth. What is not clear is how folder-based systems such as ACC can accommodate this. It is accepted that ISO 19650 focuses on the concept of 'information containers', which are not necessarily files. However, the current reality is that almost all information containers are files, and as such there is a real possibility of folder-based systems compromising the single source of truth. Further investigation is required to determine how compliance is evaluated and whether the certification of such platforms is compatible with the standard's intended outcomes.

Fourthly, process inertia within the industry appears to be due to several interrelated factors, such as a continued reliance on email communication, perceived advantages of unstructured data, underdeveloped models, and a persistent skills gap. This aligns with findings by Pinheiro (2019) who noted that handover practices often result in voluminous but unstructured documentation, with key operational requirements poorly captured or entirely missing. While emails remain a reliable method for sharing information, alternative workflows should be explored such as referencing files via links to container IDs within CDEs instead of attaching documents directly. This approach is already possible due to certain commercial software such as pro forma konekt. Implementation guidance is available on the vendor's website (Konekt, 2025). When it comes to unstructured data, their benefits to convey design intent, win tenders, or support site documentation are recognized. Future research should investigate if such benefits can be replaced by structured data. However, replacing unstructured data with structured formats will require more agile delivery methods, rather than traditional waterfall approaches that attempt to define exhaustive requirements upfront, often causing implementation to be difficult (Zou et al., 2023). To address the challenge of underdeveloped models, clients and designers must take a more proactive approach, as previously discussed. Addressing the industry's skills gap will require a coordinated educational response, including government-led training programmes, industry workshops, and the integration of relevant content into university and college curricula.

Lastly, regulatory requirements and contractual obligations significantly contribute to the persistence of unstructured data in the industry. The findings reveal that, in order to comply with building regulations, information is often submitted in standard pdf formats, running in parallel to the objectives of the Irish BIM mandate. This highlights the need to integrate regulatory processes into digital workflows and suggests that the Irish construction sector should begin to explore digital permitting as a viable solution (Beach et al., 2024). Moreover, the findings make it clear that unless structured data is explicitly stipulated within contractual agreements, it will not be delivered by the supply chain. Therefore, structured data requirements must be contractually embedded and,

CC BY

importantly, linked to financial incentives. Without such mechanisms, the motivation to adopt structured data practices across the supply chain is likely to remain low.

5.1 Limitations and future work

This study has several limitations that should be acknowledged. Firstly, it is geographically limited to Ireland, and the interviews primarily reflect the context and practices of the Irish AEC industry. Secondly, the sample size is restricted to ten interviews. Future studies could expand both the number of participants and the geographic scope to gain broader insights. Thirdly, qualitative research inherently involves a degree of subjectivity. To mitigate this, and in line with the recommendations of (Thomas and Harden, 2008), the study clearly outlines its methodological approach. While qualitative methods are suitable for the exploratory nature of this research, future work could adopt a more detailed weighting or ranking system such as that employed by (Durdyev et al., 2022) to produce more nuanced and quantifiable findings. Lastly, the participants who were interviewed for this study were professionals with over five years of experience and held senior positions. While this provided valuable strategic perspectives, future research could benefit from interviewing junior professionals, whose insights may reveal different or overlooked aspects of practice.

6. CONCLUSION

Unstructured data remains deeply embedded in day-to-day practices of the Irish AEC industry, hindering automatic information retrieval. Despite BIM mandates and push for automation, documents, emails, and 2D sketches dominate project workflows. This research identified five core reasons:

1. Limited understanding of structured data and concepts like BIM, IFC, digital twins and linked data.
2. Vague or misaligned client requirements
3. The burdensome and inconsistent application of ISO 19650 coupled with the loss of single source of truth due to multiple CDEs
4. Inertia due to reliance on emails, current advantages of unstructured data and skills gap.
5. The absence of contractual and financial motivation to produce structured data.

Analysis of the interviews has shown that practitioners often confuse semi structured data with structured data, treat BIM as 3D modelling only, and default to Pdfs for clarity, record-keeping, and legal assurance. Even when BIM models are created, they are frequently underdeveloped or disconnected from practical needs, especially for FM. This highlights that problem isn't just technical, but also organizational, contractual, and educational. This would suggest that in the Irish context, Efforts to implement structured data must account for above mentioned realities. Upskilling clients to express clear information needs, training designers to interpret and respond meaningfully, and influencing software vendors through market demand are steps that would support this transition. The provision of support to simplify standards such as ISO 19650 (in terms of language and structure), and guidance on managing multiple CDEs to reduce duplication and preserving the single source of truth would also be of benefit. Regulatory bodies could also move toward digital permitting, and structured data delivery could be financially incentivized through contracts. Additionally, educational institutions could embed core data principles in AEC curricula, and researchers should explore lightweight, agile delivery methods over rigid waterfall-style methods. Ultimately, this study highlights that structured data will emerge not just from top-down requirements but from bottom-up practices, reflecting how people actually work. This balance between formal systems and lived experiences is key to turning structured data from an exception into the industry norm.

ACKNOWLEDGEMENTS

The research team would like to thank all the subject matter experts for participating in this research and sharing their insights. The authors would like to acknowledge the support from the Build Digital Project, funded by the Department of Public Expenditure, National Development Plan Delivery and Reform, Ireland.

REFERENCES

Abanda, F.H., Balu, B., Adukpo, S.E., Akintola, A., 2025. Decoding iso 19650 through process modelling for information management and stakeholder communication in BIM. Buildings 15, 431. https://doi.org/10.3390/buildings15030431

CC BY

Abdullah, M.H.A., Aziz, N., Abdulkadir, S.J., Akhir, E.A.P., Talpur, N., 2022. Event detection and information extraction strategies from text: A preliminary study using GENIA corpus. Presented at the International Conference on Emerging Technologies and Intelligent Systems, Springer, pp. 118–127. https://doi.org/10.1007/978-3-031-20429-6_12

Abdullah, M.H.A., Aziz, N., Abdulkadir, S.J., Alhussian, H.S.A., Talpur, N., 2023. Systematic literature review of information extraction from textual data: recent methods, applications, trends, and challenges. IEEE Access 11, 10535–10562. https://doi.org/10.1109/ACCESS.2023.3240898

Aleqabie, H.J., Sfoq, M.S., Albeer, R.A., Abd, E.H., 2024. A review of textmining techniques: Trends, and applications in various domains. Iraqi Journal for Computer Science and Mathematics 5, 9. https://doi.org/10.52866/ijcsm.2024.05.01.009

Antai, I., Lenka, S., Achtenhagen, L., 2025. Digital platforms and the construction supply chain: trends and emerging themes in extant AEC research. Construction Management and Economics 43, 113–129. https://doi.org/10.1080/01446193.2024.2396592

Arantes, L. de O., Falbo, R. de A., 2010. An infrastructure for managing semantic documents. Presented at the 2010 14th IEEE International Enterprise Distributed Object Computing Conference Workshops, IEEE, Vit ria, Brazil, pp. 235–244. https://doi.org/10.1109/EDOCW.2010.17

Atkinson, R.A., Hunter, A., Car, N.J., Purss, M.B.J., Cochrane, B., 2022. Roadmap for interoperable 3d data models in ogc apis and other data exchange approaches. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLVIII-4/W4-2022, 13–20. https://doi.org/10.5194/isprs-archives-XLVIII-4-W4-2022-13-2022

Beach, T., Yeung, J., Nisbet, N., Rezgui, Y., 2024. Digital approaches to construction compliance checking: Validating the suitability of an ecosystem approach to compliance checking. Advanced Engineering Informatics 59, 102288. https://doi.org/10.1016/j.aei.2023.102288

Berners-Lee, T., 2006. Linked Data [WWW Document]. Linked Data - Design Issues. URL https://www.w3.org/DesignIssues/LinkedData.html (accessed 4.23.25).

Bilal, M., Oyedele, L.O., Qadir, J., Munir, K., Ajayi, S.O., Akinade, O.O., Owolabi, H.A., Alaka, H.A., Pasha, M., 2016. Big Data in the construction industry: A review of present status, opportunities, and future trends. Advanced Engineering Informatics 30, 500–521. https://doi.org/10.1016/j.aei.2016.07.001

Borkowski, A.S., 2023. Evolution of BIM: epistemology, genesis and division into periods. Journal of Information Technology in Construction (ITcon) 28, 646–661. https://doi.org/10.36680/j.itcon.2023.034

Bosch-Sijtsema, P., Claeson-Jonsson, C., Johansson, M., Roupe, M., 2021. The hype factor of digital technologies in AEC. Construction Innovation 21, 899–916. https://doi.org/10.1108/CI-01-2020-0002

Braun, V., Clarke, V., 2022. Thematic analysis: a practical guide. SAGE, London ; Thousand Oaks, California.

Braun, V., Clarke, V., 2013. Successful qualitative research: A practical guide for beginners. SAGE, London.

Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. Qualitative Research in Psychology 3, 77–101. https://doi.org/10.1191/1478088706qp063oa

Brinkmann, S., 2020. Unstructured and semistructured interviewing, in: Leavy, P. (Ed.), The Oxford Handbook of Qualitative Research. Oxford University Press, pp. 424–456. https://doi.org/10.1093/oxfordhb/9780190847388.013.22

Bucher, D.F., Hunhevicz, J.J., Soman, R.K., Pauwels, P., Hall, D.M., 2024. From BIM to Web3: A critical interpretive synthesis of present and emerging data management approaches in construction informatics. Advanced Engineering Informatics 62, 102884. https://doi.org/10.1016/j.aei.2024.102884

Chowdhary, K.R., 2020. Natural language processing, in: Chowdhary, K.R. (Ed.), Fundamentals of Artificial Intelligence. Springer India, New Delhi, pp. 603–649. https://doi.org/10.1007/978-81-322-3972-7_19

Durdyev, S., Mohandes, S.R., Tokbolat, S., Sadeghi, H., Zayed, T., 2022. Examining the OHS of green building construction projects: A hybrid fuzzy-based approach. Journal of Cleaner Production 338, 130590. https://doi.org/10.1016/j.jclepro.2022.130590

Eastwood, J.G., De Souza, D.E., Mukumbang, F.C., 2021. Realist research, design and evaluation for integrated care initiatives, in: Amelung, V., Stein, V., Suter, E., Goodwin, N., Nolte, E., Balicer, R. (Eds.), Handbook Integrated Care. Springer International Publishing, Cham, pp. 629–656. https://doi.org/10.1007/978-3-030-69262-9_37

El-Diraby, T., 2023. How typical is your project? The need for a no-model approach for information management in AEC. Journal of Information Technology in Construction 28, 19–38. https://doi.org/10.36680/j.itcon.2023.002

Elmasri, R., Navathe, S., 2016. Fundamentals of database systems, Seventh edition. ed. Pearson, Boston Munich.

Ghorbani, Z., Messner, J., 2024. A categorical approach for defining digital twins in the AECO industry. Journal of Information Technology in Construction 29, 198–218. https://doi.org/10.36680/j.itcon.2024.010

Giovanardi, M., Konstantinou, T., Pollo, R., Klein, T., 2023. Internet of things for building façade traceability: A theoretical framework to enable circular economy through life-cycle information flows. Journal of Cleaner Production 382, 135261. https://doi.org/10.1016/j.jclepro.2022.135261

Gorski, P.S., 2013. "What is critical realism? And why should you care?" Contemporary Sociology: A Journal of Reviews 42, 658–670. https://doi.org/10.1177/0094306113499533

Guest, G., Bunce, A., Johnson, L., 2006. How many interviews are enough?: An experiment with data saturation and variability. Field Methods 18, 59–82. https://doi.org/10.1177/1525822X05279903

Harisha, S., Bhat, S., 2024. Comparative analysis on classification of unstructured text data and specific summarization. Presented at the 2024 Second International Conference on Advances in Information Technology (ICAIT), IEEE, Chikkamagaluru, Karnataka, India, pp. 1–5. https://doi.org/10.1109/ICAIT61638.2024.10690829

Hartmann, T., 2008. A grassroots model of decision support system implications by construction project teams (Doctoral Dissertation). Stanford University, Stanford, California.

Hore, A., McNally, C., Adeleke, A.Q., Saif, W., 2023. Supporting the delivery of information management and BIM in Ireland: Learnings from Selected International Regions. Technological University Dublin. https://doi.org/10.21427/KVZJ-M250

Huang, Y., Shi, Q., Zuo, J., Pena-Mora, F., Chen, J., 2021. Research status and challenges of data-driven construction project management in the big data context. Advances in Civil Engineering 2021, 6674980. https://doi.org/10.1155/2021/6674980

Jaskula, K., Kifokeris, D., Papadonikolaki, E., Rovas, D., 2024. Common data environments in construction: state-of-the-art and challenges for practical implementation. Construction Innovation. https://doi.org/10.1108/CI-04-2023-0088

Kallio, H., Pietilä, A.-M., Johnson, M., Kangasniemi, M., 2016. Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. Journal of Advanced Nursing 72, 2954–2965. https://doi.org/10.1111/jan.13031

Konekt, N., 2025. A Guide to mastering emails and other project communications in the aeco industry [WWW Document]. URL https://www.newforma.com/publications/a-guide-to-mastering-emails-and-other-project-communications/ (accessed 5.18.25).

Li, B., Jiang, G., Li, N., Song, C., 2024. Research on large-scale structured and unstructured data processing based on large language model. Presented at the MLPRAE 2024: The International Conference on Machine Learning, Pattern Recognition and Automation Engineering, Association for Computing Machinery, Singapore Singapore, pp. 111–116. https://doi.org/10.1145/3696687.3696707

CC BY

Li, W., Song, Y., Wang, S., 2024. Intelligent Computing、Computational Power、 Computational Power Networks and Technology Ecosystems. Presented at the 2024 IEEE 14th International Conference on Electronics Information and Emergency Communication (ICEIEC), IEEE, Beijing, China, pp. 1–8. https://doi.org/10.1109/ICEIEC61773.2024.10561638

Pinheiro, S.V., 2019. Requirements specification to enable efficient environmental and energy operation of buildings (Doctoral Dissertation). University College Dublin, Ireland.

Rejeb, A., Simske, S., Fosso Wamba, S., Smart, E., Breslin, J., yapa, kosala, Ospina, A., Bradau, H., Crowley, B., Desaraju, S., Liu, K., Thakur, S., Martingale, W., Keogh, J., 2022. Charting Past, Present, and future research in the semantic web and interoperability. Future Internet 14, 161. https://doi.org/10.3390/fi14060161

Robinson, O.C., 2014. Sampling in interview-based qualitative research: A theoretical and practical guide. Qualitative Research in Psychology 11, 25–41. https://doi.org/10.1080/14780887.2013.801543

Saka, A.B., Oyedele, L.O., Akanbi, L.A., Ganiyu, S.A., Chan, D.W.M., Bello, S.A., 2023. Conversational artificial intelligence in the AEC industry: A review of present status, challenges and opportunities. Advanced Engineering Informatics 55, 101869. https://doi.org/10.1016/j.aei.2022.101869

Sakor, A., 2023. Knowledge extraction from unstructured data (Doctoral Dissertation). Leibniz Universität Hannover, Germany.

Schwartz-Shea, P., Yanow, D. (Eds.), 2015. Interpretation and method: empirical research methods and the interpretive turn, Second edition. ed. Routledge, London New York.

Serrador, P., Pinto, J.K., 2015. Does Agile work? — A quantitative analysis of agile project success. International Journal of Project Management 33, 1040–1051. https://doi.org/10.1016/j.ijproman.2015.01.006

Shamshiri, A., Ryu, K.R., Park, J.Y., 2024. Text mining and natural language processing in construction. Automation in Construction 158, 105200. https://doi.org/10.1016/j.autcon.2023.105200

Shen, X., Sepasgozar, S.M.E., Ostwald, M.J., 2024. Knowledge-based semantic web technologies in the AEC sector. Automation in Construction 167, 105686. https://doi.org/10.1016/j.autcon.2024.105686

Snyder, J., Menard, A., Spare, N., 2018. Big data= big questions for the engineering and construction industry. White Paper. First Myanmar Investment (FMI). Raleigh, US.

Soibelman, L., Wu, J., Caldas, C., Brilakis, I., Lin, K.-Y., 2008. Management and analysis of unstructured construction data types. Advanced Engineering Informatics 22, 15–27. https://doi.org/10.1016/j.aei.2007.08.011

Soman, R., Whyte, J., 2020. Codification challenges for data science in construction. Journal of Construction Engineering and Management 146, 04020072. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001846

Taherdoost, H., Farhaoui, Y., Shahamir, eyed R., Le, T.-V., Madanchian, M., Prasad, M., Okonta, E.D., Vukovic, V., Rahimian, F., Hayat, E., 2024. Prospective directions in the computer systems industry foundation classes (ifc) for shaping data exchange in the sustainability and resilience of cities, in: EAI 3rd International Conference on Smart Technologies and Innovation Management. Springer Nature Switzerland, Cham, pp. 27–42.

Thomas, J., Harden, A., 2008. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC medical research methodology 8, 45. https://doi.org/10.1186/1471-2288-8-45

Tong, A., Sainsbury, P., Craig, J., 2007. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. International Journal for Quality in Health Care 19, 349–357. https://doi.org/10.1093/intqhc/mzm042

Tripathi, A., Nassereddine, H., Sturgill, R.E., Dadi, G.B., Hatoum, M.B., Ammar, A., 2024. People, Process, and Technology maturity levels for successful technology implementation by state departments of transportation. Transportation Research Record: Journal of the Transportation Research Board 2678, 12–21. https://doi.org/10.1177/03611981231170002

CC BY

Villamin, P., Lopez, V., Thapa, D.K., Cleary, M., 2024. A worked example of qualitative descriptive design: A step-by-step guide for novice and early career researchers. Journal of Advanced Nursing 81, 5181–5195. https://doi.org/10.1111/jan.16481

Wolf, G., Wogan, M., Mohan, N., Nejat, N., Gross, R., Menzel, K., 2024. Advancing AEC Project Management: A model-based and data-driven approach for sustainable practices. Presented at the 2024 European Conference on Computing in Construction. https://doi.org/10.35490/EC3.2024.198

Wu, C., Li, X., Guo, Y., Wang, J., Ren, Z., Wang, M., Yang, Z., 2022. Natural language processing for smart construction: Current status and future directions. Automation in Construction 134, 104059. https://doi.org/10.1016/j.autcon.2021.104059

Yan, H., Ma, M., Wu, Y., Fan, H., Dong, C., 2022. Overview and analysis of the text mining applications in the construction industry. Heliyon 8, e12088. https://doi.org/10.1016/j.heliyon.2022.e12088

Yan, H., Yang, N., Peng, Y., Ren, Y., 2020. Data mining in the construction industry: Present status, opportunities, and future trends. Automation in Construction 119, 103331. https://doi.org/10.1016/j.autcon.2020.103331

Zou, Y., Guo, B.H., Papadonikolaki, E., Dimyadi, J., Hou, L., 2023. Lessons learned on adopting automated compliance checking in the AEC industry: A global study. Journal of Management in Engineering 39, 04023019. https://doi.org/10.1061/JMENEA.MEENG-5051

CC BY

APPENDIX A: INTERVIEW PROTOCOL

This appendix outlines the interview protocol employed for conducting semi-structured interviews, detailing the structure, question sequencing and key thematic areas explored.

Part 1: Contextual Profile (*first 5 minutes of the interview*)

1. Organizational Profile: Information is collected about participant's organization (type, typical projects, and client base).
2. Professional Profile: Information is collected about participant's professional background (role and experience).

Part 2: Understanding Unstructured Data (*approximately ½ of the interview*)

The following areas are investigated:

- The purposes and drivers behind document creation during projects.
- The distinction between mandatory project documents and internally generated documentation.
- Communication practices using audio, video, and email formats.
- Time spent extracting information from unstructured sources and its impact on project delivery.
- The persistence of file-based delivery despite structured approaches like BIM.

Part 3: Organizational factors (*approximately ½ of the interview*)

The following factors are investigated:

- Document duplication practices within their workflows
- Stakeholder influence on data structuring in project documentation
- Company-specific practices that contribute to unstructured data creation
- Potential measures to reduce unstructured data and facilitate digital construction transition

Part 4: Open Discussion (*last 5 minutes of the interview*)

Participants are invited to share any additional insights or important aspects of the topic not covered in the preceding discussion.

CC BY

APPENDIX B: FREQUENCY DISTRIBUTION OF THEMES AND SUBTHEMES

This appendix provides a frequency table (Table 2) and a visualization (Figure 4) showing how frequently each theme and subtheme was identified across the ten participant interviews. The table presents two metrics.

1. The number of participants who discusses each theme.
2. The total number of times each theme was mentioned across all interviews.

This distribution demonstrates the consistency of patterns across participants and provides an overview of the most prevalent challenges identified in the study. Figure 4 displays the number of mentions for each theme and subtheme through a horizontal bar chart.

Table 2: Distribution of Themes and Subthemes by Number of Participants and Coding Instances.

Themes	Subtheme	Number of Participants (n=10)	Number of Mentions
Knowledge Gaps		10	52
	Misunderstanding about Structured Data, Semi-Structured Data and Unstructured Data	10	18
	Ambiguity Surrounding BIM	10	26
	Lack of Understanding about Digital Twins	3	5
	Lack of Understanding about Linked Data	2	3
Stakeholder Dynamics		10	43
Single Source of Truth Dilemma and ISO 19650		8	18
Process Inertia		10	114
	Dependence on Emails	9	29
	Unstructured Data Usage across Project Stages	10	52
	Underdeveloped Models and Comfort Zone	7	16
	Skills Gap	9	17
Contracts and Compliance		10	29
	Regulatory Requirements	7	10
	Contractual Reasons	10	19

CC BY

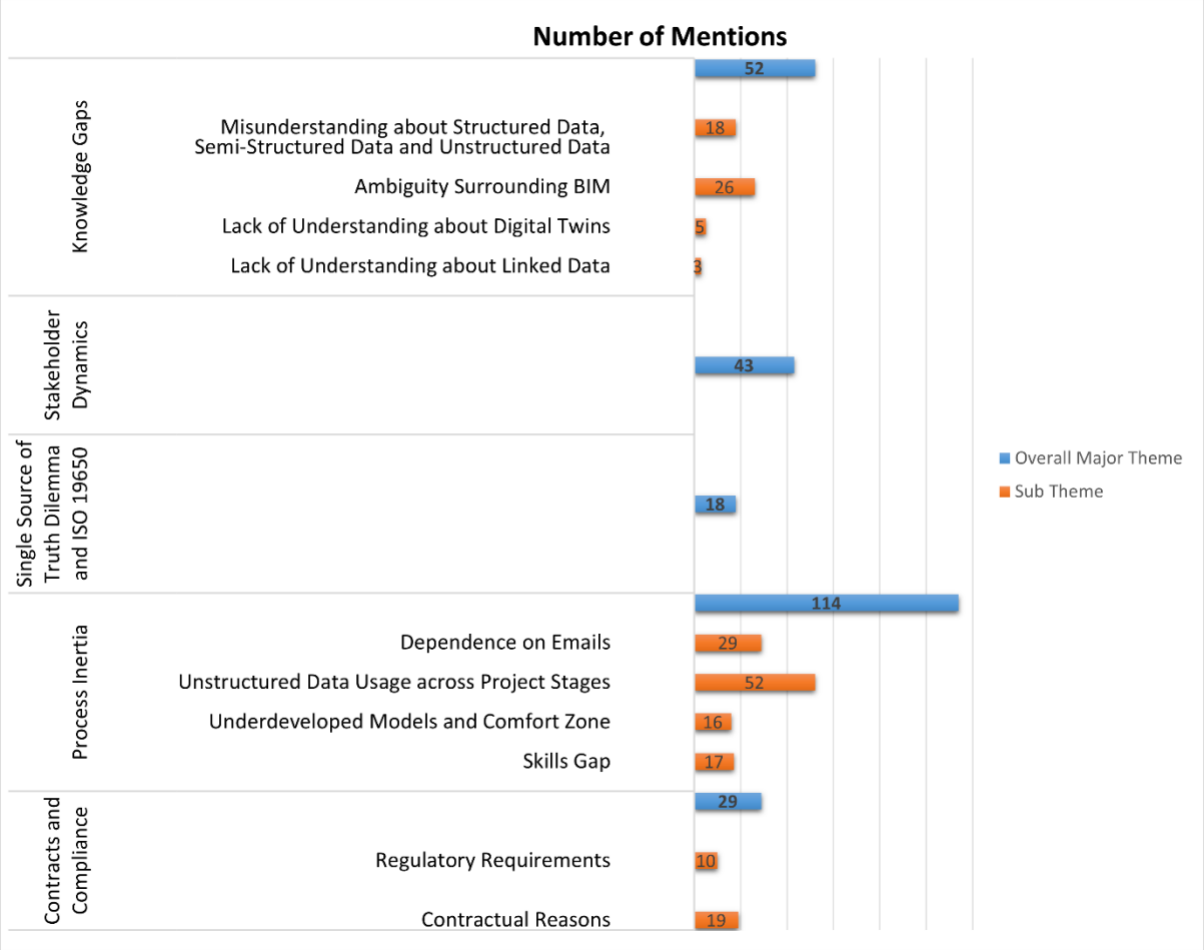

Figure 4: Frequency Distribution of Major themes and subthemes based on number of mentions across the interviews.

APPENDIX C: THEMATIC CODING TREE

This appendix represents the complete coding tree (Figure 5) developed during thematic analysis, illustrating the hierarchical organization of major themes and their constituent subthemes.

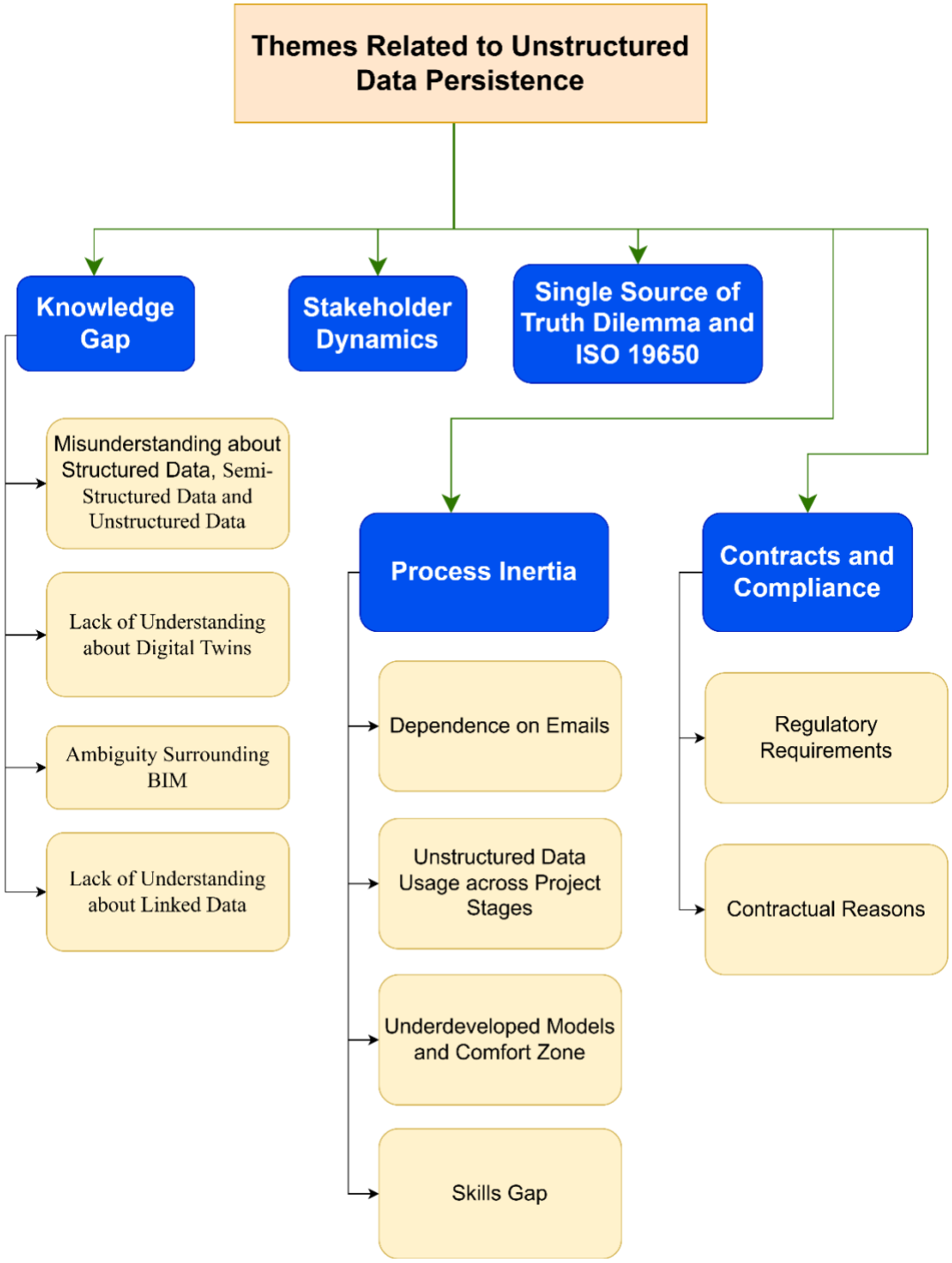

Figure 5: Thematic Coding Tree.

APPENDIX D: CALCULATION OF PERCENTAGES FOR PEOPLE, PROCESS AND TECHNOLOGY MAPPING

This appendix presents a summary Table 3 showing how the 15 identified themes and subthemes were distributed across the People–Process–Technology framework. The counts and percentages were calculated based on the total number of themes and subthemes depicted in Figure 2.

CC BY

Table 3: Calculation of Percentages for People - Process - Technology Mapping.

Category	Count of Themes and Subthemes (out of 15)	Percentage of Total (%)	Themes and Subthemes (From Figure 2)
People Only	5	33.33	4.1 Knowledge Gaps 4.1.1 Misunderstanding about Structured Data, Semi-Structured Data and Unstructured Data 4.1.2 Ambiguity Surrounding BIM 4.1.3 Lack of Understanding about Digital Twins 4.1.4 Lack of Understanding about Linked Data
Process Only	3	20	4.5 Contracts and Compliance 4.5.1 Regulatory Requirements 4.5.2 Contractual Reasons
Technology Only	0	0	-
People + Process	2	13.33	4.4.1 Dependence on Emails 4.4.2 Unstructured Data Usage across Project Stages
People + Technology	1	6.67	4.4.4 Skills Gap
Process + Technology	1	6.67	4.3 Single Source of Truth Dilemma and ISO 19650
People + Process + Technology	3	20	4.2 Stakeholder Dynamics 4.4 Process Inertia 4.4.3 Underdeveloped Models and Comfort Zone
Total Themes and Subthemes	15	100	

APPENDIX E: THEME EMERGENCE AND SATURATION ANALYSIS

This appendix provides visual evidence of data saturation through two charts illustrating theme emergence across the ten participant interviews.

Theme accumulation was calculated from the matrix coding query exported from NVivo. Firstly, the coding matrix was converted to a binary presence table, where each cell was coded as *"1"* if at least one reference to that theme appeared in the participant's interview, and *"0"* otherwise. Using this binary matrix, a cumulative sum of unique themes was computed across the interview sequence: if a theme had appeared in any previous interview, it was not counted again. This resulted in the cumulative theme accumulation curve shown in Figure 6.

Secondly, the number of newly emergent themes per interview (Figure 7) was calculated by differencing consecutive cumulative values (i.e., Cumulative(i) − Cumulative(i−1)). Together, these charts demonstrate that theme emergence stabilized by the sixth interview, with minimal new themes emerging thereafter.

CC BY

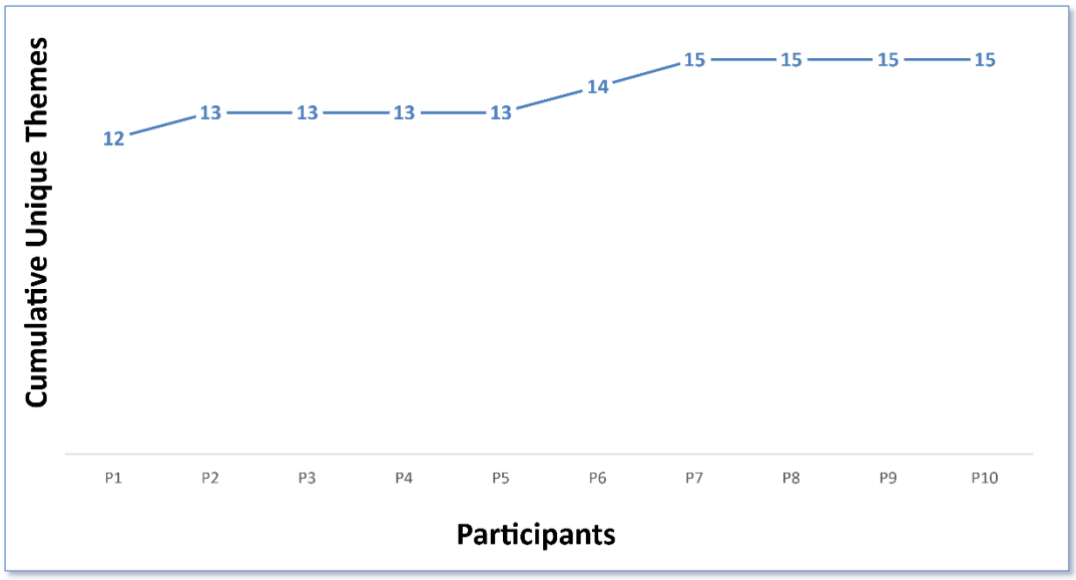

Figure 6: Cumulative unique themes identified across the ten participant interviews.

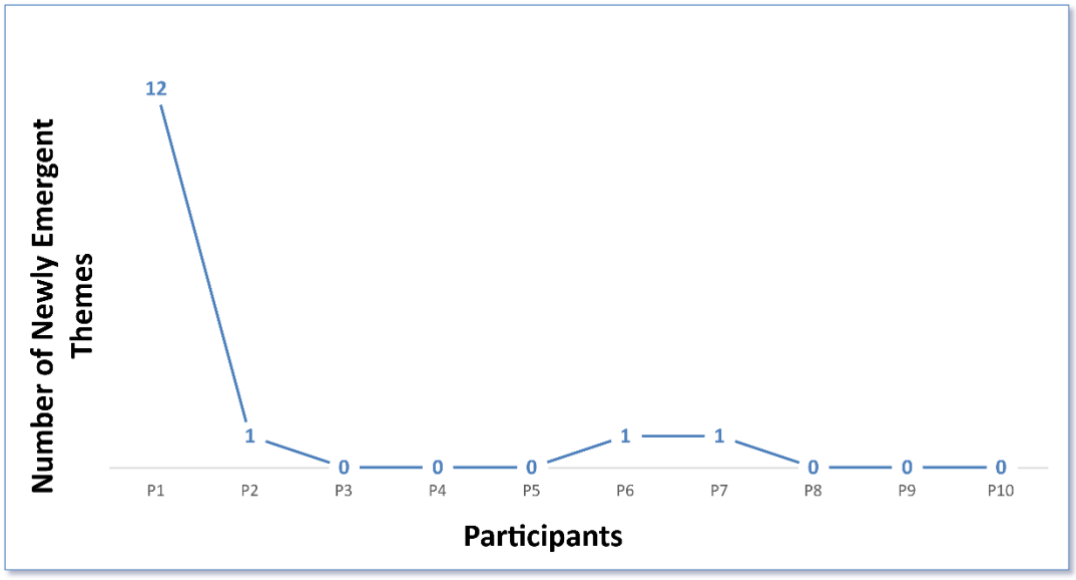

Figure 7: Number of newly emergent themes identified in each participant interview.

COREQ (COnsolidated criteria for REporting Qualitative research) Checklist

A checklist of items that should be included in reports of qualitative research. You must report the page number in your manuscript where you consider each of the items listed in this checklist. If you have not included this information, either revise your manuscript accordingly before submitting or note N/A.

Topic	Item No.	Guide Questions/Description	Reported on Page No.
Domain 1: Research team and reflexivity			
Personal characteristics			
Interviewer/facilitator	1	Which author/s conducted the interview or focus group?	
Credentials	2	What were the researcher's credentials? E.g. PhD, MD	
Occupation	3	What was their occupation at the time of the study?	
Gender	4	Was the researcher male or female?	
Experience and training	5	What experience or training did the researcher have?	
Relationship with participants			
Relationship established	6	Was a relationship established prior to study commencement?	
Participant knowledge of the interviewer	7	What did the participants know about the researcher? e.g. personal goals, reasons for doing the research	
Interviewer characteristics	8	What characteristics were reported about the inter viewer/facilitator? e.g. Bias, assumptions, reasons and interests in the research topic	
Domain 2: Study design			
Theoretical framework			
Methodological orientation and Theory	9	What methodological orientation was stated to underpin the study? e.g. grounded theory, discourse analysis, ethnography, phenomenology, content analysis	
Participant selection			
Sampling	10	How were participants selected? e.g. purposive, convenience, consecutive, snowball	
Method of approach	11	How were participants approached? e.g. face-to-face, telephone, mail, email	
Sample size	12	How many participants were in the study?	
Non-participation	13	How many people refused to participate or dropped out? Reasons?	
Setting			
Setting of data collection	14	Where was the data collected? e.g. home, clinic, workplace	
Presence of non-participants	15	Was anyone else present besides the participants and researchers?	
Description of sample	16	What are the important characteristics of the sample? e.g. demographic data, date	
Data collection			
Interview guide	17	Were questions, prompts, guides provided by the authors? Was it pilot tested?	
Repeat interviews	18	Were repeat inter views carried out? If yes, how many?	
Audio/visual recording	19	Did the research use audio or visual recording to collect the data?	
Field notes	20	Were field notes made during and/or after the inter view or focus group?	
Duration	21	What was the duration of the inter views or focus group?	
Data saturation	22	Was data saturation discussed?	
Transcripts returned	23	Were transcripts returned to participants for comment and/or	

Topic	Item No.	Guide Questions/Description	Reported on Page No.
		correction?	
Domain 3: analysis and findings			
Data analysis			
Number of data coders	24	How many data coders coded the data?	
Description of the coding tree	25	Did authors provide a description of the coding tree?	
Derivation of themes	26	Were themes identified in advance or derived from the data?	
Software	27	What software, if applicable, was used to manage the data?	
Participant checking	28	Did participants provide feedback on the findings?	
Reporting			
Quotations presented	29	Were participant quotations presented to illustrate the themes/findings? Was each quotation identified? e.g. participant number	
Data and findings consistent	30	Was there consistency between the data presented and the findings?	
Clarity of major themes	31	Were major themes clearly presented in the findings?	
Clarity of minor themes	32	Is there a description of diverse cases or discussion of minor themes?	

Developed from: Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*. 2007. Volume 19, Number 6: pp. 349 – 357

Once you have completed this checklist, please save a copy and upload it as part of your submission. DO NOT include this checklist as part of the main manuscript document. It must be uploaded as a separate file.

Themes	Subtheme	Coded Text	Participant ID
Knowledge Gaps	**Misunderstanding about Structured Data, Semi-Structured Data and Unstructured Data**	"Can you give me an example? What do you mean by Unstructured Data?" "I'm not sure what you mean by unstructured data, though what would you mean by that?"	P1
		"When you say unstructured, you mean things like PDFs, minutes, files not in models?"	P2
		"These are highly structured documents and at the various stages." "They usually impose some sort of a common data environment which is the highest structured kind of information that we can possibly share." "What would you consider unstructured data source?" "I'm just not entirely sure how I'm going to differentiate between the structured and unstructured results in this one. Because I'm getting all of them." "how much time we are spending on searching the unstructured data. Because I just generally, I'm not entirely sure what part of the kind of day-to-day searches is relating to the unstructured data we are just trying to minimize that time"	P3
		"I'm actually not even fully knowledgeable on the complexity of it"	P4
		"we consider unstructured data information that is not predefined in a data model or organizational framework, such as	P5

		PDFs, audio files, video images and So,cial media posts." "Within the ISO 19650 we address those problems through naming suitability and provision folder structure and the lifecycle of the assets and through an internal folder structure. And so,… this is the structure that we follow to name every single document, regardless of what it is, regardless of whether it's structured data or unstructured. Through this way an unstructured data becomes structured through the life cycle of the project."	
		"do you suppose adding appropriate file naming conventions makes the data structured? P6: Yeah, exactly" "do you consider a drawing as unstructured data?"	P6
		"I'd say there's unstructured, which is, you know, images, Pdfs documents, and the likes. Of course, it's structured to enable a computer to open the file. But I would call that unstructured. And then you've got structured data. It could be an Excel file. It could be Csv file could be something could be to rdf. but then you've got standardized structured data where it's actually been created and stored in a standardized manner so that a computer can actually recognize what it is in a standardized way. Excuse me, and I think there's a lot of misunderstanding between those 3 levels."	P7

		"There seems to be a perception amongst some stakeholders that just by stipulating use of ISO 19650 parts one and 2 that you're therefore can expect a consistent standardized approach to be applied across all projects."	
		"By unstructured sources you mean not in the model, is it?"	P8
		"Unstructured data, I would define as having no metadata, no container ID, and not being saved in a particular place."	P9
		"The biggest thing is understanding what data you want structured in the first place…Once you know what the data is, you can figure out how to structure it. For example, if you're looking at a PDF, you need to decide what information you want from it. Once you know that, you can restructure it for example, by moving the information into an Excel file, because Excel is easier for AI to read…You can also set up templates for a PDF so that anything in a certain window is extracted. It's important to know where that information is and how to structure it…. Another example: we wanted to extract title block information from drawings. We defined the zone on the drawing where the title text would be, then ran it through a series of drawings. That was AI structuring the data for us….But if you just say, "I want all the information from the drawing," AI won't	P 10

		know what to do. You need to define the information you want and point to where it is…So, the biggest thing is understanding the end goal: what you want AI to help you with and then working backwards to structure your data accordingly."	
	Ambiguity Surrounding BIM	"As I mentioned in some of the correspondence, we don't actually use Revit, BIM as such in the traditional sense of BIM."	P1
		"there's no reason why that proforma template couldn't be in BIM" "I'll be honest. I'm not up to speed with sharing data via IFC files. It's not something that we would share" "it may not always be a software issue; it could be how the architect has drawn the model. There might be a very specific way to draw the model so the thermal software can read it, and if the architect hasn't done it that way, the model won't work."	P2
		"they outsource their BIM deliverables to a technician or someone who's not necessarily the designer. I think that's how they overcome that step. If they're heavily involved in the project, but the one thing they need to deliver at the end is the BIM model, it kind of comes almost as an afterthought, something they produce rather than a design tool. And that's less useful for us, because we don't have it as part of our	P3

		coordination sequence."	
		"the reason why everything was done through PDF, and not through BIM" "within their philosophy, or like nowadays in many companies, BIM is Revit" "sometimes even people involved in the process, they're not fully aware of what it is. You know the process or what the outcome should be and we're dealing with contractor and design team that are learning."	P4
		"They talk to each other, So, information that is not relevant to the BIM model, would it make sense to incorporate it within the BIM model?"	P5
		"there are lots of people within the company who still only consider it as the models. But there are also a lot of key players realizing more and more that it's about the 'I' actually." "Many companies work with IFC; more and more can also handle native file formats. If you're using native files, you don't face as many interoperability issues because you're working within the same system."	P6
		"a lot of people seem to think that IFC is a is a model that you can actually work with and edit. But it's not. It's a file. It's an exchange file format that's intended not to be edited, but to be imported and exported between applications," "Like to be honest, a lot of my colleagues think	P7

		that BIM is just the file naming." "It's a very loosely used term and this kind of means different things to different people. So, we do assess here through our technical competency matrix as part of the annual training reviews. People have their own perception or knowledge of BIM and it does tend to stay quite low relative to other competencies like bridge design or road design or buildings structures design. Even though we do look to train people and the likes, they just still consider that they don't have much knowledge in the area. So, I think a lack of understanding is still a problem."	
		"It's mainly because the structural engineers haven't been upskilled in BIM, and they're not allowed to use it" "Alright. What would be your own understanding of BIM? What do you understand by BIM? P8: Jesus. I don't know." "do you consider all these documents that you produce as part of BIM? P8: That's a good point. I suppose, do they fall? Potentially, some of them. Yeah."	P8
		"For the main contractor, there's still very limited understanding." "BIM is still interpreted as being just a 3D geometry"	P9

		"I prefer to stay as native as possible in the Autodesk environment. That for me, is for very practical purpose." "the concept of IFC is that you export your native model to IFC, and straight away, you have a duplicate. So, you're moving away from single source of truth"	
		"There's probably a small minority who consider BIM just to be the 3D" "there probably are a number of people who would say BIM is just 3D models" "On some of our projects, the design team will refuse to issue us over a model and they're saying, no, you're not getting a model. You're getting the drawings. If you want a model, you need to recreate one yourself, which again, I think is maybe a shift in the industry that definitely needs to happen." "I find IFC models difficult to work with" "In terms of a use case for IFC. I haven't seen any great advantages really at the moment."	P10
	Lack of Understanding About Digital Twins	"When it comes to digital twins and asset information, we have found in the last 2 or 3 years that clients don't understand this, and they don't really want it." "I can't name the company, but we had a large pharmaceutical client. We were building a facility worth over €100 million, and the EIR asked for a full	P6

		digital twin everything captured in the model, COBie, asset information, and so on. We responded that yes, we could do that, but we suggested having an early meeting with the facility management team to understand how they wanted their data structured: what information they actually needed, and for which specific assets. It turned out to be a really good meeting, great people. And what they actually wanted was an Excel sheet. “	
		“I’m pitching that in the company as our digital twin platform” “that message is starting to sink in at our company, what the digital twin actually means to the client.”	P9
		“I don’t believe many clients are using the full model. From my experience, they take the model for visuals, but all FM is managed through Excel.”	P10
	Lack of Understanding about Linked Data	“Linked Data is very powerful, and the use of ontologies is very powerful. …But there aren't that many people who have expertise in it. And it is twice specialist, and the major software vendors haven't really adopted it and don’t seem to be particularly willing to implement it in their systems. So, I think while in theory, it's potentially very useful and while it's been shown to be useful on at a small scale, certain use cases. I'm not yet convinced it'll be adopted on a broader	P7

		scale." "The proof of concept isn't strong enough yet, in my opinion, to insist that it's implemented at a national level"	
		"Yeah. Are you aware of ontologies or the concept of linked data? It's a hot topic in academia, and they're trying to address the same challenge by making a knowledge graph and connecting data across the web. P10: Are you? R1: We have that. P10: I'm not too familiar with it."	P10
Stakeholder Dynamics		"the builders that you're engaging with, who aren't top tier builders with quantity surveyors on their team and IT managers and document controllers, they're operating largely by themselves and they're operating largely over their phone even at that scale. So, you're trying to get the information to them in the quickest, cleanest way and that's still email and phone calls, really" "have originally worked in larger offices and took a lot of the kind of discipline of heavily written documentation of large projects and government projects down with me to work on small residential projects. And my experience was that, the level of administration that I was bringing to the project was actually off putting for the contractors, and was	P1

		leading to poorer results, not better results. It was scaring them off. It was making them guarded because they were seeing this flurry of emails and confirmations of instructions. And I found that I would get better results if I kind of held back a little and wasn't quite so bureaucratic over the project." "we're doing a very large house in Cunningham Road in Dalkey. It's a conservation project. By and large, we're dealing with a conservation joiner, specialist joiner who takes away sash windows and prepares and repairs sash windows. And he doesn't do email. He told me today; he doesn't do email." "the way we disseminate information or information will be issued, will be driven by generally the biggest player in the design team. So, in our projects that, actually, funnily enough, might often be the structural engineer tends to have a bigger office and working on a larger scale of projects" "But even those guys have to shift to suit the small-scale builders like the joinery guy that I just mentioned, because they realize that you can't apply the same techniques with small scale builders that you might apply it when you're working with, say, ABC or XYZ construction, or someone like that."	
		"it's the stakeholder. It	P2

		depends on the stakeholder and client requirements. Certainly, a lot of clients now have their own specifications and requirements for common data environments. Particularly, one large infrastructure client we're working with at the moment, all the projects are delivered using ACC (Autodesk Construction Cloud). So, you would have maybe smaller clients who don't have any specific requirements." "For that particular infrastructure client, they host all the common data environments they own and host them. Whereas we might have other projects where clients don't have any specific requirements, and someone in the design team may host the common data environment. Or the third case is where the client has absolutely no requirement for a common data environment, and the design team would choose to set one up because it facilitates design collaboration. So really, it depends on the client and what the client is asking for." "it's much easier when a client comes to us and says, "This is what we want, and this is how we want it." And oftentimes, we have. Yeah, and it's better if the client can host the common data environment, we find, because we often have restrictions on who we can share information with. There can also be	

		quite a bit of time involved in administering those common data environments. So, unless we are specifically contracted to do it, we would prefer not to. It certainly is easier when a client comes to us with a prescriptive way of working and specific software they would like us to use." "there are two types of clients. Some, like a university client we worked with, are very up to speed. In that case, they dictate to us the Uniclass codes, the asset information models, and they have very specific standards, that's great. But in other cases, where the client has less understanding, they may not follow a full ISO standard but have specific requirements. We try to work with the client to figure out what they want to achieve and align what we produce to their needs. Because if the client doesn't have the capability to interpret the data, there's no point in providing it in a format they can't use."	
		"I think we would also have the client would give us their information management plan, which might set out how they'd like to receive deliverables, and that they want us to follow a certain file naming procedure. So, for example, the XYZ have their own that's specific to the ABC campus and following certain codes. But in general it follows	P3

		the ISO standard. So that would be something that would come from the clients in how they want us to name our documents. And how they would like us to revision and put suitability codes and things like that onto our document." "I think the Irish Government's requirements and the government contracts are on the opposite side of that. They are very structured and very formalised. I don't think we would be able to get away with a government client being contacted mostly by phone with no record of what was discussed. None of the public contracts would be delivered this way. I would say CWMF contracts, they are extremely formalised, and the deliverables are very clear there." "because Autodesk seems to have a monopoly on that market" "We're probably working with six different platforms right now, various CDEs. Each one is different; each one has different workflows and requirements. So, what we end up with is maintaining our own trackers to get a better overview of what's happening in these external CDEs. We provide and share the information through the CDE, we receive information through the CDE, but we still rely on our own internal tools to understand what's happening in each of	

		them."	
		"stakeholder or client. They might have a requirement in relation to how the information has to be issued to them, but, as I mentioned before, often relates more to file naming convention which we would follow," "We follow what we've been requested by XYZ (client)" "what we provided is what we've been asked by the client at the time of signing the contract with the contractor"	P4
		"Some of these problems that we're facing are through the linking between the model and the CDE, So, if you're using Software that is not by Autodesk, and you have a CDE that is not regulated by Autodesk then you end up facing these kinds of issues."	P5
		"When it comes to digital twins and asset information, we have found in the last 2 or 3 years that clients don't understand this, and they don't really want it." "I can't name the company, but we had a large pharmaceutical client. We were building a facility worth over €100 million, and the EIR asked for a full digital twin everything captured in the model, COBie, asset information, and so on. We responded that yes, we could do that, but we suggested having an early meeting with the facility management team to understand how they wanted their data	P6

		structured: what information they actually needed, and for which specific assets. It turned out to be a really good meeting, great people. And what they actually wanted was an Excel sheet. They wanted ten parameters mapped out, an asset list down the side, and data captured in a simple spreadsheet format." "We do get requests like this occasionally, but they typically come from the design team, not the client. And when we actually speak to the client, they often want something slightly different or maybe not less, but definitely different from what's been put in the EIR." "you're saying there's some ambiguity between what the client wants and what the designer proposes? P6: Yeah, there definitely can be." "Using that pharma example, I'm reading between the lines here, but I'd imagine the design team had a conversation with the client, maybe at the pretender stage, and said, "One of your options is to capture asset information and COBie," and probably explained what that meant. The client would've said, "Yes, that sounds great, we'd like that." So, the design team put that requirement into the EIR and the tender documents. When we read that, we think, "Okay, no problem, we can deliver	

		COBie and structured data." But we know we need an early meeting with the client to clarify exactly what they want. Because a couple of vague lines in an EIR can lead to us spending months gathering all sorts of information only to find we've got the wrong stuff, missed something, or given them too much." "The manager had a very interesting way of describing what assets he was interested in. He said to me, 'Imagine you're a giant, and you turn up beside that building. You pick up the whole building, turn it upside down, and shake it. Anything that falls out? That's what I'm interested in. Anything that stays stuck to the shell and the walls? Not interested.' That was a really interesting way to picture what he was looking for." "On interoperability more broadly yes, it's always a major discussion with vendors and software suppliers. Many companies work with IFC; more and more can also handle native file formats. If you're using native files, you don't face as many interoperability issues because you're working within the same system." "So, in this entire collaborative project, which stakeholders do you think influence the production of unstructured data the most? Or it could be multiple stakeholders. P6: Everybody, to a	

		certain extent. I mean, if you're talking about, do you consider a drawing as unstructured data? R1: Yes. P6: Yeah. So basically, everyone. The design team will issue their design drawings along with the model. The model is also often shared for information purposes only. The subcontractor team and the main contractor team will take those models and then further coordinate them with specific information maybe replacing dampers with the actual dampers that are going to be installed on site. Often the AHU is shown as a box. We'll update it, showing all the connections and level of detail, moving through the stages. There are drawings produced on the back of all of that. There are technical subs produced on the back of all of that. There are meetings that are minuted. And on the back of all of that, it's all unstructured data, really. So, it's the entire team, really."	
		"clients don't typically specify standards-based exchange. So even if it's structured data within models, it's not necessarily semantically rich. It's not necessarily standardized with internationally recognized classification or anything like that." "Then you've got the public clients themselves. Whether it's transport infrastructure, Ireland or the Opw. HSE	P7

		or local authorities. County councils. They influence it because whether they implement those requirements within their procurement contracts is, is going to influence what the supply chain does." "the supply chain never really believed that the client would actually implement it. So, they then will say that they'll do something, but when it actually comes to the project they don't," "where there's a business need and a market force that dictate stuff, structured exchange of structured information is valuable to the supply chain. Then they'll go ahead and do it" "some of the clients are concerned that the market isn't ready for this. but in reality, a lot of the market has been doing this for years" "The final big stakeholder is the software companies themselves because if they don't implement something, then it's difficult."	
		"Typically, our clients are big pharmaceutical companies with a lot of in house expertise. For example, one of them that I'm working with has very clear requirements, a very good idea of what they want, and a very incisive review process. They know what they want, and they know what they're doing. Other clients also in the pharmaceutical sector might not have been	P8

		around as long or aren't as big, so they don't know what they need in the same way. It can really vary." "The life sciences are very good that they know what they want, and it's very clear, and they've done it before…. whereas when I'd be working in more typical construction, you know, commercial or residential and stuff, and the clients tended to not really know what they wanted."	
		"For a long time, they didn't know what they wanted. BIM, for example, a lot of times, they would almost subcontract out the responsibility to a project manager because the clients themselves would not be aware of how to structure their information needs." "The architects would revert to industry standard documents, which were quite theoretical and idealistic, and the technology and software weren't fully there to support being able to deliver projects to this level. But the architects would prescribe this high level anyway, because that was the standard in the documents, and they were just ticking the box, I guess, for the client without any real understanding of what the value was behind it." "But technology, I think, is the big driver behind it. Once the technology makes it accessible, it'll almost force users to use	P9

		metadata” “it's primarily from the appointing party, the client. They're supposed to lay out the whole process, the whole delivery. Everyone is supposed to follow the appointing party. The problem is that they don’t understand it fully yet, so they outsource it to specialists. But again, that’s starting to change now.” “We won't do it until we have to do it, until there is a requirement there from the client side.”	
		“Some projects have nice, clean-cut client requirements, and on those, it’s easier to understand because everything is well-documented. A lot of times, it needs to be driven by the client. If the client doesn’t care about information management or the production information on their project, it can lead to design teams or subcontractors being more relaxed with their deliverables” “The client might say that they want a federated as-built model at the end of the project, but they don't know what requirements or what assets they want tagged in there. For us, that can be difficult to try to understand what the client needs and what they want to maintain within the building” “A lot of times, clients are unaware of what they actually need, so they ask their design team for an answer. Often, I feel the design	P10

		team says, "Give us everything," in case they leave something out, and the client ends up needing it. They're trying to cover themselves by saying, "Just give us everything," instead of saying, "Here are the 10 items we want in the building." "I think it's more of an issue where the design team themselves maybe don't know what the client wants either, and the kind of understanding of it is a bit vague sort of like, okay, just give me all the information, and it's a failsafe, and that if we receive everything over, well, then no one can do any wrong if that makes sense like, you know. Otherwise, if they only specified certain elements to us. Well, then, all of a sudden, if we deliver it at the end. The client might say, okay, well, we're missing X parameter from the model." the design team can say, "We were never asked for that." "Some clients are like, 'Oh, well, if they got an as-built model on their project, I kind of want one too,' but their own systems aren't set up to use a model. I think they'll still end up looking for the other documentation." "Even on some of our projects, the design team refuses to issue a model. They say, "No, you're not getting a model; you're getting the drawings. If you want a model, you need to recreate it yourself." I think that's a shift the industry needs to make,	

		but we're not there yet." "In terms of use cases for IFC, I haven't seen great advantages at the moment. It would probably need to be driven from the client side maybe if they want to integrate it with their FM system."	
Single Source of Truth Dilemma and ISO 19650		"But where there is a common data environment, I think there's still a bit of legacy that people like to follow things up with emails. Even if there is something on a common data environment where there's a comment assigned or, you know, I think there's still a bit of legacy where people like to follow that up with emails. And it's probably seen as a more formal way of doing it."	P2
		"this is the main problem with CDEs, there are so many of them, and we are asked to work with so many different CDEs. Almost every client has their own preference for that." "even within one project, you could have a client-provided CDE where you upload all of your documents for approval for the client to review, then a separate CDE for internal design team use, and then, on one project, when it got to construction stage, it was the contractor's responsibility to provide another CDE where they'd put all their contractor submittals and RFI responses. That was another thing entirely. It becomes quite unwieldy. It's	P3

		unwieldy because you have different members of the project team coming and going, needing logins to access information, and, as P3a says, tracking everything that's coming in." "We are currently working with about six different platforms, various CDEs, each with different workflows and requirements. As a result, we've created our own tracking systems to gain a better overview of what's happening across these external CDEs. While we provide and share information through the CDEs, we still rely on our own internal tools to understand the activities occurring in each of them."	
		"The more CDEs, the more problems you end up facing." "we are obliged to use a CDE that is imposed by the client and the Resources in that CDE are limited to implement these standards that I just showed you. So, at the moment we are using OB and all the capabilities that we have in BIM 360 are not on OB. Some of these problems that we're facing are through the linking between the model and the CDE, So, if you're using Software that is not by Autodesk, and you have a CDE that is not regulated by Autodesk then you end up facing these kinds of issues."	P5
		"Our main ones really are BEP, TIDP, MIDP"	P6

		“we have a document manager dedicated to developing those documents."	
		“Different CDE, software providers implement those naming conventions in different ways within their applications. It still tends to be quite a lot of human effort to manage exchange of those files between applications, the managing of metadata, revisioning status codes, all those sorts of things that really frustrates a lot of people. Where we've tried to be. We've tried to standardize it, but we've actually probably made things a bit more complicated than they need to be.” “you can lose metadata along the way. You know, in in one CDE. You might have all of the approval history for a particular package of deliverables. You might have all the, you know the who’s approved what and when, whereas when you then pull it out of one system and put it into another one, you might lose all of that history because it doesn't get carried along with it. The provenance metadata may not go along with it. Even the revisions might have to be manually inputted in the second system, because there's no way necessarily of integrating between the 2.”	P7
		“Usually, we also follow up and send stuff directly to people by email, because it gets lost in the common data	P8

		process. You end up having thousands of documents on these things, and people don't read the notifications they get from their software. Then they say they haven't seen the information or haven't reviewed it. So, you end up saying, "Oh, yeah, this has been issued on that software. And here it is again to make sure you've got it." You often end up sending things in duplicate."	
		"If I upload a drawing to Procore, it requests metadata and defines what the document is, which works fine. But when someone who doesn't have access to Procore needs a copy, we download and email it to them, resulting in all that metadata being lost." "It's too complicated, too verbose, too standard-based, and too theoretical. It's full of red tape and filler, which turns people off immediately. So, what we tend to do is extract the essential parts and condense them into cheat sheets."	P9
		"Some CDEs, are kind of like a file-based approach, and others are kind of like a folder based. And so, to give, maybe, like an example, ACC will be kind of a folder based CDE, where permissions are granted based on folders, and as a result you could have one document in maybe 2 or 3 different folders… We would prefer to have one container id for a file and its whole life cycle on the project. It	P10

		sits under that one container, so you can see when it was uploaded, when it was reviewed, when it was approved, when it was then maybe revised at a later date. And that's what we kind of prefer. But some CDEs don't allow that and they would prefer based off a folder permission." "So, if ACC is based on the folder type, doesn't that go against the "single source of truth" idea? Because you can have multiple files in different places? P10: Correct. Yeah, yeah, yeah." "One of the solutions the software providers gave us was: upload them into one folder, then copy them into another folder once they go through an approval workflow. But again, we're like, now you're losing the activity log of that file. It's now a brand-new file, and suddenly a project admin or a senior person who has access to multiple folders might see that drawing multiple times. There's a risk of them using the wrong drawing or reviewing the wrong drawing." "It's just recreating the same documents again and again, which can be quite tedious." "BCAR. Okay, so typically, a lot of the time that's managed on the client's CDE. They usually request their own CDE for that, which sometimes adds extra work for us. We have to keep a copy of the records ourselves, so	

		we upload them to our CDE in a location we can always access. Then we also upload them to the client's BCAR CDE. So, there's a bit of duplication in that process."	

Process Inertia	**Dependence on Emails**	"you're trying to get the information to them in the quickest, cleanest way and that's still email and phone calls" "But when important information has been communicated by telephone. We would follow up with an email confirming the conversation. That's just happened. This is specifically for domestic projects specifically related to something to do with cost, health, and safety."	P1
		"Email is probably still our main method of communication between design teams and between other members of our project team." "instant messaging via Teams is probably replacing email for a lot of internal communications. And my personal view is that on Teams it can be more difficult to track and search items. And so, I prefer emails for any key information that needs to be sent to people" "I think, there's still a bit of legacy that people like to follow things up with emails. Even if there is something on a common data environment where there's a comment assigned. Or you know,	P2

		I think there's still a bit of legacy where people like to follow that up with emails. And it's probably seen as a more formal way of doing it." "I suppose the main benefit of those is that if that person leaves that, you always have access to certain emails."	
		"because a lot of information is being shared through email, we have a record of that. We have the information and the exchange happening there. But it wouldn't be like a regular structured thing. This is just the ongoing life of the project. There isn't a dedicated person who is writing a short summary on a regular basis" "there may be an odd confidential correspondence that's not being shared with the full team, but normally everyone who is in the distribution group gets a copy of that email"	P3
		"Anything that is via phone call is then maybe is kind of communication tool. But then, we follow up with an email and just saying, like, whatever you know, please, confirm, or as per our phone call conversation, we know to that, you know." "There are kind of a number of email. Yeah, yeah. And we have a system. We save them all and attachment on the server. They are all saved by date, so it's very easy to go back and keep a record or looked at the information that	P4

		has been exchanged." "we don't have personal email in the company. The emails are by project. So, everybody who works in a project access that email. Or if you work in multiple project, you have access on multiple emails."	
		"And how often do you communicate the project information through emails? P5: Daily." "Okay, and what kind of information? P5: Would be any kind of information that we deem relevant to the project at the stage we're on. So, for example, drawings, instructions and data sheets, trackers." "I update to the CDE, upload it to the CDE and then share it via email through the CDE in most cases." "When it's critical data, I try to share through the CDE. When it's non-critical data, I try to communicate it through email."	P5
		"We do sometimes share information informally by email. For example, say we're coordinating a building with specific steel that has to be painted, and there's a clearance requirement maybe 50 mm. If the design is still being worked out as we're coordinating, the designers might send us a sketch showing which steel needs painting and which doesn't. That helps us keep working. But that's informal. It's not a contract document, it still needs to go through the CDE and	P6

		formal approval processes." "I don't think emails will ever become obsolete, not completely." "What I don't like is when information is shared off to the side, someone sends an email with an Excel sheet, then forwards the email but the attachment is missing, or someone else sends a different email entirely, and now we're all looking at different information."	
		"They are then shared with 3rd parties, so that could be through email, in which case they're duplicated within email systems in anyone, in whoever was at the recipient or sender of those emails."	P7
		"the client themselves within their organization, they might just send emails or sketches and things like that for ideas they have." "So, are you sharing the same information through more than one channel, like you're sharing a report both by email and through the CDE? Do you do that? P8: When you say CDE, what's that? Sorry? R1: Common Data Environment. P8: Yes, yes." "Then it goes through a formal review process. Usually, we also follow up and send stuff directly to people by email, because it gets lost in the common data process."	P8

		“Concepts, changes, and things like that tend to go across via sketches, emails, and so on.”	
		“All the time. That's a very big problem in the industry. For example, if the architect creates a BIM execution plan and EIR on behalf of the client, and they email it to us and copy nine people on it, that's 10 copies now that exist of that BIM execution plan” “The same applies for any deliverable: drawings, documents, schedules, specifications, reports. “Once they're emailed. They're being duplicated. And that still is kind of a bad habit of the industry” “if I upload a drawing to Procore, it'll ask for metadata, and it'll define what it is, and that's fine. But then somebody wants a copy of it, they're not on Procore. So, we download it, email it to them, and all that metadata is lost.” “A lot of our smaller subcontractors are just two men in a van, fitting a floor finish or whatever. They won't have a clue about all of this, they'll just email us the drawings, if they have any, and they'll work away like that. So, we have to determine where the boundaries are.” “It’s all emailed around rather than being shared in the common data environment. That happens all the time, in every discipline and sector. There’s a lot of firefighting involved for	P9

		proper information management, and it happens everywhere." "Email, I guess, is a big culprit for that. It's the default go to for how people distribute information."	
	Unstructured Data Usage across Project Stages	"In the residential sector, it tends to be a booklet presenting the general themes and ideas of the project generated from a 3D Model. And it will have written Doc information about the brief, the budget, and the design strategy." "produce planning documentation, which are drawings submitted obviously for securing planning permission for the project, for statutory approvals, fire safety certificates, disabled access certificates." "After that then, we're creating a suite of drawings, mostly for the purposes of pricing and tendering the built projects of obtaining costs for the built project and those drawings are supplemented with written specifications for the most part." "And then in a normal project, the other type of information that we would submit, prepare documents that we prepare would be as meeting minutes. Generally, the site meeting minutes, client meeting, minutes, site visit records. So, when visiting site, observing any issues to do with quality control or progress on site for the purposes of monitoring the work on site."	P1

		“I suppose, the other type of documents we do would be certification. So, it would be in related to quality control and compliance with building regulations. Those are broadly, I suppose, the types of documents we most, and there will be others, and that haven't come to mind. But broadly. That's those are the documents that we issue.”	
		“And so, at stage one, we would usually just be providing reports, maybe with some sketches to support those. They would be in PDF format, and that would form part of a stage gate approval to the client to proceed to the next stage.” “when we start to enter Stage 2, we would start to produce our actual design deliverables. So that would be our drawings and PDF, our schematics in PDF.” “again, we would produce another report at Stage 2A which would again form another element of stage gate approval onto the next stage. And then in Stage 2B, we would further develop those design deliverables, the drawings, the schematics, and also start to produce our schematics and schedules.” “at the end, our tender, our design deliverables at the end of the design process would be our specifications, our schedules, our	P2

		drawings, our layout drawings, and our schematic drawings." "at the end of the design process, we would also produce a final stage report again for stage gate approval. I suppose that's particularly for public sector clients. For private clients, we may not have the same level of reporting in terms of stage gates, but fundamentally, the design deliverables would generally be the same." "when it gets into construction stage, really the drawings are handed over to the contractor who would then develop them on. So, we have less formal deliverables at that stage. Really, our formal deliverables during construction and during handover would be reports, certification and snag lists, things like that. Usually, we might have some element of updates to our design deliverables, but no new , the bulk of our deliverables are produced at that stage, and it's really just reporting and snagging and close out." "at the end of, say, for life safety systems, emergency lighting and fire detection and alarm systems, we would have to provide a design certificate at the end of the project to certify the design of the system, which then gets packaged up with the commissioning cert as part of the handover certification document. And if it's a BCAR project we will have to produce ancillary	

		certificates, certificates of design and inspection which would form part of the BCAR handover. And if it's not a BCAR project, we would produce an opinion of compliance with the building regulations. So, it's and then we may have to certify payment applications during the, depending on the project manager on the project, we may have to certify payment applications. So, we do have a pro forma for monthly or periodic inspections, to show the value of the works that are completed so that can form part of the payment applications."	
		"A lot of presentation materials at the early stages that might be quite well developed, like visuals and things that are going to sell the project and based on images." "We still have people who are drawing manually, so sketches or watercolours will be generated at that stage as well usually scanned and digitalized." "So that could be anything that could be minutes that could be brief documentation reports definitely reports being sent to the client, inputs from the consultants, various consultants." "we would have like issues trackers, both for internal use for ourselves, or they could be shared with other consultants. We would also consider spreadsheets generally a	P3

		text document as well, because these are not calculation spreadsheets as such. They contain a lot of material like written material." "we would have like issues trackers, both for internal use for ourselves, or they could be shared with other consultants. We would also consider spreadsheets generally a text document as well, because these are not calculation spreadsheets as such. They contain a lot of material like written material." "What are other options? What could be the possible alternative to that? If not file based systems? What else could we be working with" "Keeping Pdfs so far seem to be the most robust way of keeping the project records for prolonged periods of time." "we would have our core set of drawings, in particular the GA drawings, that get developed and continue all the way through. From design development, you'd still have the same ground floor plan with the same drawing number that continues all the way through—goes to planning, goes as part of the tender package, then through to construction, and would probably have an as-built version as well. So, there would be some drawings that have continuity all the way through. But then you would have a huge amount of extra drawings for tender	

		stakeholders.”	
		“At the very, very early stage, we do sketches,” “we have external Excel document Word document. We use per presentation a lot. I mean, whatever the presentation is made via a PowerPoint Google drive. They're often then used in a PDF format, and what else, I guess would be the main one” “it depends on which stage, there are also reports that they need to be issued, which are all, but usually they are prepared via Word, and then exported in a PDF format. You have your minutes, you have your in like your Excel document to record drawing issue sheets, sketch registers, you know, is everything which forms part of the formal issue of document, or the internal advance of a project” “You know there are lots of also sketches there scanned in and put into the server.” “if you take the BCAR process. You know there is a large number of documents that they need to be produced throughout the construction stage, for whatever you know, stage applies of a project, and then at completion.” “another number of documents that are requested by the Department of ABC, and which are stage based, including cost plans and reports so we issue them but in the case, for example, of the	P4

		XYZ, because XYZ was our client, we issue them to them, and they pass them to the department." "Usually we issue all of them as a PDF format, and that's the way also the department like to receive them with the correct file naming."	
		"So, for example, for an animation file the code would be AF, for a method statement would be MS. For a presentation, So, for example, for this presentation would be PP, and here you have the roles that usually take part in a project, and it's also, defined by a set list in the ISO19650 standard." "the date for the pouring of the concrete is recorded through a photo from my booklet on site"	P5
		"we have a document, it's an appendix to our BEP, which would be the Asset Information Model (AIM) for handover." "We have documents on BIM meeting summaries, IT solutions, some standards, minimum quality and assurance checks" "our main ones really are the BEP, TIDP, MIDP, which gets built on the back of that, and then various appendices depending on the project. So, like that AIM would be one, but not all projects have an Asset Information Handover requirement, you know." "like a Communications	P6

		Protocol which I didn't mention before but is another document we'd be looking at that. The Communications Protocol would have our naming convention in it." "A lot of our projects they won't specify a CDE. So, in our tender submission, we propose our CDE. This is done by the tendering department, I don't have great visibility of it but they would. There's a whole series of documents outlining the naming convention, how we use it, the security, where the information is stored online, what servers are used, etc. We'd propose all of that at tender." "Often the AHU is shown as a box. We'll update it, showing all the connections and level of detail, moving through the stages. There are drawings produced on the back of all of that. There are technical subs produced on the back of all of that. There are meetings that are minuted. And on the back of all of that, it's all unstructured data, really. So, it's the entire team, really."	
		"obviously at the initial stages the public clients will need to consider what the what they're trying to do in terms of a project. What are they trying to achieve? What problem are they trying to solve? And then come up with initial concepts for a solution. Likely they will then engage external parties, consultants, designers, and the likes, so they'll	P7

		need to procure those they need to set out the scope for those services need to tender those. So that's all going to be you know, documents, word-based documents, text-based documents, possibly some drawings maps and the likes typically Pdf, based drawings." "It might be photographs, aerial imagery or reports, input from 3rd parties, public consultation documents and the likes route selection reports, environmental impact assessment study reports, all the way through then planning." "you'll continue to have lots of drawings, some models sitting behind those drawings. But it wouldn't typically be the case that those models are provided to clients per se. and then all the way through to detailed design." "typically, the outputs are both models and drawings, to inform contract documents, to procure tenders, or in design, and build contracts, to provide information to the client, to enable their approval of the designs prior to construction. And then all the way through construction." "Then there's a lot of unstructured information exchanged between various parties, whether it's data sheets for pABCucts that are being proposed by a contractor, or queries that are being raised by a contractor and responses from designers and clients,	

		contract administration documents, photographs. You know, site inspections, quality inspection results." "then through to handover, which is typically a very challenging phase of a project for public clients who are trying to manage asset information where they will receive a huge amount of information from a contractor as part of the handover set, whether it's in a safety file in relation to the health and safety Legislation obligations or asset information management or asset information which is to go into an asset information management system. clients in Ireland typically don't have strict rules in relation to that type of information handover. So, a lot of it is unstructured data."	
		"we create sets of drawings, and we'd create specification drawings. Specifications would be the major part in terms of getting the building built." "There are other things like reports that are developed as well. So, reports, drawings, and specifications." "Drawings obviously convey the design to everyone for coordination, costing, and then construction." "Like sketches and things like that and markups? Yes, yes. So, we would do a lot of draft sketches and	P8

		markups, probably for BIM designers. We would like changes and things would often go on a PDF markup and then be changed in the model." "depending on the scale of the change whether it's an issue or something to save money. It's usually brought up in a meeting or on site and then formally raised in an RFI. We respond to the RFI, and then it has to be approved by the client and their costing team. R1: Is the RFI a manual document, or is it handled through a BIM software? P8: In all my previous work, it's always been a manual document—written up on a template."	
		"for a tender, we would receive a tender pack, and that would contain shared resources: documents, drawings, reports, schedules, specifications, and contract particulars as well" "including ISO 19650 formatted documents. And that's quite a substantially different format, including exchange information requirements, level of information need, and different maintainable asset requirements like COBie" "about the BCAR process, how do you handle fire safety, lighting systems, and certifications required for building regulations? Do you link them in the common data environment, or is	P9

		that separate? P9: It's mostly separate. We usually hire a third party to manage the BCAR aspect. We have in-house quality, health, and safety teams. We run inspection checklists and an inspection test plan. We also do internal drawing audits, update drawings to "as built," and track observations, submittals, and RFIs through the common data environment. We have BCAR trackers for all deliverables, reports, and inspections, but the third party manages that."	
		"I suppose the purpose of all the documents issued, from a design team to a contractor, is to give over a design intent. Then, we review the documents and figure out what we need to build to meet our client demands" "That could include certain performance aspects of the building whether it's to meet XYZ's requirements, or certain LEED, GOLD, or BREEAM standards. For us, it's about understanding what we need to build, reading the documents, and making our proposal." "The documents from the client side typically come in one document like the AIR, which would ideally have a maintainable asset list and the parameters they want populated. But the actual information to populate it can come from a wide range of sources, technical submittals, drawings, maybe even a list of	P10

		RFIs raised against that asset. That information can come from many places, yeah."	
	Underdeveloped Models and Comfort Zone	"I was spending huge amount of time constructing the model, the model in a very crude, a very crude model because it wasn't really lending itself to do the fine detail. And then I was using SketchUp, actually for 3D Modelling and making huge amounts of fine details in SketchUp that I couldn't do in Revit and was using 2D line work to create all the detail. I thought this is a total waste of time, because I'm creating a model that is taking a long time to create and doesn't have any of the detail or the finesse that I'm looking for and I'm drawing on the finesse and the detail in 2D." "If this office wants to properly engage in public sector work, right? We have to do BIM right. I'm under no illusions about that whatsoever. And that's one of the things putting me off getting into public sector work. Much more is that I find my job less enjoyable using BIM."	P1
		"I think it remains dominant because that's what the industry is used to. The industry is slow to change. If you look at the different aspects of the design and estimating processes, they're all built around file sharing and PDFs. That's just how it's been done. There's no doubt it's moving forward, but it's slow." "To be fair, it may not	P2

		always be a software issue; it could be how the architect has drawn the model. There might be a very specific way to draw the model so the thermal software can read it, and if the architect hasn't done it that way, the model won't work."	
		"Because things get quite contractual for us when a project is on site. It's much better to be able to refer to a Pdf. That shows our exact design intent rather than a model that would have to be completely perfect in every single junction, rather than just having like one drawing." "I think we're still working on our models to a scale of about 1:20 or 1:10, but not down to the level of detail that we might actually issue on a detailed drawing that shows intent, shows a typical interface that might applied in multiple scenarios. And that might be 1: 5 or 1: 10, but like it would have more detail in it than we're actually modelling. Like when it comes to membranes and vapor barriers and things like that. That would actually be very difficult to include in a model file."	P3
		"the reason why everything was done through PDF, and not through BIM is that obviously, when you, when a contractor finds information which is not fully coordinated within the design team, there is an opportunity to raise an issue and claim, and for BIM to be fully coordinated is a	P4

		different type of work and depending to the project than to issue in a PDF, obviously, you're working in 3D. And sometimes in 3D you can find discrepancy that 2D they don't align."	
		"Facilities manager is not going to go off into a model. He doesn't use models. Normally. he's not going to go in there and start sectioning a model and navigating around and finding the floor and clicking on it and finding his information. He's not going to do that. He's going to go to his own manual or his facilities management software where he has his data structure the way he wants it…. He's gonna find it 20 times quicker than going off and looking in the model." "It's almost a generational thing that you need to move through. I don't think you can flick a switch and change an engineer's mindset, especially someone who's under pressure on site to meet a deadline and get a building built on schedule with no delays. You can't get him to change his whole way of looking at things. He's used to getting a PDF drawn. He's used to opening up these text files and finding it." "You can't just suddenly throw everything you've done for the last 40 years out the window and start fresh tomorrow, it can't work like that. The whole thing would fall down, from my point of view."	P6

		"Because BIM is still regarded as a cause of extra cost and over complication. The people in construction companies who tend to make the decisions about software and approach tend to be more senior members, directors, and they're accustomed to working the old way. They lean on their experience, and they look backward on their 30 years of experience as opposed to 3 years in the future" "The models are often seen as risky because they're not contractual deliverables, and there's a lot of inaccurate information in them."	P9
		"I don't want to be picking on a design team. But design team and subcontractors, they don't want to produce the models to a good enough level to be able to track and incorporate the Building Control Amendments Regulations (BCAR) information on. And there's still like a lot of elements that are not modelled or only modelled in 2D. And again, we're just coming off a project here, there was deliverable for a design team to produce a model in 3D. But they've only ended up producing it in 2D. And now we need to figure out the rest of the coordination, and they said that they may produce a model at the end of the project and give it over to us. But again, that's kind of defeating the whole purpose of our coordination and of our	P10

		exercises." "People can't produce high quality models quickly enough because projects are always under time pressure. The attitude in the industry is often, "Let's just get something into the model and move on." "If clients want to use the model for FM, or if they want to rely on it, then yes, there's definitely value. But if they're going to use Excel spreadsheets instead, there's no point. Producing accurate models takes time, picking the right materials and all that." "I think it comes down to the contract. On all of our projects, no matter how good our models are, it ultimately comes down to the drawing as the contract document. The reason for that is, it's what the design team and subcontractors are comfortable standing over."	
	Skills Gap	"I don't see why using a Revit model is going to result in a reduced construction cost and arguably it might reduce the opportunity for errors, and therefore cost arguably, but that would rely on a whole team being extremely proficient on that platform."	P1
		"probably the main constraint is access to the software and the user's ability to use that software." "I'm not up to speed with sharing data via IFC files" "It's a knowledge gap	P2

		and an expertise gap as opposed to any technical reason." "To be fair, it may not always be a software issue; it could be how the architect has drawn the model. There might be a very specific way to draw the model so the thermal software can read it, and if the architect hasn't done it that way, the model won't work."	
		"We would also consider spreadsheets generally a text document as well, because these are not calculation spreadsheets as such. They contain a lot of material like written material."	P3
		"Even the contractors that we've been working with, they had to go in training. They didn't have any knowledge of asset register or COBie."	P4
		"When I was trying to onboard this person in there, I could tell he wasn't understanding what I was talking about. So, I tried to look at it from a different angle, and I said: Look, how do you produce your drawings, your fabrication drawings, so that your team know how long to cut a piece of pipe and what angle it should be used. And he said, well, typically, I'll take the Pdf that's given to me by the designers and then I'll sit down on the drawing board with a T-square and a set square and a pencil, and I'll draw out my drawing. And I couldn't believe it, that company was still using pen and paper to get their	P6

		drawings together."	
		"linked Data is very powerful. And the use of ontologies is very powerful. it gained some traction in in the Netherlands and in in some with some Scandinavian clients as well is likely to become more widely adopted But It's there aren't that many people who have expertise in it." "we don't necessarily train all junior staff on the detail of IFC." "to implement those properly on a project needs a lot of effort on behalf of the public clients to write the exchange information requirements into the contract, but I just don't think they have the expertise to be able to do that."	P7
		"Structural engineers typically don't do the modelling…. we're still in the Pdf markup stage in ABC company at the moment… It's mainly because the structural engineers haven't been upskilled in BIM, and they're not allowed to use it…. we don't actually have a BIM workflow from analysis to modelling at the moment. So, we asked them to just model, based on sketches or markups." "But in ABC, we're not supposed to. They keep it really clear that we don't do that because not everyone has the skills."	P8
		"We have to remember that most people on a construction project aren't overly technical. They might be experts	P9

		in their trade, but when it comes to information management, they're beginners. We can't expect them to understand this level of theoretical prescription. Even the cheat sheets are a stretch for them. The standards are written at such a high, almost academic level, it's hard to see how they apply practically on a project. There's a big knowledge gap between the people doing the work and the people writing the standards."	
		"The biggest issue is the lack of skills and the time it takes to develop a model to that level of detail." "If a client specifically wants to use the model for BCAR, that needs to be clear from the start. When we're tendering, we'd need to price in the extra work: higher standards, maybe subcontractors with better skills, and tighter model management. It's doable, but the skill shortage is real." "Some tasks, like updating materials in Revit for U-values, are easy. But modelling rebar, for instance, is an advanced skill. It needs an advanced BIM modeler, not just a regular one. So, it's partly a skills issue, partly a resources and cost issue."	P10
Contracts and Compliance	**Regulatory Requirements**	"for statutory approvals, fire safety certificates, disabled access certificates. So, therefore, achieving statutory approvals" "I suppose, the other	P1

		type of documents we do would be certification. So, it would be in related to quality control and compliance with building regulations.”	
		“And if it's not a BCAR project, we would produce an opinion of compliance with the building regulations.” “for the BCAR. So, the BCAR is the building control amendment regulations. And so, there are building regs compliance certificates, they're a standard PDF template that's produced, that, you know, it's an industry standard template produced.” “There's an ancillary certificate of design commencement, ancillary certificate of design completion, and then an ancillary certificate of design inspection. So, there's 3 certs that we have to do. But each individual engineering discipline will have to produce those, and then the contractor will have to produce. There’re separate certs, the contractor has to submit. They all are packaged together and given to the assigned certifier, who then produces his overall cert, which is given to the local authority to say that you know the certified of the building has followed the building regulations.”	P2
		“even if you take the BCAR process. You know there is a large number of documents that they need to be produced throughout the	P4

		construction stage, for whatever you know, stage applies of a project, and then at completion."	
		"the building authorities do not regulate the use of certain unstructured data, and Some it does. So, for example, anything that has to relate with quality or with safety management that would be regulated through the contract, and if it's deemed necessary"	P5
		"We don't yet have model-based planning submissions in Ireland."	P7
		"We have BCAR trackers for all deliverables, reports, and inspections, but the third party manages that."	P9
		"BCAR. Okay, so typically, a lot of the time that's managed on the client's CDE. They usually request their own CDE for that, which sometimes adds extra work for us. We have to keep a copy of the records ourselves, so we upload them to our CDE in a location we can always access. Then we also upload them to the client's BCAR CDE. So, there's a bit of duplication in that process. As for the other parts, I'm not fully involved in all the BCAR requirements, but generally, we need two copies: one we maintain for our own records, and another that we issue to the client for their records."	P10
	Contractual Reasons	"No one's paying me to learn BIM. And no one's paying me for the extra	P1

		cost to use it in the office." "we're not being incentivized to use BIM. We're being penalized for not using BIM. So, we're being told you can't have government work unless you use BIM, and you need to pay more money to train your staff and buy the software. I think that's unfair. Because we're the people in making the investment, not the government, like the private individuals, are making the investment, not the government."	
		"There can also be quite a bit of time involved in administering those common data environments. So, unless we are specifically contracted to do it, we would prefer not to."	P2
		"But there may be a contractual obligation to provide certain types of documents. So, for example, like specifications would be needed for tender. That would be always part of that project issued." "because things get quite contractual for us on when a project is on site. It's much better to be able to refer to a PDF that shows our exact design intent rather than a model that would have to be completely perfect in every single junction, rather than just having like one drawing that kind of says, this is This is a typical detail or a design intent detail, and this should be applied across the whole project."	P3
		"ABCF one, I have to	P4

		say, though, that even though the project was carried out in Revit within the design team, the project was not Revit with the contractor, and that because of the risk associated, and even the ABC at the time was recommended when we entered the contract, so there was no Revit exchange whatsoever with the contractor throughout all the process of construction." "I'm just saying is, and the project we are talking about was fully Revit is just the interface on the public form of contract that at that time when we signed the contract with our contractor wasn't within a Revit. So, the contractor didn't use Revit." "But again, what we provided is what we've been asked by the client at the time of signing the contract with the contractor." "The BIM process is not covered into the contract in the public form of contract"	
		"whenever you begin a project, you would have to specify in the contract what information you want to store, So, that would be reflected in the safety file." "if it's deemed necessary, it should be specified in the contract prior to beginning the project."	P5
		"Because a couple of vague lines in an EIR can lead to us spending months gathering all sorts of information	P6

		only to find we've got the wrong stuff, missed something, or given them too much."	
		"A broad ranging insistence on IFC, is a little bit, I have a bit of difficulty with it you know. The question is, why and what's the purpose? And what's the benefit? And who's going to pay for it? Are they willing to pay? And is it actually going to be once you're in a project, as someone actually going to insist on it being done and someone going to monitor and validate it that it's been done properly?" "They didn't make it worthwhile for the contractors, while all the payments were related to actually delivering the construction, and at the end of the job, it wasn't really in the contractor's interest to spend lots more resources and time tidying off all the information and handing over to the client to comply with the contractual obligations. So, it's important, if clients really want this information exchanged in a particular way, they have to associate that to a meaningful amount of money in the contract. Basically, they have to be willing to pay for it, and they have to put that clearly into the contract."	P7
		"There has to be something in it for everyone.yeah. And then, yeah, there's also how you share models and stuff. Actually,	P8

		sharing models is always an issue. Before, in previous companies, it was an issue. It's not so much in ABC, because everything's in house. But we'd strip out all the information usually when we issue models so that people couldn't take our IP and things like that. So, they just get the geometry. So, like this data rich thing, and then you strip out all the data before you share it"	
		"The models are often seen as risky because they're not contractual deliverables, and there's a lot of inaccurate information in them." "I would encourage the government bodies to stick with their current format. For example, we received tender documents from a government body, a school and the Department of Education has the Information Protocol, and that's appended to the main contract. That makes everything contractual, and that's a big step up. That means we have to play ball, and that means we have to get compliance from our services. It increases the cost, but we have to make it work because of this Information Protocol. It takes the casual nature out of it. I'd imagine there will be pushback on this, particularly from SMEs and smaller contractors bidding for things like schools who aren't certified. I think the government may actually step back from this one, but I'd encourage them not to	P9

		stick with it, because that's the only way the industry is really going to step up to the plate, when it's enforced at a contractual level."	
		"I think it comes down to the contract. On all of our projects, no matter how good our models are, it ultimately comes down to the drawing as the contract document. The reason for that is, it's what the design team and subcontractors are comfortable standing over. Drawings are always the contract." "For example, we're on a project under the Capital Works Management Framework, and while a BIM process is required, the deliverable is the drawings. The response is, "Here are the drawings, now you need to produce your own models again from scratch." That's a problem. There should be a framework where the model passes over to the contractor at a milestone. Otherwise, you're paying a design team to produce a model, and then paying the contractor to create the exact same model again just updating the level of detail."	P10