MANAGING DATA FLOWS IN INFRASTRUCTURE PROJECTS - THE LIFECYCLE PROCESS MODEL

Heikki Halttula, Managing Director
Topcon Technology Finland Oy, Espoo, Finland
Heikki.Halttula@topcon.com

Harri Haapasalo, Professor
University of Oulu, Industrial Engineering and Management
Harri.Haapasalo@oulu.fi

Risto Silvola, Director
Konecranes Finland Corporation, IT and Business processes
risto.silvola@konecranes.com

SUMMARY: Productivity in the construction industry (both houses and infrastructure) has not been improving as expected, while other industries have been able to improve their productivity significantly. The appropriate use of building information modelling (BIM) technologies brings several benefits and advantages to construction projects. The main challenges of project efficiency emerge in the form of numerous requests for information during the construction project, which are considered to be waste in the processes. This highlights the need for a practical process model to plan the information flow for BIM-based projects. The main aim of this study is to propose a model to plan the flow of project information among primary stakeholders especially in infrastructure projects. Our main findings are firstly, the foundation for data management starts from defining unified one data for the product and the for the process. Unified data means one single repository of data – all stakeholders use the same unified data. It is also essential that data responsibilities and ownership are defined. Secondly, we found that the biggest challenges are that the data needs are not planned beforehand, resistance to change, difficulty receiving existing data and data must be modified before use. As a whole, it seems sometimes that the technology on data transfer is more important that what has been transferred and why. Finally our construction, the life cycle model for data flow originates from one data to all stakeholders, single data repository must be updated along the life-cycle of the object covering also the operations and maintenance, where the data has to be updated through the whole life-cycle. This new approach is intended to enable the early involvement of maintenance stakeholders in designing product data for a project lifecycle perspective. The model helps to change the current information flow and gain the benefits that a BIM-based process can offer. This study is based on case studies and is qualitative in nature and naturally needs more validation.

KEYWORDS: BIM, life cycle, data flow, infrastructure project, data model, master data, PDM/PLM

1. INTRODUCTION

Productivity development in the construction industry has been very weak during last 40 years, while in some industries it has increased by up to 200% since 1964 (Eastman et al., 2011, Pekuri et al., 2011). Digitalisation and information and communication technology (ICT) have played a major role in this improvement. Value-added time in construction is 10%, while it is 62% of the total work in the manufacturing industry (Eastman et al., 2011). Digitalisation is rapidly changing the way people work in the industry. Construction placed last out of 15 studied trades for the adoption of digitalisation (Friedrich et al., 2013) Digitalisation in construction has confronted trends like building information modelling (BIM), mobility, cloud technology, big data, 3D printing, internet of things (IoT), robotics and software as a service (SaaS) solutions.

There are several definitions for BIM (American General Contractors 2006; Penttilä, 2006; Succar, 2009; Eastman et al., 2011; BuildingSmart, 2017). According to Eastman et al. (2011), the acronym BIM means either building information model, building information modelling or building information management. In general, the acronym BIM is understood to mean not only a product model but also a modelling technology with relevant processes to develop, communicate and analyse building models. Some references also discuss the process of building. Kunz and Fischer (2009) have introduced virtual design and construction (VDC), presenting a computer-based description of projects. The VDC project highlights the product, the organisation and the process that the organisation teams will follow. Although the role of BIM is larger than a technical description of the data model, it cannot solve the process efficiency challenges alone. According to Halttula et al. (2015), BIM's role in construction is more like a tool or an enabler that helps to improve process efficiency.

While most frequently reported benefits of BIM are related to cost reduction, better quality control and significant time savings throughout the project lifecycle, challenges are mainly focused on BIM software interoperability and the challenges of handling big data sets (Bryde et al., 2013). The UK Cabinet Office has declared that BIM was a significant contributor to the savings of £840m in construction costs in 2013 and 2014 (HM Government, 2015). Most of the investments are made in the operation and maintenance phases, but their data needs are not yet in focus. The use of data models in the construction industry is concentrated on individual phases. The lifecycle approach has received minor attention. Project managers failure to identify what data flows between teams is critical and can lead to design process problems and reworking on-site. Project's contractual arrangements result in bottlenecks in the flow of requests for information (RFI) (Love et al., 2008). The use of BIM has many benefits, but poor lifecycle approach to data management hinders realising these benefits. HM Government's (2015) Digital Built Britain strategy pinpoints yearly investments in construction at £89bn and in maintenance and operations at £122bn in the UK. For service providers who use the infrastructure, the share of GPD (gross domestic products) is £597bn. From this point of view, it would be interesting to reinforce the operation, maintenance and service providers role in the early design phase.

In the electronics industry, product data has played a key role, but in the construction industry, it has not been widely recognised. In manufacturing enterprises, product data management (PDM) has become one of the main features of business efficiency concepts. All product lifecycle phases need different type of data, and all data must be managed to deliver correct information to multiple stakeholders when needed (Yang et al., 2007, Rachuri et al., 2008). Silvola et al., (2011) nominates this as "one master data". PDM is the basis for digital operations management systems in organisations (Stark, 2005). In addition to PDM systems, there are product lifecycle management (PLM) systems that take care of product management through the lifecycle. In addition to PLM and PDM systems, there are other systems that process product data like enterprise resource planning (ERP), customer relationship management (CRM) and computer-aided design (CAD) applications (Loshin, 2001). However, in construction industry the information systems nor data does not create continuum during planning, design and construction, not to mention maintenance and renewal (Ghaffarianhoseini et al., 2017, Haapasalo, 2018, Halttula et al., 2015 Halttula et al., 2017, NIST 2004, Rostami & Oduoza, 2017, Succar 2009). Complexity is high in electronics products (like telecommunication networks), therefore it is useful to learn data management over life-cycle from there. Telecommunication network deliveries are project business and kind of infrastructure projects also.

Data and data management provide significant possibilities for improving efficiency in every industry. However, the construction industry has not been able to grasp its share, even when there have been positive efforts towards this. One of the main reasons for this has been sub-optimisation and putting effort only into some lifecycle phases, leaving large parts of the process underutilised. The situation is basically the same with house or infrastructure

projects, even when the object to be built are very different. This study aims to analyse the main challenges of data flow and to construct a foundation for a process model of organising data management in infra-projects' life cycles. These objectives are operationalised on the following three research questions:

RQ1   What are the cornerstones of data management over the project life cycle?

RQ2   What are the main challenges of data flow in the infra-construction project life cycle?

RQ3   How should data management be organised in infra-projects?

This study is qualitative case study (Yin, 2003). At first, we review literature on cornerstones of data management over the project life cycle to create foundation for our empirical case studies. Data management is reviewed from the generic literature to learn practices that already work in practical operations. Secondly, we have selected our empirical three case projects (Yin, 2003) from the infrastructure industry to keep the scope narrow. The main contribution of the paper is to construct overall model (see Eisenhardt, 1989) for organising and managing the flow in an infra-project, based on generic earlier literature and empirical cases from infra sector.

2. LITERATURE REVIEW ON DATA MANAGEMENT

2.1 Digitalisation and BIM

Digitalisation, automation and data management are embedded in everything in our modern economy. There are uncounted numbers of gadgets and services that offer individual services for people and organisations to improve the efficiency of processes and activities. However, in many cases, these solutions only see efficiency from one point of view, not from the point of view of the entity. This rarely leads to an optimal situation. Digitalisation means renewing and developing activities in a way that improves processes and the whole logic. In many cases, renewal has only resulted in the partial automation of processes. (Haapasalo, 2018). Former opportunities and problems have also been seen in construction.

As a response to the growing complexity of construction projects, the application of ICT solutions is seen as a solution for the building industry. The benefits of BIM have been listed in several studies (Azhar, 2011, Antwi-Afari et al., 2018, Bryde et al., 2013, Merschbrock and Munkvold, 2012, Autodesk 2002, Eadie et al., 2013, Barlish and Sullivan, 2012, Dakhil et al., 2019), which claim that BIM increases the quality and speed of building processes while decreasing the cost. Challenges of utilisation and implementation are also numerous and studied widely in relation to not only technology but also change in processes and organisational settings (Eadie et al., 2013, NBS, 2016, Ghaffarianhoseini et al., 2017, World Economic Forum, 2018, Bryde et al., 2013, Eastman et al., 2008, Yan and Demian, 2008, Azhar, 2011, Merschbrock and Munkvold, 2012). Analogical to the utilisation of CAD, BIM has been expected to solve problems like the cookbook recipe (Haapasalo, 2000), which is not, of course, the case. From the literature, it is not actually possible to see whether "building" in BMI is understood as a noun, meaning the product in construction, or as a verb, meaning the construction process. These are, of course, very closely related, but their origins are very different. Like many authors (Eastman et al., 2008, Azhar et al., 2012) propose that the focus should not only be on technology but actually more on processes.

The problem is very challenging. To simplify the complexity, we must first concentrate on the product and then process (Silvola, 2018; Haapasalo, 2018). The origin of data management (i.e. BIM) during the life cycle is the product — the object to be created or the object having the whole life cycle — to be modified in different phases. Silvola (2018) describes these life cycle phases of the built object: as planned, as designed, as built, as sold, as operated and as maintained. All phases (or stakeholders) should be dealing with the same product data. These life cycle phases or processes are, however, no less important than the product itself. Silvola (2018) emphasises the meaning of one data as the DNA of the product during the life cycle, meaning that there can be one correct product data at the time. Every stakeholder should use this one data. If data is stored in different systems or stakeholder silos, it will most likely differ from the original.

On approach to linking data with practical processes is to implement enterprise architecture in an organisation. Enterprise architecture is a systematic way to describe principles, methods and models that are used in the design and recognition of an enterprise's organisational structure, business processes, information systems and infrastructure. Enterprise architecture offers consistency and integration between business, IT (information technology) strategies and processes. (Lankhorst et al., 2009, Aziz et al., 2005). Business architecture shows the strategy, models, processes, services and organisation. Comparably, information architecture identifies, documents

and manages information needs, assigns ownership and accountability and describes how data is stored and exchanged between stakeholders. Technical architecture defines the strategies and standards of technologies and methods used to develop, execute and operate the application architecture. Application architecture outlines the specifications of technology-enabled solutions in support of the business architecture (Berson and Dubov, 2007).

2.2 Data governance

Cohen (2006) describes data governance as a practice in which an organisation manages the quality, consistency, usability, security and availability of its data. It covers a whole enterprise's data management, business processes and risk management. The goal is to create a positive control for the policies of data creation, retention, usage and archival. Wende (2007) and Weber et al. (2009) describe these roles in a RACI model:

- Responsible (R): a role that is responsible for executing a particular data quality management (DQM) activity. The accountable role determines the degree of responsibility.
- Accountable (A): a role that is ultimately accountable for completing a DQM activity or authorising a decision.
- Consulted (C): a role that may or must be asked to provide input and support for a DQM activity or decision before it is finished.
- Informed (I): a role that may or must be notified of the completion or output of a decision or activity.

The roles for data governance described by Cohen (2006) are data steward, data owner, data manager and data user. Data stewards take care of data governance policy and advise data owners and data managers about how to apply those rules. They develop, monitor and control policies for data and are overall coordinators for enterprise data delivery efforts. They continually improve the data flow and measure the performance. Data stewards can be one person or a stewardship committee.

Data owners are data stewards co-workers, and they are responsible for defining enterprise information requirements. Every business function has a data owner. They develop standards for the storage, retention and disposal of corporate information. They ensure information quality and availability. Data managers, or custodians, carry out the data delivery function. They work closely with the data stewards and data owners to implement data governance policies. There are several data managers in an organisation, and they help users with current applications and technology. They are responsible for gathering process improvement ideas. On the technical side, data managers follow the policies defined by data owners, and they capture, store, retain and dispose of enterprise information. They design the technical infrastructure accordingly to data owners' information requirements. Data users are not part of the official data governance organisation, but they are an essential part of the system because all the policies, requirements, delivery mechanisms and technical architecture designs are made for users. Without users, there would not be a need for data governance. The desires of the user community drive the need for data governance. (Loshin 2001, Cohen, 2006, Khatri & Brown 2010) Silvola et al., (2011) introduce the product data owner network (PDON) concept to govern data, describing definitions, policies and guidelines, data management activities, data owners, roles and responsibilities and enterprise's business functions. Implementing a PDON requires an understanding of the relevant data owner roles.

2.3 Product structure

In the manufacturing industry, businesses rely on the product structure. It is the backbone of PLM and a critical component in creating a PLM system (Saaksvuori and Immonen, 2008; Kropsu-Vehkaperä et al., 2009). Product structure describes the parts or components and their properties, service features, documents, installations and the relationships between them. Product structure and the items connected with it are the basis for many functions of a system (Saaksvuori and Immonen, 2008; Crnkovic et al., 2003).

Product structure is usually a hierarchical structure with a division of parts into a hierarchy of assemblies and components. An assembly consists of subassemblies and components. A component is the lowest level of a structure. Product structure is a conceptual model that describes product-related information and how information is connected to other product information objects (Saaksvuori and Immonen, 2008). In the construction industry, the bill of materials (BOM) is a well-known concept. Product structure may include common manufacturing and maintenance configurations, whereas BOM describes a particular product material breakdown structure (Svensson and Malmqvist, 2002, Peltonen, 2000).

The conceptual method that is typically used to design and describe product structure is called object-oriented. Objects are data elements that define a product component, product element, module, subsystem or assembly. A structure consists of objects that can have functional or hierarchical dependencies in relation to one another. The product structure, which may include different levels, contains joint hierarchies of various objects. In the object hierarchy, properties inherited from father to son: lower object classes have the properties of higher classes together with some additional or changed features. Objects have attribute information that describes their properties (Saaksvuori and Immonen, 2008).

2.4 Master data

The electronics industry has a long history of managing product data, related transactions and processes. There is a culture of managing product information for production, customer service and maintenance purposes. Master data is intended to organise information that is regularly needed. Master data defines the different entities utilised by the industry as parties (customers, dealers, employees or citizens), places (locations, offices, regional alignments or geographies), and things (accounts, assets, policies, products or services) (White et al., 2006, Moss, 2007). Master data defines business-related features together with connected metadata, attributes, definitions, roles, connections and taxonomies (Loshin, 2009, Dayton, 2007). Only the subset of elements required for data sharing and standardisation are master data; master data objects are the key business elements that matter the most. Master data helps enterprises realise the factors and trends that may affect business (Berson and Dubov, 2007). Otto and Huner (2009) have categorised four features of how master data differs from other types of data:

- Master data always describes the basic characteristics (e.g. the age, height and weight) of objects from the real world. Unlike transaction data (e.g. invoices, orders and delivery notes) and inventory data (e.g. stock on hand and account data).

- Master data usually remains unchanged. If the characteristic features always remain the same, there is no need to change the respective master data.

- Instances of master data classes (e.g. customer data) are quite constant regarding volume, at least when compared to transaction data.

- Master data forms a reference for transaction data. It does not need any transaction data in order to exist.

"One master data" is a concept of finding a set of data that does not change and is created only once. One product master dataset is created by the product portfolio management process. It is seen as the DNA of the product (Silvola, 2018).

Master data management (MDM) is a process where business units use information systems to harmonise, cleanse, publish and protect common information assets that must be shared across the organisation (White et al., 2006). The goal for MDM is to form an integrated, accurate, timely and complete dataset to manage and grow business (Berson and Dubov, 2007). Master data management is a discipline to define and standardise key business data and manage changes to those definitions over time (Dayton, 2007; Moss, 2007). It is often divided into two parts: operational MDM and analytic MDM. Operational MDM integrates operational applications, such as ERP, CRM and supply chain management, in upstream data flow. Analytic MDM is seen in practices that involve data warehousing, such as customer data integration and financial performance management. Together they form the enterprise MDM (Apostol, 2007).

2.5 Product data

Engineering and manufacturing corporations have recently been focusing on PDM, and it has become one of the main features in their business efficiency concepts. There also is a trend in business development to concentrate more on the organisation's key competencies, which is leading to increased collaboration with partners, suppliers and contractors (Ameri and Dutta 2005, Stark 2005, Saaksvuori and Immonen 2008). At the same time, the amount of data produced and transferred during the product's lifecycle is enormous and continues to increase (Ameri and Dutta 2005).

There are several phases of the product lifecycle which each require different data: for example, design requires product data, construction requires transaction data, customer interface requires customer or user data and maintenance requires data of renovations. All this data must be managed in an integrated and systematic way to

deliver correct information when needed to stakeholders (Yang et al., 2007; Rachuri et al., 2008). Organisations that are not able to manage their product data will have difficulty managing their products (Stark 2005). Many organisations are not aware of what product data they have, how critical it is, where the critical data is stored and how redundant their product data assets are (Khatri and Brown, 2010).

Product-related data can be organised, stored, maintained and utilised with PDM systems. PDM is the basis for digital operation management systems in organisations (Stark, 2005). In addition to PDM systems, there are PLM systems that take care of product management through the lifecycle. Other systems also use product data, including ERP, CRM and CAD applications (Loshin, 2001). Although organisations have up-to-date systems for managing product data, their data may include errors (Kropsu-Vehkaperä et al., 2010). According to Godfrey (2002), data quality problems are one of the most prominent challenges organisations are facing.

2.6 Cornerstones of data management

As a synthesis from earlier literature digitalisation and data offers possibilities for major leaps in industry, supply chain or business level, the information and data are still servants, not the master. This holds even when they enable the full re-engineering on business processes. Business processes and value chains utilise data and information and, therefore, define the overall framework for the entity as receiving business processes. Data and its definitions are important in a quality sense; otherwise, the entire system fails. It is critical to understand product (built object) data, master data, transfer data and, especially, data governance: who produces, modifies, utilises or owns the data and how these roles possibly change during the life cycle of a built object. If we look at the business processes and the product, it is obvious that the product (the built object) is the starting point. It provides services to customers that the infrastructure in question should provide. It is also the object that needs to be planned, built, utilised and maintained. The correct and updated data concerning the object forms the foundation for everything. At a single organisational level, it might be easy to understand the product that the organisation is delivering, but the bigger challenge emerges when different organisations combine their contributions on an entire project.

3. RESEARCH PROCESS

This research is qualitative in nature. It consists of a literature review, the analysis of projects as case studies to identify challenges and the generation of rough level data flow for infra-projects (Eisenhardt, 1989; Yin, 2003). The literature review focused on data-related issues both at the industry digitalisation level and at a more detailed data management–level. (FIG. 1.) We utilise synthesis of literature in twofold, first analysing the cases trough literature findings to find out the challenges and using the findings of operational models directly to create data management model for infra-projects. The purpose of this study was to provide a balanced understanding of the life cycle level needs, the utilisation of data and the requirements of data. Digitalisation, BIM, guidelines for the infra-project life cycle phases were used to outline the big picture for infra-projects. Comparably, master data, product data and data governance set the guidelines from a more detailed perspective.
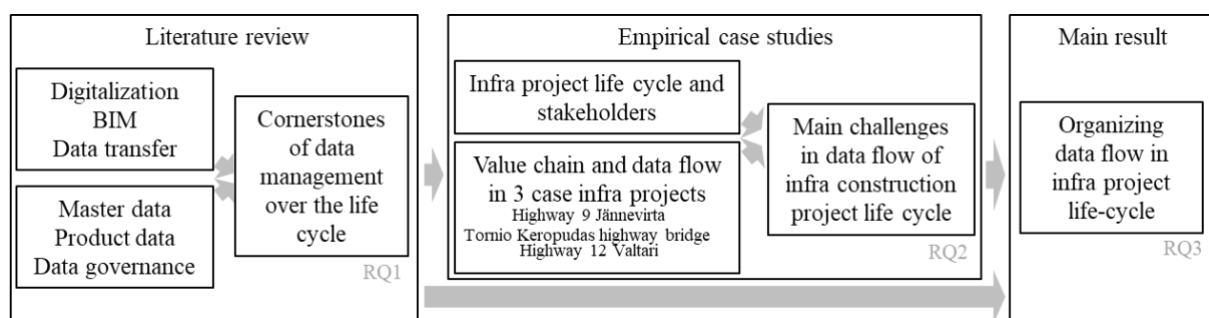

FIG. 1. Research process and the logic of the research.

Three empirical case studies were carefully planned based on the literature review a compiled understanding from the review of earlier research (Eisenhardt, 1989, Yin, 2003). As can be seen from the case studies, the specific features of the infrastructure industry were documented at a rough level by the life cycle phases and stakeholders. Material for the case studies was then collected. Empirical material for all three cases was collected in the same way; the main source of material was interviews of the key personnel in the project, i.e. project managers. Additional sources of information were project and public documents. Swim lane diagrams for the main

information flows were completed during the interviews to guarantee valid understanding. After the interviews, additional material was utilised to finalise the case descriptions. Challenges of data flow were synthesised from the interviews and the additional material. Based on the literature and the main challenges, we created a main-level organisation for information management in infra-projects and a more detailed description of the model.

4. DATA MANAGEMENT OF THE INFRA-PROJECT LIFE CYCLE

4.1 Infrastructure project life cycle and stakeholders

In this study, we concentrate on the infrastructure project life cycle, which consists of phases: preliminary design, design, construction and maintenance. According to Investopedia (2018), the term "public works infrastructure" was adopted by the U.S. National Research Council and refers to functional modes, including highways, airports, water supply and resources, telecommunications and the joint systems these elements include. Infrastructure can include a diversity of systems and structures with the condition that there are physical components. Infrastructure projects have different kinds of data models and data transfer systems. Many infra-projects use a LandXML-based data transfer format; in building projects, IFC are used. Many infra-projects include structures, like tunnels and bridges, that are modelled using Industry Foundation Classes (IFC). These sub-models must be included in infra-LandXMl sub-models in a real-world coordinate system. This presents a different data management challenge for infra-project than for sole building projects.

According to the PMBOK (2000), a project has distinctive features: "a project is a temporary endeavor undertaken to create a unique product or service". Temporary means that project has an exact start and exact end. A project is unique; it does not resemble any product or service. A project team is created to perform the project, and it is usually separated after the project is over. Although there are repetitive elements, they do not change the uniqueness of the project. Thousands of highways have been built, but each highway is unique, having different clients, different designers, different constructors, different terrain and so on. Organisations typically divide projects into phases for better management control and to better build relations with the organisation's current operations. Together, all the project phases form the project life cycle (PMBOK, 2000). Morris (1988) describes the construction project life cycle by dividing the project into five phases: feasibility, planning, design, construction-manufacturing, turnover and start-up. "Feasibility includes project formulation, feasibility studies, strategy design, and approval. Planning and design include base design, cost and schedule, contract terms and conditions, detailed planning. Construction includes manufacturing, delivery, civil works, installation and testing". In Finland, the highway construction lifecycle has five phases: preliminary investigation, general planning, traffic engineering, execution (detail design and construction) and maintenance (FTA, 2010).

In the current study, we divided the project into four phases: preliminary design, design, construction and maintenance. Preliminary design includes the phases before the decision of the construction has been made, including feasibility studies, general planning, environmental impact assessment and traffic engineering. Design means a detailed design for construction. Construction means the actual execution of construction work, and maintenance means actions to keep the route at the desired condition level for operation. Together these project phases are called the life cycle.

According to Aapaoja et al. (2014), there are several groups of stakeholders in a construction project along their salience (key players - primary team members; keep informed - key supporting participants; keep satisfied – like tertiary stakeholders; minimal effort - extended stakeholders). The main types of stakeholder in one infra-construction project can go up to several tens not mention individual stakeholders. To ensure the best results and added value, the project should focus on the most salient stakeholder requirements. Stakeholder groups can include customers, end-customers, main contractors, subcontractors, property operators, main designers, other designers, public authorities, material suppliers, residents and neighbours, construction consultants and sponsors. In this study, we have concentrated on analysing key players as primary team members as case specific. We have analysed these stakeholders according to life cycle phases: preliminary design, design, execution and (operation &) maintenance. This classification is adequate for studying data flows between life cycle phases, not the individual stakeholders' requirements, with the goal of producing a model for planning the data flow of a life cycle level.

4.2 Value chain and data flow in an infra-construction project

4.2.1 Highway 9 Jännevirta Bridge

Kreate is one of the leading infrastructure construction companies in Finland. Kreate has 300 employees, and its turnover is approximately 180 M€. Our first case study project was the 35 M€ Highway 9 Jännevirta Bridge renewal project in Kuopio, Finland. The client was the Finnish Transport Agency (FTA). The project was the Design Bid Build (DBB) project where the old bridge over the Saimaa Lake deep fairway was replaced by a new concrete bridge. The road geometry was also improved 3 km on both sides of the bridge.

In the DBB projects, the client organised the first bid for design work. After the winning design organisation finished the design phase, the client gave design information to the construction companies for bidding. After the bidding phase, the winning company began the construction work. After the construction work, the constructor produced quality assurance and as-built measurements. As-built models were then given to the client.

In FIG. 2., the preplanning documents are used as the basis for the bid of design (1). The winning design stakeholder executes the design work and delivers project data to the client according to the agreement (2). The client gives the design information to the construction companies for bidding (3). There is a possibility to get design information straight from the designer to the constructor, but liability issues reject the amount and quality of the data (4). The construction bidding team delivers information to the construction team (5). Construction company obtains subcontractors (6). Constructor delivers the agreed as-built information to the client (7). The data need for the maintenance phase is unclear (8).

The great majority of the Finnish infra-projects are DBB projects. Data flow between stakeholders has bottlenecks in this type of project (and in general in DBB projects in Finland) due to fragmented contact (Haapasalo 2018). Constructors have difficulties receiving all the data needed from the client and the design organisation. In DBB projects, the client has the first bid of the project design work. The FTA has a policy that requires BIM to follow its modelling guidelines.

In Jännevirta project, the design information was given in IFC, XML (Extensible Markup Language) and infra-model formats. Constructors had to edit the geometry information of roads to get the surface models to the automated machinery. The structure's IFC-model quality varied depending on the used design applications. One challenge is that the design organisation provided only models that did not include any uncertainties because they were afraid of claims from the construction company.
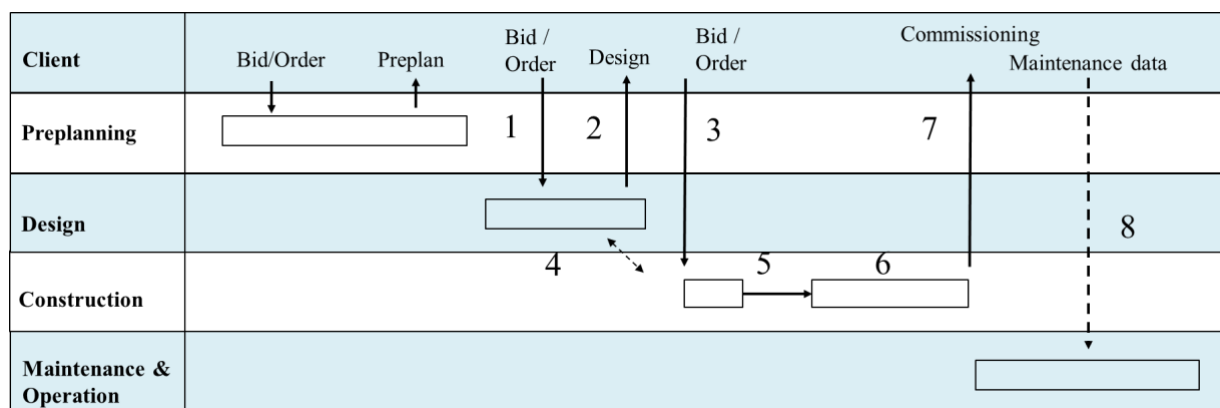

FIG. 2. Rough data flow in Highway 9 Jännevirta bridge. The project was a DBB project where the client (FTA) divided the project into separate phases: preplanning, design, construction and maintenance. The client was responsible for how and what data was transferred to stakeholders between life cycle phases.

The data management was not handled well in the bidding phase. In our case project, it was hard to get data from the client and the design organisation. The internal data flow in the construction phase from the bidding foreman to the construction team was weak. The bidding foreman was used to having the design information in a pdf format and did not require the model data even though it was available.

In the construction phase, Kreate used specific software to produce and enhance collaboration models. An application was used to create machine control models. The site engineer, site foreman and responsible site

foreman had communication problems. A common model for the whole project team that showed the differences in the designed model could have helped communication. In the construction phase, there is no specific application for mass haul planning. This work was done based on the work experience of construction supervisors; no specific applications were in use.

There are no actual requirements for maintenance phase data. Constructors produced as-built models for quality assurance as requested by the client. Some clients wanted pdf files of different road construction surfaces that were printed in a paper format and archived. The design models worked for an as-built model, but if the tolerance was significant, the construction had to be modelled again. Usually, all the possible as-built information is delivered to the client.

4.2.2 Tornio Keropudas Highway #9 bridge renewal project

Our second case study project looks at a railway bridge renewal project in Tornio Keropudas Finland. The client was the FTA, and the project delivery type was design built (DB). In the project, Kreate built the new railway bridge in place of the old bridge. The old bridge was first jacked away, and then the new bridge was jacked in the location of the old bridge. The requirement was that the interference to the railway traffic be as short as possible. The design of the bridge was started in early 2018, and the total project finished at the end of 2018. The design engineering consultant was Suunnittelu Kide. The geotechnical engineering consultant was Pöyry.

The value stream of DB projects is that clients can use preplanning results as the basis for the bid on DB projects. The winning DB consortium's designer executes the design work and delivers project data to the constructor according to the constructor's requirements. Designers and constructors work in parallel and influence one another's work. The constructor gets design information straight from the designer in the format he or she wants. The DB consortium delivers the agreed as-built information to the client.

The client (FTA) set a requirement that the project have to be done using BIM. In DB projects, Kreate can establish requirements for the information they receive from design engineers. It can specify what information it wants and in what format. In the case study project, the bridge design information was transferred using an IFC format and geotechnical information in a DWG (drawing) and 3D DWG format. In a DB project, the data delivery problems are easier to solve than in DBB projects. Maintenance operators needs were not discussed in the case study project.

Design engineers used different applications for structural engineering and work planning of the project. In the IFC data of the project, there are right metadata and the right classification. The design engineers modelled the rebar of the bridge, which helped the construction. The production of printouts was optimised, which saved costs. The constructor got the design data easier than in a DBB project. One major change is that there were experienced designers available during the construction phase. They were able to use different design applications to solve on-site problems. The tools can be more specific for design editing and management than in DBB projects.

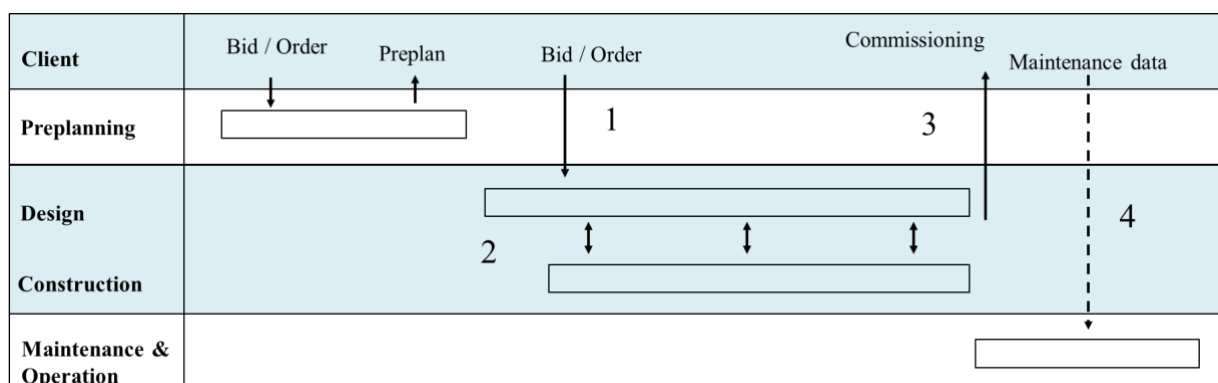In FIG. 3, the client orders the preplanning of the project.

FIG. 3. Data flow in Tornio Keropudas Highway 9 Bridge renewal project. The project is a DB project where the client (FTA) has divided the project into separate phases: preplanning, design, construction and maintenance.

The preplanning results are used as the basis for the bid of the DB project (1). The design stakeholder executes the design work and delivers the project data to the constructor according to the constructor's requirements. Designers

and constructors work in parallel (2). The constructor gets design information straight from the designer in the format they want (2). The DB consortium delivers the agreed as-built information to the client (3). The data need for the maintenance phase is unclear (4).

4.2.3 Highway 12 Valtari Alliance project

The Finnish company Skanska Oy employs 2,086 people. This case study interview was done with a project manager and ICT specialist. Highway 12 Lahti's southern ring road development is divided into three projects: Part 1A DB project, Part 1B Alliance project and Part 2 DBB project. The total value of the development is 275 M€; Part 1B is estimated to cost 170 M€. The Valtari Alliance group consist of the FTA, Pöyry Finland Ltd., the City of Lahti, Skanska Infra Ltd. and the Municipality of Hollola. In the Valtari project area, there are 4,5 km highway 2 by 2 lanes, 3 interchanges, two tunnels (first 0,5 km and second 1,0 km), streets, bicycle lanes and 12 new bridges. The project area is on the Laune groundwater pool.

In the Valtari Alliance project delivery agreement, the client, design and construction stakeholders form a consortium where they promise to fulfil project Part 1B together. In the bidding phase, there were four consortiums competing. Two of the candidates quickly dropped off according to tender documents. The remaining two candidates had a two-day development session with the client, which was the basis for finally choosing a Valtari consortium.

In the agreement, there is no dispute clause. The bonus and sanctions are mutual for all consortium members. There is a good incentive to help all co-workers on the project. The maximum bonus for the project is 2% of the total project value if the quality points are 100. The deadline for the project was decided by the Alliance team to be June 2018. In the first design phase of the alliance, the goal was to find a solution on a rough level to reach the right cost level. The Alliance project delivery type made significant changes possible when members reached an agreement. The Alliance leader group, which includes the municipalities and the FTA, has powerful tools that can be used to change the final engineering plan of the project if needed to reach the best possible result.

All project stakeholders were held to the same agreement with the same bonus and sanction terms. There were no obstacles to data flow based on agreements. A project team of 50 persons worked daily in the Big Room and helped one another reach the mutual alliance goal. All project documents, like minutes of meetings, were stored in the SharePoint project portal. There was a lot of information on different computer's hard discs and databanks in different folders. That data was difficult to use because it was not possible to find the information based on metadata or the location of the project. The FTA required the use of BIM even with requirements that are more detailed than the constructor needs for quality assurance. The FTA wanted to collect as good data as possible for the maintenance phase. The maintenance operator has not yet been chosen.

Pöyry used a highway design software in detail design. In the final engineering phase, two other applications were used. All software vendors have a modified version of the transfer data format. In general, there are no accurate models of 3D viewing software currently in use. Designers used an application designed for building construction — not ideally suitable for infra-projects. Today, road models can be transferred; bridges are transferred using IFC. Instead of specific class detection software, designers use highway design applications. There is no need for major mass haul optimisation. The case study project was in the city area, so the major changes required due to mass haul optimisation were not easy. The geotechnical conditions had the potential to cause a reason for additional design work if the soil was different than in the initial site investigation.

In the project, there were subcontractors' machines, which used machine control. A measurement foreman used an application for editing the on-site data for machines. Machine control models could also be transferred straight from the highway design system. The FTA had not yet named the maintenance operator for tunnels in while case interviews took place. The data model for the old road register was insufficient, and not all data could be transferred. New data management systems for road data are in the development phase and could help the situation in the future.

In FIG. 4, the value streams of the Valtari project are drawn. The client orders the preplanning of the project (1,2). The preplanning results are used as the basis for the bid of the Alliance project (3). The client chooses the best consortium based on, e.g. expertise and team working skills (3). All stakeholders (clients, design and construction) work as equal partners towards the project goals (5). Because the clients are part of the Alliance team, it is possible to go back to the final engineering plan result, if needed, to improve the outcome (4). The alliance-consortium

delivers the agreed as-built information to the client (5). The FTA has unambiguous and detailed instructions about the as-built data to provide to the maintenance operator (6).
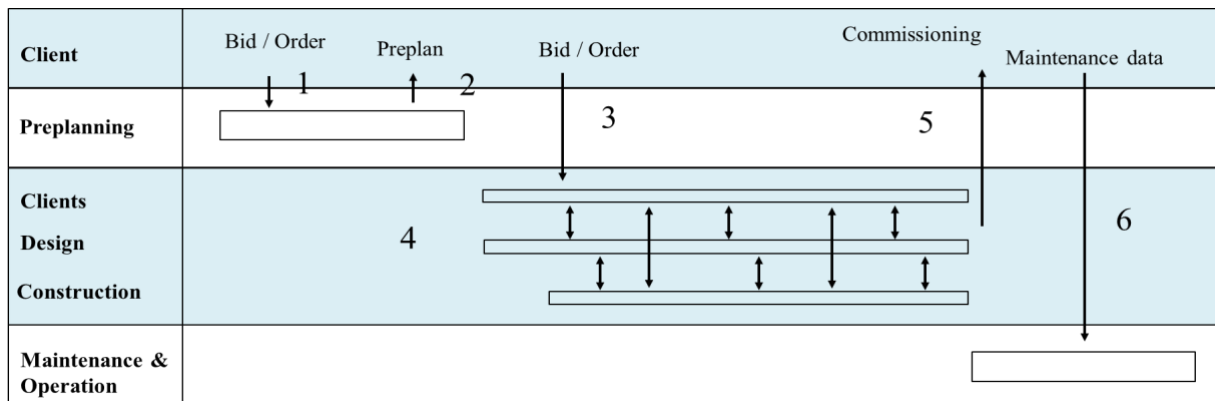

FIG 4. Value stream and data flow of the Valtari Alliance project. In the alliance, all project parties are in the same agreement where the pain and gain incentives of the project are divided equally with no dispute clause.

4.3 Main challenges of the data flow of infra-construction projects

Table 1 summarises the main challenges of data flow.

Table 1. The main challenges of data flow in the infra-construction project life cycle from the case projects.

The data needs are not planned beforehand	Resistance to change	Difficulty receiving existing data	Data must be edited before use
• Data management has been poorly organised in the bidding phase. • The internal data flow in the construction phase from the bidding to the construction team is weak. • The format of the required model for the maintenance phase is unclear. Construction companies have no connection to the maintenance organisation. • Maintenance operator's needs have not been discussed in the project. • In the 3D viewing software, there are not accurate models in use. • The FTA has not named the maintenance operator for tunnels. • The data model of the old road register is insufficient, and not all data could be transferred. • The data flow for asset management and maintenance have many challenges. • The communication between stakeholders is weak. • Data flow between project phases is inadequate. • Data management is a problem. • Specifications are missing. • Maintenance operators all have their maintenance systems, and coordination is based on tacit knowledge. • Contractors have different systems for data collection.	• The bidding foreman is used to having the design information in pdf format and does not utilise the model data even though it is available. • In the construction phase, there are no specific applications for mass haul planning. • Some clients want pdf files of different road construction surfaces that are then printed in a paper format and archived. • Characteristic features for current maintenance systems include activities that are based on physical paper documents. There are service manual books and a maintenance programme in service cards. • As-built data is very often delivered in pdf format, which means that some data is lost when digital information is printed as pdf files.	• Constructors have difficulty receiving all the needed data from the client and the design organisation. • The challenge is that the design organisation hand over the models that do not include any uncertainties because they are afraid of claims from the construction company. • Constructors must ask several times for more specific models during the construction project. • There is a lot of information on different computer's hard discs and databanks in different folders. That data is hard to use because it is not possible to find the information based on metadata or the location of the project.	• Constructors must edit the geometry information of roads to get the surface models for the automated machinery. • The structure's IFC-model quality varies depending on the design applications used. • All software vendors have a modified version of the infra-model data format.

The biggest group is "The data needs are not planned beforehand". When the data needs are not planned beforehand, it causes extra work and waste. Stakeholders are not aware of what data they can use and what data they should deliver to the next phase of the project. Most of the challenges are related to the maintenance phase. Maintenance stakeholders' requirements are not considered during the earlier lifecycle phases. "Resistance to change" is the second main group for bad data flow. Clients or project personnel require data in pdf files because that is the way they have traditionally received done it. The amount of information in the pdf file is limited compared to 3D BIM information or 3D surface models. The third typical reason for challenges of data flow is "Difficulty receiving existing data". Data exists, but liability issues in agreements reject the data flow, or the project personnel cannot find the data from different hard disc drives or document data banks. The fourth main group is "Stakeholders must edit the data before use". Data might exist, but it needs editing before use, which is waste.

In the current data delivery practice, not much attention is paid to the maintenance phase of the project. Maintenance stakeholders have to collect the required data from various sources that design and construction stakeholders have created for their individual purposes during the project. Maintenance stakeholders do not have the possibility for early involvement concerning requirements for the maintenance phase. This problem could be avoided if the maintenance stakeholders were part of the integrated project team. ICT systems in the maintenance phase, all the information would have to be in a digital format. Even if data was originally in a digital format, it is delivered to the next phase of the project in a pdf or even in paper format. Table 2 shows that the more collaboration there is in the project delivery, the better the flow of data. In all the case studies, maintenance stakeholders' requirements were not taken care of. There are liability issues, particularly in DBB type projects, that seem to reject or hinder the data flow from one stakeholder to another. In alliance-type projects, there are no limitations related to agreements between stakeholders that were parties in the alliance group.

Table 2. In the case studies, the data flow quality varied from weak to fair on a scale of weak, fair, satisfactory, good and excellent. D = Design, C = Construction, and M = Maintenance.

Project	Project type	Data flow D = > C	Data flow D, C => M	Comments
Jännevirta Bridge	DBB	weak	weak	Liability issues reject data flow. No requirements for maintenance. No plan for life cycle data delivery.
Keropudas Bridge renewal	DB	fair	weak	Cooperation enabled between D and C stakeholders. Not enough construction specific applications in use. No requirements for maintenance. No plan for life cycle data delivery.
Valtari Highway 12	Alliance	fair	weak	Cooperation enabled between all stakeholders. Not enough construction specific applications in use. No straight requirements from maintenance stakeholders. No plan for life cycle data delivery.

In Table 3, the different applications that were used in each project are listed. There were many applications in use, but they did not all fit as infrastructure projects. There were interoperability challenges, and some applications had quality problems. When the project delivery type enabled the collaboration between stakeholders, the quality of the data flow was fair. The lack of preliminary data flow planning prevented better data flow quality.

Data governance is a concept in which an organisation manages the quality, consistency, usability, security and availability of its data. Cohen's (2006) roles for data governance in the case projects are identified in Table 4. It is interesting to note that the case projects themselves were not able to identify these roles (e.g. typical comment: Roles are unclear for the constructors). There were some related roles in the projects, like BIM manager, BIM coordinator and BIM consultant, but these roles were not adequate for effective life cycle data governance. If the construction industry decides to use PDM, there need to be data governance roles. The naming of responsible persons and the specification of their tasks could be part of the project planning.

In the BuildingSmart Finland (2015). (Common InfraBIM Requirements), there is a general description of roles in data model–based infra-projects. There is also a list of initial data models and models that are the results of each project phase. In the case studies, these roles and guidelines were not clear and not comprehensively used. The

technical data modelling in infra-projects is one part of the total data management challenge. In addition to design and construction models, data model should cover the needs of related business processes, like customer–stakeholder relations, production, maintenance, HR and finance, more comprehensively.

Table 3. Summary of the ICT systems that were in use in case projects.

Case study	Summary of the ICT systems in use
Jännevirta Bridge	• Application for editing automated machinery surfaces • Application for scaffold design • Application for 3D DWG managing • Application for managing IFC and XML data transfer formats • Application for collaboration models and clash detection • Application for counting the number of trucks • Application for quality assurance and control • Application for volume calculation
Keropudas Bridge renewal	• Application for managing IFC and XML data transfer formats • Application for structures and rebar design • Application for scaffold, shifting and casting • Application for highway design, editing and management
Valtari Highway 12	• Project portal • Application for highway design, editing and management. Also for automated machinery surface editing, • Application for managing IFC and XML data transfer formats

In Table 4, there is a suggestion of data governance roles in infra-projects. The client has the challenge of establishing the project goals, bonus and sanctions so that all stakeholders can make decisions for the best of the project. Each project stakeholder has its own organisation with its own company-specific goals and bonus systems. The client's duty is to harmonise project incentives so that stakeholders win only when the whole project wins. Otherwise, it is difficult to reach project goals. The same applies to data governance roles: the project data stewards have to plan data governance for the benefit of the whole project, not only for their individual benefit.

Table 4. Data governance roles in the different phases of an infra-project.

Project phase role	Preliminary planning	Design	Execution	Operation	Maintenance
Data Steward	Client's CTO or committee of main stakeholders' CTOs	Client's CTO or committee of main stakeholders' CTOs	Client's CTO or committee of main stakeholders' CTOs	Client's CTO or committee of main stakeholders' CTOs	Client's CTO or committee of main stakeholders' CTOs
Data Owner	Client's department leader	Design stakeholder's project manager	Construction stakeholder's project manager	Operation stakeholder's responsible leader	Maintenance stakeholder's responsible leader
Data Manager	Principal designer of preliminary planning	Principal designer	Surveying foreman / site engineer	Owner's traffic manager	Maintenance foreman
Data User	Pre-engineer	Designer	Foreman / site foreman	Traffic contractor	Maintenance workers

5. MODEL FOR THE INFORMATION FLOW

The main reason for problems in the data flow of infra-construction projects stems from the lack of organisation of the flow of data for the entire life cycle. Typically (and also supported by our case studies), projects do not have a plan for how to organise, monitor and control data during the life cycle. This partially leads to separated data silos or bunkers. Silos seem to emerge in the preplanning and planning phases, causing more silos to emerge in

the building phase. It seems to be rare to plan for the data needs of the maintenance phase. Problems in data quality and usability are consequences of proactive planning.

Based on earlier research and our case studies, we created a life cycle model for data flow (FIG. 5). The origin for data management is defining the product — the object whose lifecycle we try to optimise and manage — and the product data. All other partial products, services or combination of both during the life cycle must ultimately serve this purpose if we want to minimise total cost versus benefits. The information flow between processes should be twofold (e.g. to provide design data from a CAD system to PDM and ID (identifier) data from PDM to CAD, enabling E-BOM for the as-designed object). The idea of the PDM process is to serve as an evolutionary data repository for planned, designed, built, sold, used and maintained built objects — as the "DNA" of built objects. There must be one centralised PLM/PDM system to act as a central data repository for the product master data; otherwise, there will be no practical possibilities to achieve updated quality for data. Naturally, this requires a predefined product structure for the built object based on what the stakeholders contribute during the life cycle. Logically, this also means that even when different stakeholders in the life cycle utilise different systems (CAD, BIM, CRM, ERP, etc.), they must comply and connect on one product data, when utilising it and when modifying it, according to their roles and responsibilities.

As the up-to-date product data form the first step of the data flow during the life cycle, other types of data are also required. Modern integrated business processes like construction and maintenance transfer a lot of business-related data for improving activities. The ERP system mainly serves the purposes of the construction process by controlling the progress of work, sourcing and subcontracting. In order to operate, ERP requires as-designed data from PDM and ERP systems. This might partially work in some projects, but complete utilisation is generally rare. Production as a stakeholder should have a significant role in the design phase to contribute to the sense of "design for construction". This would naturally improve the usability of product data for construction.

It is typical in the construction business that use and maintenance are operated by different organisations than design and construction. It is also typical that, after construction, product data is shared in paper documents from design team. These documents typically lose intelligence created in design and, in the worst case, these documents are not fully contained as-built data. It is very rare that maintenance systematically contributes to design in "design for maintenance" (this does not mean that designers think of objects in the sense of maintenance). The operations and systems in use and maintenance could improve overall efficiency if they were involved in the design and construction phases. It is surprising that data is defined in design and defined again in construction but is usually lost and need to be rediscovered in the maintenance phase (materials, components, maintenance instructions and supplier information). This is especially critical when failures or malfunctions occur; with the right data, reparations can save a lot of time and money once the detailed component or material level data is available "as maintained".
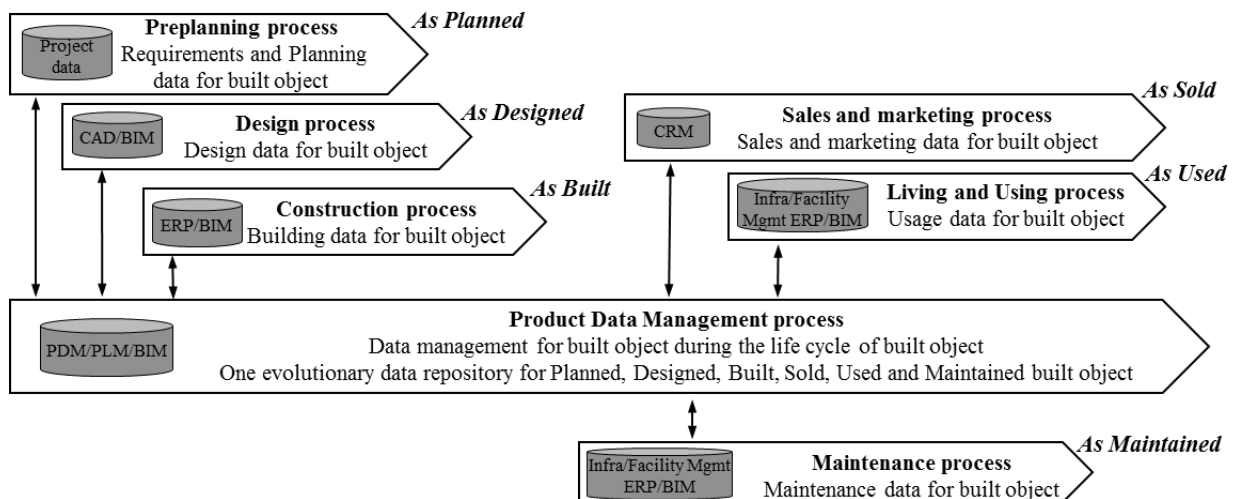

FIG. 5. Organising and managing the flow in an infra-project.

Master data management is a set of the best data management practices that organise key stakeholders, participants and business clients (Loshin, 2009) for the data management of the entire life cycle of a product (built object). A central repository for all product-related data gives the project participants (temporary organisation) a single version of the truth for their product master data, enabling improved operational processes for business efficiency. In MDM, different organisations in project and information systems cooperate to harmonise, cleanse, publish and protect mutual data assets to be shared across the project organisation (see White et al., 2006). The first step in data integration is defining product master data, since all stakeholders use product data in some phase of the life cycle, noting that there are several steps in true utilisation.

The main purpose of data governance is to ensure that data models, data quality and actions done are managed within the project (built object). There are several players that must participate in the governance work. Usually, these players represent the data owner, IT and actual data maintenance persons. The data maintenance process is very operative and is linked to other main processes. Here, the starting point should be the minimum viable product (minimum viable built object) containing the necessary data for the life cycle.

6. DISCUSSION AND CONCLUSIONS

Data and data management provides significant possibilities to improve efficiency also in infrastructure projects. Avoiding sub-optimisation and fragmentation of the data is one of the clear improvement opportunities. Data and digitalisation, naturally, must be servants, not the master. They can provide significant efficiency improvements but still the integrated project model is the entity to be optimised. However, data must be available, correct and stored in one common repository for the project's life cycle – unified for stakeholders. All this begins with generally understood definitions for data and data processes. If we look at the business processes and the product, it is obvious that the product (built object) is the starting point. It is also the object that needs to be planned, built, utilised and maintained.

According to our case study as the main challenge of the data flow of the infra-project life cycle reveal that data needs are not planned beforehand. There are also difficulties receiving existing data and when received data must be edited before use. These can be tackled by carefully planning the data flow from the lifecycle perspective — especially from a maintenance perspective. One interesting observation from the case projects is that all participants discussed "infra-projects" (focusing on planning, design and construction), even when the life cycle of the road or bridge was significantly longer. However, changing view for the entire life-cycle is not easy since the industry has operated with a certain logic for more than 30 years and is resistant to change.

We also noted that data silos are partially due to existing contact models not supporting the initiatives to pursue life cycle data management even when theory says to do so (it does not matter who's Excel is right – they are all equally wrong and outdated). Relational contract models and life-cycle contracts are partially driving for integration and collaboration also in data issues (see e.g. Aapaoja et al., 2013, Herrala et al., 2012, Hietajärvi et al., 2017, Pekuri et al., 2013). Not causing contractual barriers for data and data flow. Thera are different national regulations driving for unified data flow. Maybe the only practical possibility is to follow the UK practice in which a customer requires as-built data models for use and maintenance phases. The current practice is not optimal overall because customers ultimately pay for the data waste and other waste caused by inefficient processes. The other way would be that the companies in construction or infra industry re-engineer their processes and even business models enabled by more efficient data management (see e.g. Aapaoja et al., 2013, Pekuri et al., 2013). Data assets might be able to provide novel value for the customers or at least considerable cost efficiency.

Based on the earlier research and our case studies, we have created a life cycle model for data flow from a data and data utilisation perspective. The origin for data management is defining the product — the object whose life cycle we try to optimise and manage — and the product data. The idea of a PDM process is to serve as an evolutionary data repository for planned, designed, built, sold, used and maintained built objects: the "DNA" of the built object. No matter what the excuse, data needs to be stored in one place to achieve good data quality, even in theory. There are a lot of improvement possibilities, but we should start with very simple steps, like in FIG. 5. The model needs to reveal the roles for data governance presented by Cohen (2006) (data steward, owner, manager and user) or the RACI model (responsible, accountable, consulted and informed stakeholders related to data) (Costello, 2012). The evolutionary phases of an infra-project life cycle increase the complexity of these but still need to be resolved. Data must have an owner, no matter what phase it is in, and the process needs to be carefully planned infra- by infra-project.

Our study discusses findings from three in-depth case studies and is rooted in an extensive literature review. Our findings are in line with earlier data management research. According to our case studies infrastructure projects are still surprisingly behind practices of other industries (see, e.g. Silvola et al. 2011, 2016, 2019). The achievable benefits are not only cost savings, but increased value provided during the life cycle. Our research continues to explore the knots of data management in project-oriented business, noting that we had only three cases in our study however, we were leaning on large generic data management literature.

REFERENCES

A Guide to the Project Management Body of Knowledge (PMBOK). (2000). Project Management institute, Four Campus Boulevard, Newtown Square, USA.

Aapaoja, A., Herrala, M., Pekuri, A. & Haapasalo, H. (2013). The Characteristics of and Cornerstones for Creating Integrated Teams. *International Journal of Managing Projects in Business*, 6(4), 695 - 713.

Aapaoja, A., & Haapasalo, H. (2014). A Framework for Stakeholder Identification and Classification in Construction Projects. *Open Journal of Business and Management,* 2(1), 43–55.

Ameri, F. & Dutta, D. (2005). Product Lifecycle Management: Closing the Knowledge Loops. *Computer-Aided Design and Applications* 2 (5), 577-590.

American associated general contractors of America AGC. (2006). Contractors' Guide to BIM 1 (2006) 3.

Antwi-Afari, M., Li, H., Pärn, E. & Edwards, D. (2018). Critical success factors for implementing building information modelling (BIM): A longitudinal review. *Automation in Construction,* 91, 100–110.

Apostol, C-G. (2007). Enterprise master data management trends and solutions. *Informatica Economică,* 3(43), 35-38.

Autodesk. (2002). Building information modeling, Technical report, Autodesk Inc., San Rafael, California, United States.

Azhar, S. (2011). Building information modeling (BIM): Trends, benefits, risks, and challenges for the AEC industry. *Leadership and management in engineering,* 11(3), 241–252.

Azhar, S., Khalfan, M. & Maqsood, T. (2012). Building information modelling (BIM): now and beyond. *Construction Economics and Building,* 12(4), 15–28.

Barlish, K. & Sullivan, K. (2012). How to measure the benefits of BIM—a case study approach. *Automation in construction,* 24, 149–159.

Berson, A. & Dubov, L. (2007). Master Data Management and Customer Data Integration for a Global Enterprise, McGraw-Hill, New York, NY.

Bryde, D., Broquetas, M., & Volm, J. (2013). The project benefits of Building Information Modelling (BIM). *International Journal of Project Management,* 31, 971–980.

BuildingSmart Finland. (2015). Yleiset inframallivaatimukset YIV2015, Osa 1, Tietomallipohjainen hanke.

BuildingSmart. (2017). BuildingSMART International. Retrieved from https://www.buildingsmart.org/. [4 Dec 2018]

Cohen, R. (2006). What's in a Name? Data Governance Roles, Responsibilities and Results Factors. Retrieved from: https://www.information-management.com/news/whats-in-a-name-data-governance-roles-responsibilities-and-results-factors. [11 Jan 2018]

Costello T. (2012). RACI—Getting Projects "Unstuck". *IT Professional,* 14(2), 64-63.

Crnkovic, I., Asklund, U. & Persson-Dahlqvist, A. (2003) Implementing and Integrating Product Data Management and Software Configuration Management. Artech House, Norwood, MA.

Dakhil, A., Underwood, J. & Al Shawi, M. (2019). Critical success competencies for the BIM implementation process: UK construction clients. *Journal of Information Technology in Construction*, 24(5), 80-94.

Dayton, M. (2007). Strategic MDM: the foundation of enterprise performance management. *Cutter IT Journal*, 20(9),13-17.

Eadie, R., Browne, M., Odeyinka, H., McKeown, C. & McNiff, S. (2013). BIM implementation throughout the UK construction project lifecycle: An analysis. *Automation in Construction,* 36, 145–151.

Eastman, C., Teicholz, P., Sacks, R. & Liston, K. (2008). BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors. John Wiley & Sons, Hoboken.

Eastman, C., Teicholz, P., Sacks, R. & Liston, K. (2011). BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors. Wiley. Kindle Edition. Location 484.

Eisenhardt, K.M., (1989). Building theories from case study research. *Academy of management review*, 14(4), 532-550.

Finnish Transport Agency (FTA). (2010). Tiensuunnittelun kulku. Retrieved from: https://www.liikennevirasto.fi/. [1 Dec 2018]

Friedrich, R., Groene, F., Koster, A., Maekelburger, B. (2013) The 2012 industry digitization index. Strategy&Press Room. Retrieved from: http://www.strategyand.pwc.com/media/. [1 Dec 2018]

Ghaffarianhoseini, A., Tookey, J., Ghaffarianhoseini, A., Naismith, N., Azhar, S., Efimova, O. & Raahemifar, K. (2017). Building information modelling (BIM) uptake: Clear benefits, understanding its implementation, risks and challenges. *Renewable and Sustainable Energy Reviews,* 75, 1046–1053.

Godfrey B (2002) Managing Information Quality – A Critical Process for Most Organizations. *Proceedings of the Seventh International Conference on Information Quality, Cambridge: Massachusetts Institute of Technology*

Haapasalo, H. (2000). Creative Computer Aided Architectural Design - Internal Approach to The Design Process. Department of industrial engineering. Acta Universitatis Ouluensis C 149. Dissertation. University of Oulu.

Haapasalo, H. (2018). Käyttäjälähtöinen tiedonhallinta. In eds. Haapasalo H, Aaltonen K, Kähkönen K. & Saari A. 2018. Rakentamisen Integraatiomekanismit. *Research reports in Industrial Engineering and Management*, 1/2018. pp. 48-61.

Halttula, H., Haapasalo, H., Aapaoja, A. & Manninen, S. (2017). Early involvement and integration in construction project – Benefits of DfX. International Journal Knowledge, Management and Learning, 6(2), 216-237.

Halttula, H., Haapasalo, H., & Herva, M. (2015). Barriers for achieving benefits of BIM. *International Journal of 3-D Information Modeling,* 4(4), 16-33.

Herrala, M., Belt, P., Haapasalo, H. & Pakkala, P. (2012). Organising infrastructure networks – two descriptive case studies on value creation. *International Journal of Services, Economics and Management*, 4(2), 104–121.

Hietajärvi, A-M., Aaltonen, K. & Haapasalo, H. (2017). Managing integration in infrastructure alliance projects: dynamics of integration mechanisms. *International Journal of Managing Projects in Business*, 10(1), 5 – 31.

HM Government. (2015). Digital Built Britain - Level 3 Building Information Modelling - Strategic Plan. Retrieved from: https://www.cdbb.cam.ac.uk/news/2015DBBStrategy. [3 Dec 2018]

Investopedia. (2018). Infrastructure. Retrieved from: https://www.investopedia.com/terms/i/infrastructure.asp. [1 Dec 2018]

Khatri, V. & Brown, C. (2010). Designing Data Governance. *Communications of the ACM*, 53(1), 148-152.

Kropsu-Vehkapera, H., Haapasalo, H. & Vahtiala, M. (2010). Exploiting product data management through product lifecycle for intangibles, *Proceedings of the 17th Annual EurOMA Conference, Managing Operations in Service Economies, June* 6. – 9.2010, Porto, Portugal.

Kropsu-Vehkapera, H., Haapasalo, H., Harkonen, J. & Silvola, R. (2009). Product data management practices in high-tech companies. *Industrial Management & Data Systems*, 109(6), 758-774.

Kunz, J., & Fischer, M. (2012). Virtual Design and Construction: Themes, Case Studies and Implementation Suggestions. CIFE Working Paper #097, 1.

Lankhorst, M. (2009). Enterprise Architecture at Work: Modelling, Communication and Analysis. 2nd Ed. Springer-Verlag.

Loshin, D. (2001). Enterprise Knowledge Management: The Data Quality Approach. Academic Press, Morgan Kaufmann Publishers.

Loshin, D. (2009). Master Data Management. Morgan Kaufmann, Burlington, MA.

Love, P., Edwards, D., Irani, Z. (2008) Forensic project management: an exploratory examination of the causal behavior of designinduced rework. *IEEE Transactions on Engineering Management,* 55(2), 234–247.

Merschbrock, C. & Munkvold, B. E. (2012). A research review on building information modeling in constructionan area ripe for is research, Association for Information Systems.

Morris, P. (1988). Managing Project Interfaces: Key Points for Project Success. In Cleland & King, Project Management Handbook, 2d ed. Englewood Cliffs, NJ,: Prentice-Hall.

Moss, L. (2007). Critical success factors for master data management. *Cutter IT Journal*, 20(9), 7-12.

NBS (2016). International BIM report 2016, Technical report, RIBA Enterprises Ltd.

NBS (National Building Specification) (2017). What is COBIE". Retrieved from: https://www.thenbs.com/knowledge/what-is-cobie. [12 Dec 2018]

NIST. (2004). Cost analysis of inadequate interoperability in the U.S. capital facilities industry, in: A.C. Gallaher, J.L. Dettbarn Jr., L.T. Gilday (Eds.), M.P.O.C, National Institute of Standards and Technology.

Otto, B. & Huner, K. (2009). Functional reference architecture for corporate master data Management. Working paper 21, Institute of Information Management, University of St Gallen, St Gallen.

Pekuri, A., Haapasalo, H. & Herrala, M. (2011). Productivity and performance management – managerial practices in construction industry. *International Journal of Performance Measurement*, 1(1), 39-58.

Pekuri, A., Pekuri, L. & Haapasalo, H. (2013). The Role of Business Models in Construction Companies. *Australasian Journal of Construction Economics and Building*, 13(3), 13-23.

Peltonen, H. (2000). Concepts and an implementation for product data management. Acta polytechnic Scandinavica, Espoo.

Penttilä, H. 2006. Describing the changes in architectural information technology to understand design complexity and free-form architectural expression. *Journal of Information Technology in Construction,* 11(29), 395–408.

Rachuri, S., Subrahmanian, E., Bouras, A., Fenves, S., Foufou, S. & Sriram, R. (2008). Information sharing and ex-change in the context of product lifecycle management: Role of standards. *Computer Aided Design,* 40(7), 789-800.

Rostami, A. & Oduoza, C.F. (2017). Key risks in construction projects in Italy: contractors' perspective Engineering, Construction and Architectural Management, 24(3) 451-462.

Saaksvuori, A. & Immonen, A. (2008). Product Lifecycle Management. 3rd Ed. Berlin Heidelberg: Springer-Verlag.

Silvola, R. (2018). One product data for integrated business processes. University of Oulu Graduate School; University of Oulu, Faculty of Technology Acta Univ. Oul. C 686,

Silvola, R., Kemppainen, T., Haapasalo, H., Kropsu-Vehkapera, H., & Jaaskelainen, O. (2011). Elements and implementation of product data ownership network, Proceedings of TIIM2011 Conference. Oulu Finland.

Silvola, R., Harkonen, J., Vilppola, O., Kropsu-Vehkapera, H. & Haapasalo, H. (2016). Data quality assessment and improvement. *International Journal of Business Information Systems,* 22(1), 62–81.

Silvola, R., Jaaskelainen, O., Kropsu-Vehkapera, H. & Haapasalo, H. (2011). Managing One Master Data - Challenges and Preconditions. *Industrial Management & Data Systems*, 111(1), 146–162.

Silvola, R., Tolonen, A., Harkonen, J., Haapasalo, H. & Männistö, T. (2019). Defining One Product Data for a Product. *International Journal of Business Information Systems, accepted, forthcoming*.

Stark, J. (2005). Product lifecycle management: 21st Century Paradigm for Product Realisation. London Limited: Springer-Verlag.

Succar, B. (2009). Building information modelling framework: A research and delivery foundation for industry stakeholders. *Automation in Construction,* 18, 357–375.

Svensson, D. & Malmqvist, J. (2002). Strategies for product structure management in manufacturing firms. *Journal of Computing and Information Science in Engineering,* 2(1), 50-58.

Weber, K., Otto, B. & Österle, H. (2009). One Size Does Not Fit All - A Contingency Approach to Data Governance. *Journal of Data and Information Quality*,1(1), 1-27.

Wende, K. (2007). A Model for Data Governance – Organising Accountabilities for Data Quality Management. 18th Australasian Conference on Information Systems.

White, A., Newman, D., Logan, D. & Radcliffe, J. (2006). Mastering Master Data Management. Research, Gartner Inc.

World Economic Forum. (2018). An action plan to accelerate building information modeling (BIM) adoption, Technical report, World Economic Forum, Geneva, Switzerland. REF 130218 - case 00041654 V2.

Yan, H. & Demian, P. (2008). Benefits and barriers of building information modelling. Proceedings of the 12th *International Conference on Computing in Civil and Building Engineering* (ICCCBE XII) & 2008 *International Conference on Information Technology in Construction* (INCITE 2008), Beijing, China, 16th-18th October 2008.

Yang, X., Moore, P.R., Wong, C-B., Pu, J-S. & Chong, S.K. (2007). Product lifecycle information acquisition and management for consumer products. *Industrial Management & Data Systems*, 107(7), 936-953.

Yin, R.K. (2003). Case Study Research: Design and Methods, 3rd ed., SAGE Publications, Beverly Hills, CA.