Issues and challenges of knowledge management in online open source hardware communities

Jason Xinghang Dai, Jean-François Boujut, Franck Pourroy, Philippe René Marin

Design Science

Issues and challenges of knowledge management in online open source hardware communities

Jason Xinghang Dai[1], Jean-François Boujut [iD][1], Franck Pourroy[1] and Philippe Marin[1]

1 *CNRS, Grenoble INP, G-SCOP, Université Grenoble Alpes, 38000 Grenoble, France*

Abstract

Knowledge management (KM) is the process of creating, sharing, using and managing the knowledge and information of an organization . In this paper, we investigate KM practices and needs in Open Source Hardware (OSH) communities. The aim is to provide insights into the KM issues and challenges facing these communities. Our analysis is based on interviews carried out with the participants of 22 projects. Interview transcriptions were coded and tagged with concepts extracted from the literature. Text mining was used to uncover the main concepts embedded in the corpus. From this analysis, knowledge sharing emerged as one of the top-rated concepts. Codification and personalization also proved to be important KM approaches, both requiring support in the OSH design process. Using a dendrogram, we highlighted the benefits and challenges of codification together with some interesting concept associations. High contributor turnover, little or no standardization and weak project structuring are still key challenges for OSH communities when it comes to ensuring awareness, making decisions and sharing knowledge.

Key words: Open Source hardware, Open Design, Knowledge management, Knowledge codification, Knowledge personalization

1. Introduction

In the 1990s, academics drew attention to the fact that knowledge would become the main source of wealth in the future, replacing capital in the new economy (Stuwart & Ruckdeschel 1998). This transition is already taking place, especially in the engineering design industry. Today, engineering design has become highly knowledge-intensive and collaborative. The stakeholders of a product development project are usually geographically dispersed, while knowledge sharing through digital information systems has become an essential project success factor (Ouertani *et al.* 2011). One of the key challenges for engineering design project knowledge management (KM) is the difficulty of capturing design knowledge. In order to preserve product development experience and share it afterwards, designers are required to formulate and explicitly document their design intentions. Three knowledge sharing barriers were identified by (Riege 2005): individual, organizational and technical. On an individual level, despite the importance and necessity of knowledge documentation in the engineering design industry, people find it too difficult and time-consuming (van der Ven *et al.* 2006). On an

organizational level, a company may lack specific KM strategies and, on a technical level, a barrier can be thrown up by poor information technology (IT) infrastructure to support design knowledge sharing (Huysman & Wulf 2006). Since the emergence of computer-aided design (CAD) and computer-aided manufacturing (CAM), IT tools have played an increasingly important role in the engineering design industry. Complex design projects are usually carried out by multiple, geographically distributed stakeholders. IT and the Internet are essential to enable distributed collaboration. Many researchers have emphasized the importance of information technology for knowledge sharing in an organization. IT tools can be used to capture design rationale (Aurisicchio & Bracewell 2013; Rohde, Štorga & Marjanović 2015), to facilitate interpersonal communication (Ackerman *et al.* 2013) and to support collaborative design (Chen, Chen & Chu 2008). However, in spite of both academic and industrial efforts to codify knowledge using IT tools, most of the knowledge in an organization lies within the memory of each individual. Thus, complementary to knowledge codification, face-to-face socialization still plays a crucial role in knowledge sharing.

The open source principles have expanded from software design to hardware design in the last decade. Powered by affordable machine tools, free online collaborative work platforms and abundant design education web resources, numerous maker communities have emerged across the globe (Voigt, Montero & Menichinelli 2016). The Maker Movement blurs the borderline between professionals and amateurs, maximizing knowledge sharing, accelerating innovation and bringing a 'do-it-yourself' culture back to daily life and work (Zhuoxuan *et al.* 2017). The emergence of maker communities is linked to the birth of open source hardware (OSH). This new phenomenon has attracted the attention of many researchers although the terminology applied to it has not yet stabilized. In this article, we shall use the term OSH to refer to the open source characteristic of tangible products.

The new OSH paradigm has upset the traditional institutionalized vision of knowledge building and transfer. Rather than being based on professional communities of practice, today valuable knowledge is formed and passed on thanks to a distributed, peer-production process (Powell 2015), which is why IT is vital to the development of open source hardware. Knowledge in an OSH community should be documented and shared among the makers in order to scale up collaborative design. According to the principles of OSH, members of the community build a knowledge sharing platform where they create, access and share OSH related knowledge. However, in traditional engineering design practice, it is still quite challenging for designers to document their design rationale, especially for novice designers whose limited practical experience makes it difficult for them to capture and share relevant knowledge. This topic was already recognized as important in the 1990s (Klein 1993; Lee 1997) and the authors point out the benefits of design knowledge capture to support collaboration and reuse through better documentation and design supports. The effort to provide codification support by Klein (1997) has been materialized in a design rational capture system for improving knowledge codification and Bracewell *et al.* (2009) proposed a tool based on the IBIS model for structuring design rational and design knowledge capture. These researches also show that IT tools cannot entirely replace the socialization process that leads to knowledge sharing and that members of design teams still rely heavily on being able to directly access to colleagues and experts' knowledge.

The aim of this research is to investigate the barriers for knowledge sharing in an OSH project. In other words, what are the specific challenges of open source hardware design to support both socialization and codification strategies?

To answer this question, we shall first describe the context surrounding OSH and present the state of the art in KM for open source product development. We shall then identify the specific challenges of KM in OSH projects based on a review of the literature. Next, we shall introduce a corpus of interviews with OSH practitioners, which we shall then proceed to analyse both quantitatively and qualitatively. Finally, we shall outline the insights drawn from our analysis and present our findings regarding the requirements for knowledge support for OSH projects.

2.   Context of OSH and research question

Although OSH has undergone rapid recent development and gained attention both from practitioners and researchers, it is still in its infancy. Today, innovation in the digital world of 'bits' is beginning to spill over into the physical world of 'atoms'. In this section, we shall present the context of open source hardware design and outline the main characteristics of OSH. We shall then overview the literature related to the challenges of KM for open source products. As research on OSH is still limited, we shall also refer to the literature on OSS research in order to explore the nature of open source products in general. Finally, we shall identify the specific challenges of KM in OSH projects.

2.1.   Open Source Hardware

Open Source Hardware (OSH) can be defined as tangible artefacts: machines, devices or other physical things whose design has been released to the public in such a way that anyone can make, modify, distribute and use that design.[1] Despite its considerable growth, OSH is a relatively young phenomenon that has only emerged over the last decade (Raasch, Herstatt & Balka 2009). Like open source software (OSS), the development of OSH relies heavily on the Internet and IT platforms (Fjeldsted *et al.* 2012). However, compared with software, which is entirely digital, hardware products are physical objects comprising a broad spectrum of complex features that cannot all be digitally translated (Raasch *et al.* 2009; Howard *et al.* 2012).

Despite its open source nature, OSH is not necessarily based on participatory product design. Depending on the product's complexity and the initiator's wish to foster collaborative design, the level of collaboration within an OSH community varies. For example, Thingiverse is a platform where 3D printing design files can be shared by individuals or organizations. Users of Thingiverse can download these design files and, for example, leave comments, but they are not able to participate in the product development itself, rather the user can create a 'remix' of the product by creating a modified version of the object. Github, on the other hand, is a platform originally designed for collaborative OSS development and is frequently used for OSH development. Users of Github can download OSH design files but also participate in the development of the product by 'forking' and 'pulling' the platform's project repository. For collaborative OSH development

---

[1]https://www.oshwa.org/definition/.

projects, individuals proactively self-organize, choosing their own roles, tasks and period of contribution (Maher, Paulini & Murty 2010). Not all the members in an OSH community are 'designers' or 'developers' contributing directly to product development. Some are followers, replicators or community managers (Bonvoisin, Thomas, *et al.* 2017). Different roles will naturally result in different knowledge sharing behaviour.

2.2. Knowledge management and open source products

Creating and sharing knowledge is essential for fostering innovation and is the key challenge of the knowledge-based economy (Pawlowski & Bick 2015). KM has been identified as one of the key enabling factors for distributed engineering organization (McMahon, Lowe & Culley 2004; Vuletic *et al.* 2018). KM strategies for engineering design can be generally divided into two categories: codification and personalization (Hansen, Nohria & Tierney 1999). The codification strategy involves codifying knowledge. This means that knowledge is externalized and converted into a form allowing storage in a database so that it can then be accessed and reused by people in an organization. The personalization strategy encourages the sharing of knowledge through person-to-person socialization. This is supported by an IT infrastructure so that the right person is connected to the right problem. Recent scientific works in KM promote knowledge sharing within or across organizations through management and technologically-oriented approaches. According to Powell (2015), our contemporary communication environment has created greater knowledge complexity leading to changes in knowledge access and distribution modes and the way we collaborate using knowledge. In turn, these changes have increased the openness of knowledge production environments. The Internet and IT platforms are intrinsic to OSH projects. OSH project knowledge can be found in various digital forms: project specification web pages, technical file repositories, online forums, wikis, and so on. Since OSH practice relies heavily on online IT platforms, it is imperative for contributors to document their knowledge in order to share it with others. The collaborative design mode is preconditioned by the free access to knowledge produced in an OSH project (Aitamurto, Holland & Hussain 2015). This means that KM is crucial for open source product development. Documenting product development has proven to offer several advantages (Lee 1997): it provides better design support for designers, better product maintenance support and greater learning support. However, managing knowledge in open source product development has its own challenges, as we shall see.

2.2.1. *Learnings from open source software*

It is also useful to consider the characteristics and challenges of OSS before attempting to identify those pertaining to OSH. OSS is characterized as being intensely people-oriented (Lethbridge, Sim & Singer 2005) and knowledge-intensive (Giovan Francesco & Michle 2003). The process of open source product development usually involves the continuous modification and improvement of a current design. OSS projects can be extremely complex, and the knowledge needed for the software development process is very broad and unlikely to be held by any one software developer or small group of developers. Understanding exactly how knowledge is produced, shared and reused during the process of OSS development is a challenge in itself (Sowe, Stamelos & Angelis 2008). Licorish & MacDonell (2014)

argue that open source projects are usually led by a core team comprising a few individuals in charge of project communication, source code changes and major decision-making. This has been also shown by Boujut *et al.* (2019), who studied the distribution of design activities among the communities. Hence, knowledge is mainly created by this core team and shared among its members.

However, open source projects also entail a 'do-ocracy' work process, that is, a governance model allowing anyone to initiate ad-hoc solutions as long as they are willing to provide them. In this case, the knowledge used for decision-making is implicit (Bonvoisin, Thomas, *et al.* 2017). Any individual can make a change to an OSH project without providing justifications. This change can then be overridden by other contributions. However, the reasons behind such changes are not given. According to Rigby *et al.* (2016), one of the challenges of OSS project KM stems from the close relationship between the author and the authored source code. This means that a software development project can suffer knowledge loss if authors leave the project with their code unfinished. Their replacements tend to be less productive since they have to work on an unfamiliar code base. This phenomenon, called 'turnover', also applies to OSH projects as we will see later. The open source product work force (usually unpaid volunteers) is transient leading to a high turnover rate and project knowledge loss (Rashid, Clarke & O'Connor 2017). So, like OSS projects, OSH projects are also knowledge-intensive and driven by a transient work force usually composed of volunteer makers. The OSH development process can be organized to resemble that of OSS to a considerable degree (Raasch 2011). Therefore, it can be assumed that OSH projects will inherit the same KM challenges as OSS projects.

2.2.2. *Specific questions raised by open source hardware*

Furthermore, due to the physical nature of tangible products, another particular KM challenge for OSH projects lies in the difficulty of sharing knowledge related to physical product development using digital IT platforms (Boisseau, Omhover & Bouchard 2018; Chandrasegaran *et al.* 2013).

In short, three main research questions can be identified for OSH projects related to KM challenges:

(i) The OSH community is usually largely self-organized and comprises members occupying a diversity of roles. However, according to the literature, an OSH project is often dominated by a core team made up of a few people in charge of major project decisions and that is carefully managed by the project coordinator (Raasch 2011). Project contributors who are not part of the core team have limited access to knowledge production and knowledge sharing. As Aitamurto *et al.* (2015) argue, the data produced during the design process should be available as open data in order to expand knowledge access. Therefore, is there a challenge in scaling up knowledge sharing across the entire OSH community?

(ii) An open source project team is usually initiated and steered by the core team while there is a high turnover among other volunteer participants (Foucault *et al.* 2015). This constant organizational change leads to knowledge loss (Rashid *et al.* 2017), since not all contributors are aware of the overall project status and other contributors' activities (Treude & Storey 2010). Does this turnover issue also apply to open source hardware?

(iii) Since OSH development aims to produce tangible products (Müller-Seitz & Reger 2010), do 'offline' activities documentation such as prototyping and testing represent an issue for OSH developers?

Limited research has been performed to address the KM challenges in OSH projects. This paper contributes to better understanding of how OSH communities manage their project knowledge before suitable KM solutions can be put forward. A systematic analysis of OSH practitioners' interviews highlighted interesting issues raised by OSH specificities, such as volunteer-based participation, high turnover or very heterogeneous profiles and motivations of the participants. In turn we found a significant commitment to information structuring, documentation sharing which face important challenges due to the specific context of OSH. Socialization could be seen as an interesting alternative. Unfortunately, the geographical dispersion of the participants seriously hinders the process. The paper supports the idea that KM approach could provide an interesting way to overcome the limitations observed in this study. Based on the background, characteristics and challenges outlined above, we shall put forward and explore several concepts related to OSH KM.

3. Data processing methodology

In this study, we applied both a qualitative and quantitative approach. Our research was based on a corpus of interviews with OSH practitioners collected as part of the research project OPEN![2] (Bonvoisin *et al.* 2017, 2018; Bonvoisin, Mies, Boujut, & Stark, 2017; Bonvoisin, Thomas, *et al.* 2017). This interview corpus served as a basis to understand the OSH phenomenon and provide empirical data for OSH research. Through the corpus we explored several themes, including KM in an OSH project. Since this paper focuses mainly on KM, excerpts annotated with concepts related to KM were extracted and used to build a KM themed interview corpus. We have proceeded in three steps:

 (i) In a first step, we have collected the results of the 23 interviews (from 22 projects) carried out within the project OPEN! as exposed in Section 4 (see Table 1).
(ii) In a second step, we have identified a subset of answers excluding irrelevant questions (business, financial or legal aspects), which is addressed in Section 5.2.
(iii) In a third step, we have carried out an exploratory quantitative study which is exposed in Section 5.3.

According to Kohlbacher (2006) the qualitative contents analysis is a very interesting method if the contents of the text is highly semantic and if the research question is relatively focused. On the other hand, the authors consider it as not suitable for open-ended and explorative research questions. In our case, we considered the KM-related question as specific enough to apply the method. Other limitations of qualitative analysis obviously stem in the non-fully replicable

---

[2]OPEN! was a joint research project initiated by TU Berlin and Grenoble INP which aimed at studying open source hardware product development processes in order to characterize the practices and needs for supporting tools.

Table 1. List of OSH communities interviewed in the OPEN! project

Name	Aim of the project	Profile of the interviewees	Name	Aim of the project	Profile of the interviewees
Arbalet	The aim of the project is to create pixel art-works based on visual programming and a physical flat surface filled with LED.	Computer scientist	L'atelier s'adapte	Develop usable and sustainable solutions for disabled people.	IT entrepreneur
Apertus Axiom	Open hardware for versatile scientific high-end camera.	Multimedia expert, IT specialist	Open Source Ecology	The interviewee chose to discuss a brick press project as subject for the interview as it is the most advanced project of OSE to date.	Founder, Physicist and product designer
EchOpen	Portable sonographic imaging equipment.	Physicist, worked in humanitarian organizations	Open Source Ecology Germany	OSE Germany is described as a movement or an organization basically performing three types of activities in the context of energy, agriculture and all the basic needs.	Engineer in renewable energy
E-Nable	Worldwide community that produces customized prosthetic device for children or amputees needing an artificial hand for very simple usage.	IT consultant	Open source BeeHives	Open Source beehives develops and provides beehives entirely open and available either through downloads or buy purchase.	Architect, fablab coordinator
Fair Cap	Open source 1$ antibacterial water filter, to provide clean drinking water. As a small portable filter that can be screwed onto a bottle.	Economist	Ozon Cyclery	Small business in Berlin-Lichtenberg offering workshops for interested customers to make their own tailored bamboo bicycle frames which consists of many very detailed steps.	Product designer

Table 1. Continued

Name	Aim of the project	Profile of the interviewees	Name	Aim of the project	Profile of the interviewees
Farm Hack	Worldwide community of farmers that build and modify their own tools. Catalogue of open source projects/tools.	Farmer and researcher	QroKee	The idea is to focus on 2D. The principal is to use a digital tablet, with a screen facing downwards.	software engineer, co-founder
FIRE	Platform to facilitate exchanges and innovation between individual designers, researchers and manufacturers in the UK.	Professor, researcher, project leader	Raid	Company that develops Technical clothes and equipment for trail runners based on a community approach.	R&D manager
Hovalin	3D-printed functional violin that can be printed with standard consumer printers	Software engineer and musician	Sunzilla	An open-source, portable, easy-to-use solar-powered generator to replace diesel gensets for temporary/off-grid power supply.	Electrical engineer
InMoov	Open Source 3D printed life-size robot. The initiator is a skilled designer. This project gathers a worldwide community of makers, educators and hobbyists.	Sculptor and Industrial Designer	Tinker bike	A motorcycle frame that allows building a motorbike out of a wide range of salvaged parts that would otherwise be scrapped.	Product designer
Knitic	A low scale mechanized machine that knits tubular garments with programmable patterns.	Artist	Ultrascope	An Automated Robotic Observatory (ARO) developed for amateurs to help identify and characterize Near Earth Asteroids	Industrial designer
La Cool Co	Open Source connected greenhouse, build on the need to not kill house plants.	Industrial designers	Waterzilla	A collaboration to make a system that allows to create clean drinking water with Open Source tools and technologies and it is run by solar panels.	Environmental Engineer

character of the results and partially subjective interpretation of the data. This is of course mitigated by the cross checking and multiple coding methods but in fact never fully disappears. To end up the section with a more positive note we can notice that empirical qualitative content analysis is gaining in popularity among the researchers today as this is one of the most efficient methods to obtain relevant data from complex human centric processes.

4. Corpus of interviews with OSH communities

Within the context of the OPEN! Project, a corpus of 23 interviews with different OSH practitioners was put together in order to investigate the empirical state of open source hardware design. The semi-guided interviews were carried out with representative OSH practitioners and recorded. The recordings were then transcribed and translated into English if necessary. An annotation study was then performed on the text data according to a code book.

In all, 23 project initiators from 22 OSH communities were interviewed for approximately 45 minutes each (more than one person was interviewed for some projects). Table 1 lists the OSH communities addressed as part of the interview campaign. These project initiators were based in France, Germany, England, the United States, Finland, Spain and Estonia. The average age of the interviewees was 33 (min = 23, max = 64), 86% of whom were male and 14% female. The interviewees were mainly project owners. They were makers and hobbyists mostly with technical background. However, a few of them had no real technical background or had a background that did not correspond to the project they were promoting. All the projects we interviewed were relatively advanced in their developments and most had already some working prototypes (FairCap, water-zilla, Sunzilla, Tinker bike). Others were close to production or dissemination (Echopen, Apertus axiom, InMoov, Farm Hack, Hovalin, ultrascope) (Table 1). Five interviewers took part in the campaign. The interviews were semi-structured and comprised a series of open questions. These concerned seven OSH themes designed collaboratively by the five interviewers. Two experienced researchers were members of the interviewers' team in order to guarantee data reliability as well as facilitate interview recording and preprocessing. The interview questions can be found in Appendix A.

Once all the interviews had been performed, the interviewers transcribed each interview recording. The three interviews in French were translated into English. The corpus of interview data was then coded with eight macro concepts, corresponding to the seven themes defined in the interview guide and an extra category entitled 'other'. The macro concepts are shown in Table 2. Each corpus was annotated by two coders according to the qualitative content analysis method (Zhang & Wildemuth 2016), and using the content analysis software Nvivo. Cohen's Kappa was used to assess the inter coder reliability of this part. Kappa range was [0.646–1] for each subcategory (Table 2) and for all categories included Kappa = 0.701, showing a satisfactory level of agreement of the coding. The basic content analysis unit was defined according to each individual concept unit rather than linguistic text unit (e.g., words, sentences or paragraphs). The conflicts of interpretation were identified and discussed among the coders and the outcome of the discussion recorded in a code book (see Appendix A).

Table 2. Description of the macro concepts

Concept	Description
Interviewee	Information regarding the interviewee: age, formal education, occupation, role/position in the community
Openness	Information regarding the three criteria defining an open source project: accessibility, transparency and replicability
Community	Information linked to the community: reason for joining or creating a community, interaction within a community, organization and size
Product/project/service	Descriptive information about the product (components, complexity, use, maturity level), project history, standards and quality
Design process	Information regarding the design process, KM, timeline management, decision-making and collaboration.
Tools and Functionalities	Information about tools, platforms and their functionalities
Business models	Information regarding business models, marketing, competitors, revenue models, identity (vision and values), intellectual property and licenses.
Other	

This corpus was developed with the aim of understanding OSH practices with respect to the seven different themes, including KM (Bonvoisin, Thomas, *et al.* 2017). Since the subject of this paper is KM we chose to focus on excerpts from the corpus annotated with 'Design Process' and 'Tools and Functionalities'. The first theme, 'Design Process', includes the topics 'knowledge management', 'decision-making', and 'collaboration', which are all closely related to knowledge sharing. The second, 'Tools and Functionalities', describes what knowledge sharing platforms do, how OSH communities use them and how these communities share knowledge through them. In the following we will describe the specific work that has been achieved on the KM related data set and that follows a qualitative data analysis approach as described in the next section.

5. Knowledge management related corpus analysis

5.1. Methodology

The KM related corpus was annotated by a researcher specialized in KM according to a qualitative analysis approach (Smith & Firth 2011). The aim of a qualitative study is not to identify a statistically representative set of respondents, but rather to generate or develop analytical categories, concepts and insights in order to extract information from the data. Following a contents analysis approach (Vaismoradi *et al.* 2013), the categories have been derived inductively, obtained gradually from the data, and used deductively to address the data in a recursive process as quote by Vaismodrali *et al.* (2013): *It should be noted that data analysis in both approaches [ndlr. Thematic or contents] are not linear, simply moving from one phase to another phase, but should be recursive with frequent review.* In our case, the quality and credibility of the categories and concepts have been ensured, following Schwandt, Lincoln & Guba 2007 recommendations by:

   (i) regular and systematic peer debriefing among the four authors,
  (ii) ensuring a prolonged engagement in, and persistent observation with open source communities from the second third and fourth author,
 (iii) cross-checking of data by first, second and third author using KM literature sources, mentioned in the results section.

The classic grounded theory analysis requires researchers to identify analytical categories from raw data rather than defining them a priori (Glaser, Strauss & Strutzel 1968). In our study, we used the framework approach (Smith & Firth 2011). Compared with grounded theory, the framework approach requires the researcher to establish a thematic framework by identifying all the key issues, concepts and themes used to examine the data. This is carried out by defining a priori issues and questions derived from research objectives as well as issues revealed by the data (Pope *et al.* 2000). There are five stages of data analysis in the framework approach:

   (i) First, a researcher familiarizes him or herself with the raw data by studying the corpus text.
  (ii) Second, the thematic framework is built from concepts generated from research objectives as well as corpus data.
 (iii) Third, the thematic framework is applied to annotate the corpus data. The second and third stages usually overlap because the data relevant to each category is identified and examined using a process called constant comparison. In this process, each item is compared with the rest of the data iteratively to define analytical categories until the thematic framework stabilizes (saturation Figure 1).
 (iv) Fourth, the corpus data is rearranged according to the appropriate part of the thematic framework to which it relates.
  (v) Finally, the researcher can use the rearranged data to define concepts, create typologies and find associations. In this stage, the data is interpreted to generate insights.

In our research, the thematic framework was developed from concepts found in the literature review as well as in the interview corpus data as we will see in the next sections.

5.2. Knowledge management related corpus annotation

To begin with, a preliminary analysis grid was set up to define multiple concepts: knowledge sharing, codification, personalization, knowledge loss, awareness, unstructured information, turnover, physical attributes and decision-making. The definition of these concepts that has been derived from the literature and introduced in the previous sections served as a means of organizing the data. The corpus was then annotated with concepts from the preliminary analysis grid. During the annotation process, new concepts as well as sub-concepts emerged from the data. We identified two types of decision-making: core team decision-making, which was the dominant form, and 'do-ocracy' decision-making. Three causes were attributed to the general theme of 'knowledge loss': organizational 'turnover', 'unstructured information' and low project 'awareness'. With respect to the category codification strategy, the sub-concepts were 'challenges' and 'benefits',
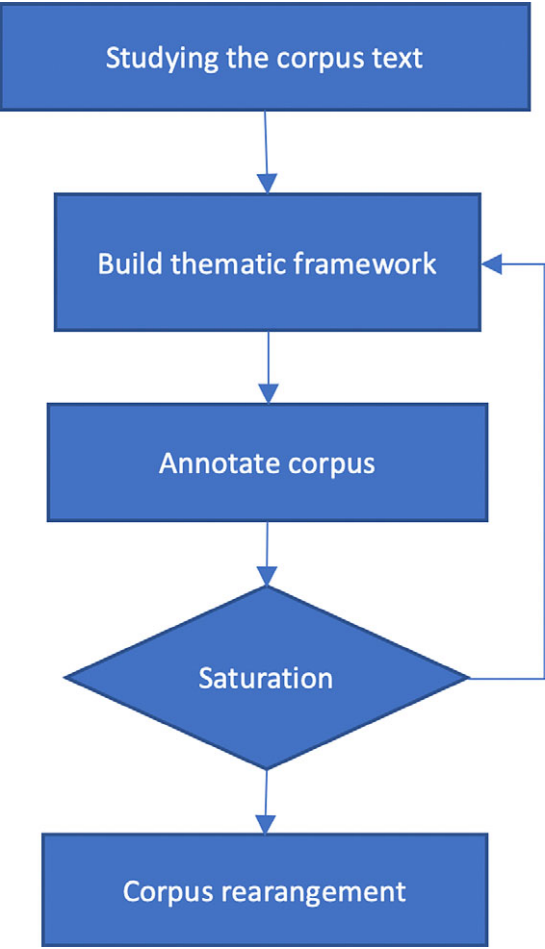
Figure 1. Corpus annotation process.

codified 'content' and 'tools' used for codification. As for the category personalization strategy, the corpus data analysis revealed three sub-concepts: OSH 'workshops', face-to-face 'meetings' and computer-mediated communications (CMC). All the sub-concepts were associated and grouped into a concept tree (Figure 2), where the 'tree leaves' represent the most detailed sub-concepts. This concept tree completely covers the three OSH KM challenges as defined in Section 2.2. The code book used is presented in detail in Appendix B.

The basic analysis segment used was the answer to the question (1 segment = 1 answer) and applied to the corpus annotation analysis presented above. Each segment is annotated with at least one subconcept from the 'tree leaves' of the concept tree, as well as concepts linked to these subconcepts on the 'tree branch'. An example of this analysis is shown in Table 3 where the first segment is annotated with the two subconcepts 'Turnover' and 'Unstructured information', both of which are linked to the 'knowledge loss' concept through the same branch in the concept tree. Therefore, all three concepts are attributed to this segment.
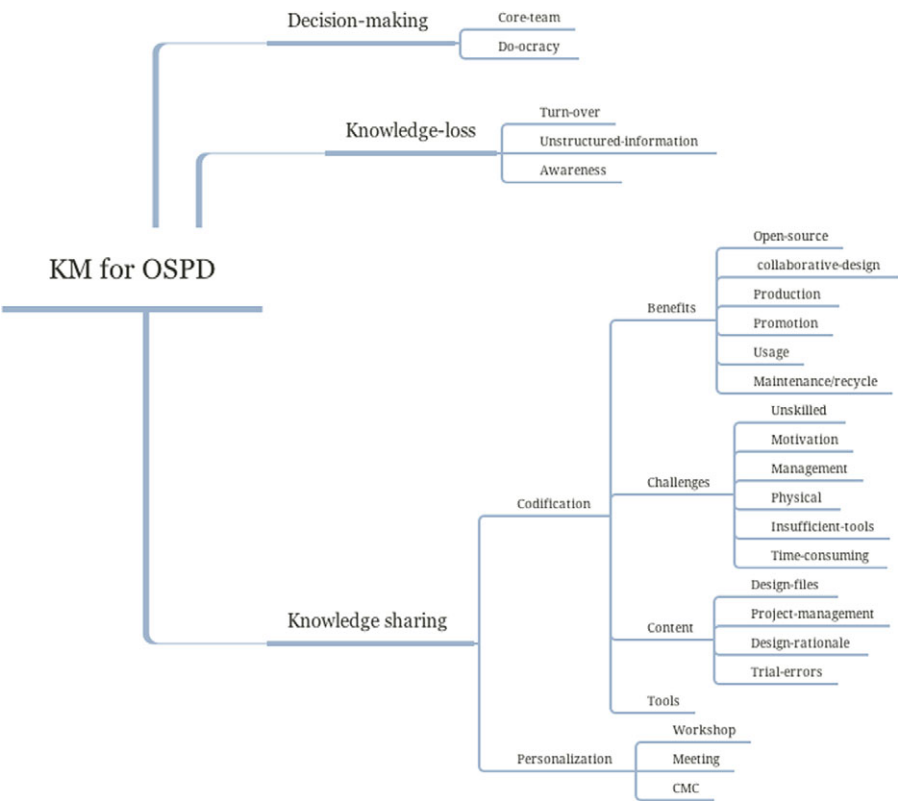
Figure 2. Concept tree of KM related concepts.

5.3. Exploratory quantitative study

The annotation analysis can be considered as a series of operations representing the data and reorganizing it into forms that lend themselves better to interpretation (Sandelowski 1995). In order to draw insights from an annotated corpus, we need to make the familiar more familiar, make the familiar strange, and/or reveal what is hidden (Patton 1990). This means eliciting recurrent concepts and searching for associated concepts to reveal hidden patterns in the corpus. Usually, this process is done by an experienced qualitative researcher who strictly follows the guidelines of a qualitative analysis methodology. In our case, a text-mining approach was used to explore the data owing to the substantial volume and complexity of the corpus. Statistical insights into the annotation concepts can be pulled from this analysis without the content of the corpus actually being interpreted. The general goal of this exploratory analysis is to navigate the corpus content with the guidance of annotated concepts. This first highlight the most frequent concepts in the corpus. The frequency of these concepts may be linked to their importance. The technique then reveals which concepts are frequently associated in the corpus. In our case, the situation with respect to knowledge management varied from project to project, with each interviewee talking about different KM aspects in the semi-directed interviews. Each project interview corpus was annotated with multiple concepts, while frequently associated concepts seemed to indicate a certain pattern in the corpus. The analysis

Table 3. Example of KM related corpus annotation

Project name: Apertus Axiom

Concepts	Corpus segments (excerpts)
#Knowledge-loss #Turn-over #Unstructured-information	*A lot of people working on the project came and left… they contributed something, but maybe it was unfinished…*
#Knowledge-loss #Unstructured-information #Knowledge-sharing #Codification #Challenge #Insufficient-tools	*…they just wrote it into the forum thread and that's just a chronological list of comments basically, so it's really difficult to filter out the essence of the information…*
#Knowledge-sharing #Codification #Content #Trial-errors	*…our approach is more to also (apart from achievements) share the challenges, the problems, things that we learned that did not work out, which I think is part of open source…*
#Decision-making #Do-ocracy	*We have organized a 'Do-ocracy', so there is no kind of hierarchy… the people who are doing something are in charge of what they are doing… but yes, it works.*
#Knowledge-sharing #Codification #Challenge #Physical	*You cannot touch materials or see how the weight and the grip of something works online, we did not find how to translate that to the internet properly.*

provides indicators based on concept association useful for interpretation later on, while already allowing an interesting qualitative interpretation.

Our analysis was performed using Rstudio with the tm and quanteda packages. The KM corpus analysed comprised 22 different projects, 89 content analysis segments and 32 unique concepts. Each segment was annotated with multiple concepts as shown in Table 3. The Document Term Matrix (DTM) applied is one of the most common formats for representing a text corpus in a bag-of-word format (Welbers, Atteveldt & Benoit 2017). A DTM is a matrix in which rows represent documents and columns represent terms or words in the corpus. In our case, a document represents a project, that is, a row in the DTM, while the annotation concepts used for that project are terms, that is, columns in the DTM. Our DTM contained 23 rows and 32 columns. For text mining purposes, it thus yielded a total of 23 documents with 32 unique terms. This DTM was used to generate term frequency analysis and term association analysis, as described in the following sections.

5.3.1. Term frequency analysis

As already stated, the aim of this analysis is to identify the most frequently annotated concepts. A Term Frequency Matrix (TFM) can be calculated by counting how many times a term appears in the DTM. This TFM can be visualized as a bar plot as shown in Figure 3, where each bar shows how many times a concept appeared in the annotation.

The aim of visualizing concept frequency is to pinpoint the most discussed concepts in the interview corpus. Interesting insights can then be drawn from the corpus content annotated with these concepts. Since each time a sub-concept is
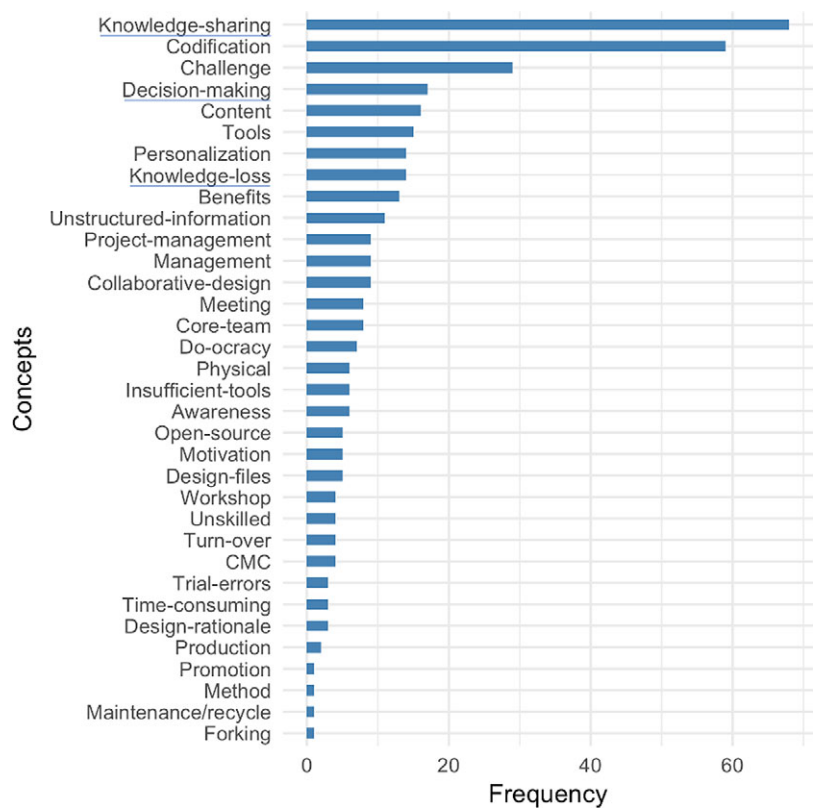
Figure 3. Bar plot of concept frequency.

annotated all the root-concepts related to it are also annotated, root-concepts are naturally more frequent than subconcepts. Figure 3 shows that, in our case, 'knowledge sharing' is the most frequent concept while 'codification' is frequently talked about as a knowledge management strategy followed by the concept 'challenge', which refers to the difficulties of knowledge codification. It appears that many 'codification' sub-concepts occur, for example, codification 'content', codification 'tools' and codification 'benefits' are all in the top 10 concepts. The top 10 concepts represent 69% of the occurrences and the path knowledge-sharing, codification, challenge appears to be the main one in terms of number of citations. It is worth mentioning that among all the subconcepts, the concept of knowledge codification 'tools' occurs the most frequently. 'Personalization', the other KM strategy, is also one of the top 10 concepts. The 'knowledge loss' phenomenon is also a frequently recurring concept in the corpus. This concept is often accompanied by the sub-concepts 'awareness', referring to project members' low awareness of a project's status, 'unstructured information', suggesting that project information is unstructured, and turnover, that is the phenomenon of transient community membership. Next come concepts detailing the benefits, challenges and methods of knowledge codification. For example 'physical' relates to the tangible features of physical products that throw up new problems for knowledge codification; 'motivation' is used in reference to the lack of motivation to document project knowledge; 'open source' is tied to knowledge codification as a precondition for

open-sourced projects, etc. With respect to 'decision-making', the subconcept 'core team' occurs slightly more frequently than 'do-ocracy'.

In summary, this analysis shows that 'knowledge codification' is frequently talked about in our corpus, especially its 'challenges' and its 'benefits'; 'knowledge loss' is an important issue; and 'core team' dominated project decision-making and self-organized 'do-ocracy' both exist. The frequency of these concepts gives us a general idea of the themes in this corpus. This means that we can now search for associations, that is concepts that appear frequently together.

5.3.2. Term association analysis

Another way to analyse corpus data is to search for repeatedly associated concepts to verify the existence of patterns in the corpus data. Usually patterns are identified by a researcher according to a qualitative analysis methodology. However, the cognitive process of pattern identification is implicit. In this exploratory study, we used text-mining techniques to help us identify these patterns. In the DTM, each document may contain multiple concepts, some of which often appear together in a document while others are relatively scarce. One way to visualize frequently associated concepts is to use hierarchical clustering techniques to group documents based on their Euclidean distance (Huang 2008). In our case, the aim was to calculate the distance of concepts (terms) rather than documents. This meant transposing the DTM into a Term Document Matrix (TDM), in which rows are concepts (terms) and columns are documents. The Euclidean distance between these concepts can be calculated; the shorter it is between two terms, the more likely they appear together in the same document. Furthermore, as documents are interviews of projects' participants, each of them reflects the point of view of one project. Therefore, the analysis provides information on the proximity of projects with respect to a certain concept. If two concepts are close to each other this means that an important number of projects share the same association.

Based on the Euclidean distance, the concepts are hierarchically clustered according to Ward's minimum variance method (Ward 1963). The results are displayed in a dendrogram in Figure 4. This diagram should be read from left to right and bottom to top. Each branch point encountered is the distance at which a cluster merge occurs. Consequently, the most distinct clusters are those with the largest separation, in other words those that are furthest from each other in the dendrogram. On the other hand, those closest to each other, often linked by branches from a low branch point, are those with the highest similarity (or Euclidian distance in our case). Our aim now is to identify the concepts displaying a high similarity, that is, those belonging to the same cluster or close to each other in the dendrogram.

There are six clusters of concepts in this dendrogram, each shown by a different colour and surrounded by dotted lines. The first cluster at the bottom of the diagram puts 'motivation', 'collaborative design' and 'project management' together. This cluster in fact covers codification benefits, collaborative design improvement, codification content, project management information, codification challenges and lack of motivation. To our knowledge, there is no literature showing this association and no semantic connection can be found in the corpus regarding these three main concepts. Therefore, this cluster was considered irrelevant.
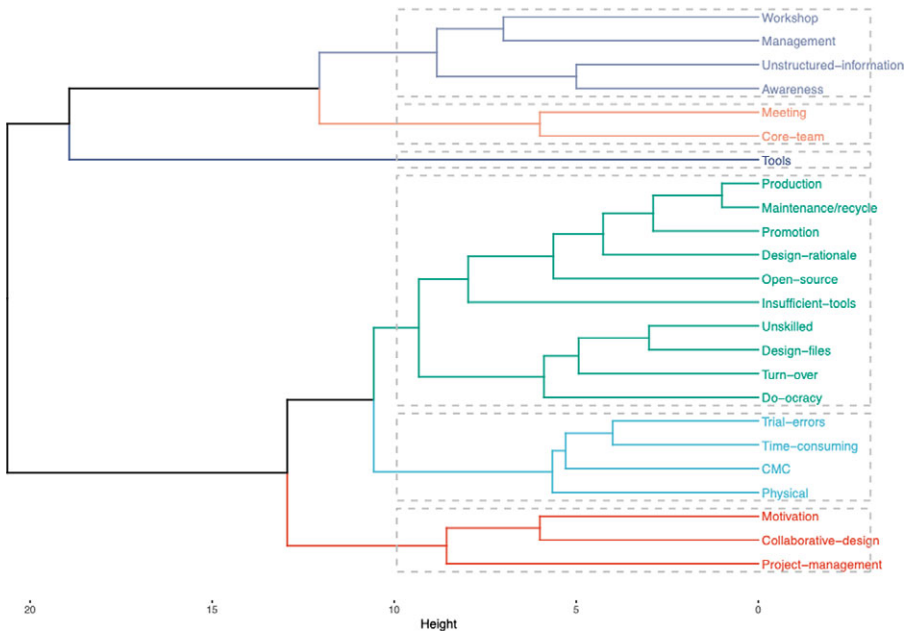
16/31

Figure 4. Dendrogram of concept association.

The second cluster (in blue) includes an interesting codification challenge: the physical nature of tangible products, the trial and error process involved in making them and the time-consuming nature of prototyping are associated with computer-mediated communication (CMC). This highlights the difficulties raised by transforming everything into digital format. According to the literature, knowledge codification for OSH is harder than for OSS owing to the physical aspects of tangible products. Our results show a connection between a certain type of codification content (trial and error) and CMC, which requires another type of codification. Also, the trial and error process involved in product prototyping is a typical feature of a physical product, one which naturally renders its codification time-consuming and makes its translation into the digital world (CMC) difficult. This aspect is related to our third research question and will be developed in the next section.

The third cluster (in green) includes 'turnover', 'do-ocracy', 'unskilled' and 'design files'. A high level of turnover in an OSH community naturally leads to potential issues in a do-ocracy based decision-making organization since there are many volunteer participants. The concept 'unskilled' refers to people lacking knowledge about design rationale documentation. This concept is therefore naturally linked to the codification content of 'design files'. The connection between 'turnover' and 'unskilled' may imply that volunteer amateur OSH participants lack the proper training for design rationale documentation which corresponds to our second research question. The third cluster also basically includes most of the concepts related to the benefits of codification. The association between the 'open source' and 'design rationale' concepts should be noted as it may indicate that documenting 'design rationale' helps to make a project open source.

In the fourth cluster (in orange), the 'meeting' and 'core team' concepts are grouped together. In the corpus, many interviewees mentioned that the core team of an OSH project usually has regular meetings to make decisions. Meetings are a classic personalization strategy for knowledge management and the core team usually plays an important role in decision-making in small sized OSH projects.

The last cluster (in dark blue) at the top of the dendrogram includes four concepts: lack of proper 'management' in an OSH project, 'unstructured information', low project status 'awareness' and 'workshops'. Due to the unstable organization of OSH projects, there is much unstructured information produced by volunteer participants. This naturally lowers people's awareness of project status since unstructured information is hard to interpret. In the corpus, some interviewees mentioned that lack of management is one of the challenges of knowledge codification. The association of 'management', 'awareness' and 'unstructured information' indicates that this lack of proper management may be one of the main reasons for knowledge loss. These remarks clearly highlight the issue of scaling up knowledge sharing across the community which is our first research question. As already stated, workshops are one of the personalization strategies used in OSH projects. It is interesting to see that it is associated with concepts representing knowledge loss and knowledge codification challenges. This suggests that it would be worth searching the corpus for the semantic relations between 'workshop' and the other three concepts in this cluster.

This exploratory provided us with insights into the corpus that helped us to navigate the voluminous and complex data set by generating indicators. These indicators tell us where we should focus in order to generate findings. Based on this analysis, we therefore reviewed the corpus content and our findings are presented in the next section.

6. Findings

In this section, we shall present our findings with respect to the corpus content. Guided by the exploratory analysis presented above, our interpretation of the corpus content allowed us to draw insights principally in relation to three aspects: OSH knowledge management strategies, codification challenges and the influence of 'turnover' on OSH knowledge sharing. These findings are in line with the three challenges of OSH knowledge sharing identified in Section 2.2.

6.1. The essential role of the codification strategy for OSH knowledge management

Both OSH knowledge management personalization and codification strategies were revealed in the corpus data. As indicated in the exploratory analysis, the codification strategy is mentioned much more frequently than the personalization strategy when it comes to the subject of knowledge sharing. This does not necessarily imply that it is more important than the personalization strategy. However, based on the content of our corpus, we believe there are three main reasons why the codification strategy is essential for OSH knowledge sharing.

6.1.1. Codification is a precondition for project openness

In order to render a project open-sourced any project information must be accessible to a large public. Documenting project information and sharing it publicly via the internet is one of the fundamental features of an OSH project. Therefore, knowledge codification is considered essential to guarantee a project's openness. As one of the initiators of Knitic put it:

> *We worked a lot with the documentation, open source means other people can make sense of all the information (of the project).*

A participant to project Apertus Axiom also stated:

> *And our general approach with the communication aside from the source files is that we try to be very honest and open about what's going on in the project, so even though some people don't understand this way our approach is that we share the good and the bad things, so when big companies report about products or about progressing development, they always only mention the good things and if they don't mention something than you can imagine that something went wrong. But our approach is more to also share the challenges, the problems, things that we learned that didn't work out.*

6.1.2. Codification is a precondition for personalized production, maintenance and recycling

One of the advantages of open-source hardware is its potential for personalized production. However, this is only possible if knowledge about the product design is well documented enabling OSH users to understand and modify the design of a product. In our case, some interviewees argued that OSH product documentation should go beyond product design files to include the whole project process. After all, the life cycle of an OSH product is not over after production. There is still maintenance and recycling to be considered. This point is underlined in what one of the initiators of OSE Germany said:

> *What we try to do is really document the whole process, its development, design, really everything. Even if you have already built it, you need maybe in the end to recycle it, so it's a circle. It is not just how to build it; it is not like DIY or something, it is really more.*

An Ultrascope participant adds:

> *there's a lot of people that have made their own telescopes and have actually pretty well documented their telescopes. That making it, their telescopes, so you can actually get some ideas of how it works. And then there is a lot of field tests.*

This aspect clearly illustrates the need for codification in order to allow scaling up knowledge sharing strategies in OSH communities (first research question).

Codification can mitigate the negative influence of 'Turnover'

Turnover issues have been raised in research question 2. Interviewees mentioned the difficulties of managing unfinished contributions from unstable participants in an OSH project. Knowledge codification is usually evoked as a solution to facilitate collaborative design given this unstable organization. When newcomers join an OSH project, they can gain project experience from project documentation. When they decide to leave the project, their unfinished work can be continued by other

people if their work has been well documented. The importance of documentation is reflected in the words of one of the initiators of Open Source Ecology:

> *We try to document everything so when the next person comes in, you can take over from where they left off.*

Apertus Axiom adds also:

> *A lot of people working on the project came and left again and they contributed something, but maybe it was unfinished, they raised some questions or something they found out and they just wrote it into the forum threat and that's just a chronological list of comments basically, so it's really difficult to filter out the essence of the information, so we went through it all and extracted the essence of the information and put it into a structure that will provide an easy overview.*

6.2. Valuable personalization strategy

In spite of the predominant position of the codification strategy in OSH practice, socialization is also highly valued as it allows people to exchange ideas, resolve conflicts, search for innovations or simply get to know each other. Our corpus underlined a close connection between socialization and decision-making with 'workshop' identified as the major form of personalization strategy making it possible for OSH practitioners to share knowledge offline.

The connection between socialization and project decision-making

Several interviewees emphasized the importance of talking to each other for project decision-making. The most common form of decision-making in an OSH project is made by the core team, which has regular meetings. These meetings can be face-to-face, or people can communicate with each other by email, skype or other forms of computer-mediated communication (CMC). Collaborative decision-making can also take place on a larger scale via a coordinating committee. The coordinating committee is also organized according to the principle of socialization meeting regularly to make project decisions. Note that socialization in the context of OSH always takes the form of CMC. Google hangout and Skype are frequently referred to as the two major tools for communication, especially for big OSH projects with a relatively large and dispersed team. As for collocated project teams, physical meetings are more common.

The importance of workshops to expand knowledge sharing

Workshop organization was mentioned repeatedly by several project initiators as a way to exchange ideas, gain new ideas, increase project visibility and develop networking. It is important for an OSH project to have a certain level of visibility in order to attract new participants and expand knowledge sharing across a large perimeter in the community, which refers explicitly to research question one. The OSH online community builds reputation and visibility mainly in two ways: online social media promotion and workshops at conferences, seminars and design fairs. For example, LaCoolCo project uses workshops as a means of promoting their project because 'everyone is there'.

> *So if they have makerspaces and they want, they are thinking about making a specific workshop for example with a specific target, public from specific ages or in a specific*

context. That makes us, that forces us to adapt what we're producing, to be useful for them as well… When we go to maker fairs or when we go to a workshop, everybody is there.

Ozoncyclery uses workshops as a way to enrich design ideation. This is neatly summed up by one of the initiators:

Knowledge (comes) not just from me but also from lots of people contributing ideas that we teach in the workshop…

Opensource ecology integrates this as part of its business model:

We prefer the idea of the immersion training workshop in manufacturing, where you organize the workshop. We have twelve people or so, they pay you to build it, they get a immersion training and you can sell the product. It's a dual revenue model, where you're catching revenue for manufacturing as well as education, and that's what we're doing.

6.3. The challenges of codification

Codifying knowledge has always been a challenge in engineering design, even for experienced designers. With respect to the OSH community new codification challenges have also emerged.

6.3.1. *Lack of management within the community*

Some interviewees reported their 'frustration' with the open source hardware design process. As mentioned in the section on 'turnover', free volunteer participation at any point in an OSH project is the signature of open source design. As already highlighted, this unstable form of participation raises several challenges. Many practitioners reported that one of the difficulties of collaborative documentation is the absence of a proper task management protocol within the community. Without a coordinating mechanism there is no clear documentation standard, even for sharing design files. There is no proper 'workflow management', especially when the stakeholders of a project are geographically distributed. This problem is reflected in the following words of one of the initiators of OSE Germany:

Multiple people working on the same project and documenting the process… You can't do that step by step… You have to set up a workflow.

Apertus axiom insist on the interest of a proper management tool:

Yes, and help find people who can take care of specific tasks because maybe some other parts of the community didn't realize that it needs to be done.

Task management and coordination mechanisms are classical problems in engineering design and are currently addressed in product life cycle management and project management tools. Our results clearly show that the problem remains in OSH communities as the tools available do not fully cover this need and the maturity of the collaboration practices has not reached a sufficient level today.

'*We do not know where to start*'

Some people said that despite their motivation and awareness of the importance of documentation, they lacked the training or skill to do it. Many OSH contributors

are amateur designers and DIY makers whose design process is more like that of a craftsman than an engineer. They usually begin their design process by trying out different solutions directly on a prototype rather than starting with a clear ideation and conceptual design phase. This kind of design process introduces subsequent problems for design knowledge documentation. According to an initiator from Ultrascope one of the main challenges is precisely this lack of documentation expertise:

> That's not one of my skills to be honest…

Even if they are technically skilled persons, the profile of the OSH participants makes it difficult to follow a structured engineering process. This is a clear drawback for OHS communities. Even for those capable of design knowledge documentation, there is still the problem of how to capture design rationale. Design rationale is an important research subject that has been addressed extensively in engineering design literature (Rohde *et al.* 2015; Eng, Marfisi & Aurisicchio 2011; Bracewell *et al.* 2009; Schaik *et al.* 2011). It reflects the reasons behind design decisions. Capturing it involves taking into account not only decisions about product design but also why those decisions are made. A number of interviewees complained that some design knowledge documentation cannot be understood because the context of that knowledge is missing, in other words the design rationale is absent. One of the initiators from FarmHack emphasizes the importance of Q&A knowledge sharing forums where the OSH community context is preserved, making the knowledge more comprehensible:

> Documentation exists in the context of the community which can answer questions and provide additional documentation upon request… documentation in the context of a community shouldn't be lost.

6.3.2. Participants are not motivated to document

Some participants are not motivated to document their project information. They do not realize what the benefits of knowledge codification are and are especially discouraged by how time-consuming it can be. One specific interviewee from Hovalin argued that OSH projects for him were a hobby and *you do not document what you decide* about your hobbies.

FarmHack member complains:

> There are a number of projects that are half-documented, and some are not really complying with the intention of it. So even when their intention was expressed early on to post and fully document.

This can also be connected to the issue of design rational capture that has been clearly identified as an issue in engineering design for long as highlighted in the previous section. Both skills and motivations are two factors that hinder the documentation process and appear as relatively typical from OSH communities. In more traditional engineering contexts, professionals are trained and documentation is part of the engineers' duty making these difficulties less salient.

6.3.3. The physical aspect of OSH

Compared with software design, OSH development involves 'offline' activities linked to the tangible attributes of physical products. Many interviewees reported

that they usually tried to make a prototype directly without working on the conceptual design on a computer. By directly handling a physical product, the designer can feel the material, make quick adjustments and explore different alternatives. This kind of activity does not require extensive knowledge of CAD. However, it is challenging and sometimes impossible for designers to put all this tangible experience into a documentary form. As one of the initiators of Apertus Axiom said:

> *You can't touch materials or see how the weight and the grip of something works online, we didn't find a way to translate that into the internet properly.*

Another particular feature of a tangible product compared with an entirely digital software product concerns its manufacturing, usage, maintenance and recycling, all of which are completely 'offline'. The knowledge related to these offline phases is usually lost. As one of initiators from InMove put it:

> *Although you know there are many people that download the parts, reproduce the robot, they do other things with it, I don't even know what they are doing.*

These nondocumented activities clearly represent a problem for OSH communities in terms of knowledge sharing (scaling up knowledge sharing) and in terms management. This aspect seems typical from OSH communities and absent from engineering design literature.

6.3.4. Tool limitations
Some of the interviewees reported that the tools they use are not adapted to OSH. Currently, there are no collaborative platforms specifically dedicated to OSH. Practitioners mainly rely on GitHub, which is a platform developed for open source software design and, as such, it lacks many essential features for tangible product knowledge sharing, especially CAD file sharing. A number of people interviewed said they use a series of management and file-sharing tools to support their project, for example, google drive, basecamp and wiki. However, a specially designed collaborative work platform is needed to scale up participation in an OSH project. Recently, more and more initiatives have emerged to build potential alternatives dedicated especially to OSH, for example, *Wikimedia, Wikifactory*.

6.4. The infamous 'turnover' phenomenon

The 'turnover' phenomenon is a well-established term in open source software design to describe the unstable organization of an open source software design team. Thanks to the openness of a design project, volunteer designers can join as well as leave the project team at any time. This freedom of participation is one of the signatures of open source design projects while posing serious challenges for their success. From the perspective of knowledge management, the freedom of participation grants all the participants access to project information, the opportunity to contribute to it, as well as the possibility of putting an end to their contribution whenever they wish, without paying attention to whether or not their contribution can be understood by other people or not.

In the interview corpus, the turnover phenomenon was identified recursively in several different projects. One of the initiators of the Apertus Axiom project aptly summarized this phenomenon as follows:

A lot of people working on the project came and left…they contributed something, but maybe it was unfinished…they just left…

'Turnover' is generally perceived as negative by the interviewees. It is associated with 'unstructured information', a concept that emerged frequently in our corpus alongside that of 'knowledge loss'. Both are seen as the undermining consequences of the unstable participation in an OSH project. Volunteers may begin to work on part of a project without fully understanding the project's past experience. Their work may be left undocumented, be poorly documented or even unfinished, making it very challenging, if not impossible, for other people to continue the work. This challenge is reflected in the words of one of the initiators of Open Source Ecology:

If somebody comes in as a volunteer, in order for that to be very effective, they'll have to understand the entire process…

Another downside of 'turnover' is low project status awareness. Several interviewees mentioned the lack of traceability with respect to other people's work, project progress and project vision. The unstable participation of volunteers results in a dynamic but volatile project status, which most of the project members are not fully aware of. In the words of one of the initiators of OSE Germany:

You put it to one side…then what you did was basically lost… it's such a little thing but it makes a big difference

Most of the OSH practitioners interviewed were perfectly aware of the negative impact of this high turnover. Two 'defense mechanisms' aimed at reducing its negative impact were identified.

1. Core team decision-making. Interviewees said they found it difficult to include unstable volunteer participants in project decision-making. The most common decision-making strategy in OSH projects is to restrict this activity to a small project core-team. This is perfectly summed up by one of the initiators of Raid:

   If we include everyone (when making a decision), we can't move forward. It's more for the core team and the most invested people within the project.

With small core teams like Sunzilla decision making is consensual:

so we just talked and, we don't have like hierarchies so we try to make decisions the five of us together and even if it is just for… we never do something against one person he or she is really against and we don't do it. But sometimes you take decisions and four of us are for the decision and one person is neutral and it's ok.

2. Standardization of participation. Some initiators described how certain standards can be imposed in a project, for example requiring participants to document and share their design files using a specific template. Since all project members are then obliged to share their contributions using the same template the amount of unstructured information is reduced. For example, the Echopen project has set up a series of templates to help project participants to share their contributions. Farm hack is also conscious of that:

Yes, so what we've talked about, our ideal, that we'd have standard tools, APIs, a database standard for tool documentation and those would be exchangeable between

platforms. That's sort of our goal within FarmHack even, so that our partners don't even need to go to FarmHack website, just reference their library of tools, and have their community contribute.

Turn-over has been identified as a predominant issue in OSH communities raising up many issues in terms of knowledge sharing, management and decision making. This phenomenon seems specific to OSH communities given the nature of the commitment of the participants. Indeed, in companies stakeholders are tightened by contract to participate to the company's effort, when participation in OSH projects is based on a volunteer involvement.

7. Conclusion

The OSH phenomenon is growing at a very fast pace, changing the way in which product design knowledge is created, shared and reused. In this paper, we identified the challenges of knowledge management through a review of the literature on open source software design and engineering design. This, together with a corpus of interviews with OSH practitioners, allowed us to develop a concept tree and perform further qualitative analysis. In order to facilitate the qualitative analysis, text mining technologies were applied to the concepts in order to identify frequent concepts and frequent concept associations. Guided by the text mining analysis, insights were drawn from the corpus allowing us to establish and present a number of findings. We were able to confirm that codification plays an essential role in OSH knowledge management and that for a project to become truly open source the participants need to have a minimum level of knowledge documentation skills. Personalization is also a valued strategy for OSH practitioners. It involves various socialization means such as workshops, meetings and computer mediated communications. Although different strategies are implemented in OSH development projects, there are still many challenges. First, OSH projects are usually managed by a small core team but involve many other volunteer participants. This leads to a high community membership turnover rate, unstructured project information, low overall project status awareness and, ultimately, knowledge loss. Second, compared with an OSS project, knowledge documentation for OSH is even more challenging since it is more difficult to translate the physical aspects of tangible products into the digital world, and there are no platforms specifically developed for OSH.

While this paper contributes to OSH research with its empirical findings and its demonstration of text mining to facilitate qualitative corpus analysis, we are nevertheless conscious of the limits of our work. Our qualitative analysis is based mainly on what the important contributors have to say, that is those community members making up the OSH project core team. In the future, our findings could be potentially enriched by including the more occasional OSH participants.

Acknowledgements

Appendix A

A.1. Interview questions

 (i) What information/files do you share (CAD files, schematics, project specifications, description of the concept, design brief)?
 (ii) If not, is it something you plan to do?
 (iii) Do you share the entirety of the product description, or are there some components you keep closed?
 (iv) Is it possible for anyone to reproduce your product?
 (v) Can people make changes to your documents? (CAD files, BoM, assembly instructions)
 (vi) What made you want to join this community?
 (vii) Could you describe your community?
 (viii) Who contributes to this project? Could you describe them (e.g., professionals in their own time, hobbyists, employees from firms with vested interests)?
 (ix) How has the community developed (stable member base, growth, stagnation, etc.)?
 (x) How many members are in the community?
 (xi) How big is the core team of the community?
 (xii) What networks is the project engaged in (research institutions/networks, maker spaces, R&D alliances, other projects, etc.)?
 (xiii) Who are the project's main stakeholders in order of relevance (members, end user groups, suppliers, founders, society at large, etc.)?
 (xiv) Is there a commercial interest in your project?
 (xv) How much time do you spend on this project (offline/online)?
 (xvi) Do you meet with other community members (personally/physically/online)?
 (xvii) What is the product you are working on? Could you describe it?
 (xviii) What is your interest in this product? How did you choose to work on these products?
 (xix) What is your vision? What are your values?
 (xx) What does your project solve?
 (xxi) What is the potential impact of this product?
 (xxii) Could you describe the design process (staged versus highly iterative)?
 (xxiii) Do you have clearly set milestones (i.e., rigid, frequent versus flexible, less frequent design reviews)?
 (xxiv) Is there a timeline for the entire project?
 (xxv) How are decisions made in your project (committees, regular meetings, type of meetings, approval guidelines, etc.)?
 (xxvi) What particular knowledge management activities are done to facilitate sharing (Wikis, Lessons learnt, checklists, cross project communication, continuous improvement, etc.)?
 (xxvii) What major issues have you come across during the development process?
 (xxviii) Could you elaborate on any failures within the project? Which external and internal factors led to these?
 (xxix) What are the internal weaknesses you have come across?
 (xxx) What are the external threats facing your project?

(xxxi) Do you use an online design platform?
(xxxii) If not, what do you use?
(xxxiii) What enables your capacity to contribute to the design activity in the community?
(xxxiv) What does the perfect platform look like to you?
(xxxv) In your understanding how does this project create value?
(xxxvi) What is your target market/segment? Who is your customer specifically?
(xxxvii) What is your revenue model?
(xxxviii) What are your customer acquisition channels?
(xxxix) What are your main cost drivers?
(xl) What makes your company and its offer unique?
(xli) What income sources do you activate?
(xlii) Age
(xliii) Gender
(xliv) Formal education
(xlv) Occupation/relevant experience (job or otherwise)
(xlvi) What competencies do you bring to the community?
(xlvii) What is your role in your project? Leader/organizer, contributor/participant.

Appendix B. Code book for KM corpus annotation

Concept	Description
Decision-making	Interviewee talks about the activity of making decisions.
Core team	The decision-making activity is executed by a restricted number of people in a project.
Do-ocracy	The task is distributed in a self-organized manner, decisions are made by people who are doing the task.
Knowledge loss	Interviewee talks about the phenomenon of losing track of knowledge produced in a project.
Turnover	The phenomena whereby volunteer participants unexpectedly join and leave a project.
Unstructured information	Project information is not structured.
Awareness	Project members are not aware of project status.
Knowledge sharing	Interviewee talks about the activity of sharing knowledge
Codification	Interviewee talks about the activity of knowledge documentation.
Benefits	Interviewee mentions the benefits or the importance of knowledge codification.
Open source	Codification is good for project openness.
Collaborative design	Codification facilitates collaborative design.
Production	Codification facilitates production of a product.
Promotion	Codification helps to promote the project.
Usage	Codification facilitates usage of a product.
Maintenance/recycle	Codification facilitates product maintenance and recycling.

Design Science

Concept	Description
Challenges	The challenges of knowledge codification.
Unskilled	Participants lack knowledge codification skill.
Motivation	Participants lack knowledge codification motivation.
Management	The lack of project management makes it difficult for knowledge codification.
Physical	The tangible features of a physical product introduce new challenges relating to knowledge codification compared with software
Insufficient tools	The tools participants use for open design are insufficient for documentation.
Time-consuming	The interviewee specifically mentions their frustration with respect to documentation because it is too time-consuming.
Content	The interviewee talks about the content of documentation.
Design files	The content is in files related to product design, usually in the form of CAD, diagrams, drawings, 3D models, etc.
Project management	The content concerns how the project is managed.
Design rationale	The interviewee specifically mentions the context of design knowledge and the reasons for design decisions.
Trial-error	The interviewee specifically mentions that the content mainly concerns the trials and errors of a product prototype.
Tools	The interviewee talks about the tools they use for documentation.
Personalization	The interviewee talks about the activity of sharing knowledge through P2P communication.
Workshop	The interviewee mentions the workshop as an activity for knowledge sharing.
Meeting	The interviewee mentions meetings as an activity for knowledge sharing
CMC	The interviewee mentions using computer mediated communication as a way to share knowledge.

References

Ackerman, M. S., **Dachtera, J.**, **Pipek, V.** & **Wulf, V.** 2013 Sharing knowledge and expertise: the CSCW view of knowledge management. *Computer Supported Cooperative Work: CSCW: An International Journal* **22** (4–6), 531–573; doi:10.1007/s10606-013-9192-8.

Aitamurto, T., **Holland, D.** & **Hussain, S.** 2015 The Open Paradigm in Design Research. *The MIT Press Journal*, **31** (4), 17–29.

Aurisicchio, M. & **Bracewell, R.** 2013 Capturing an integrated design information space. *Journal of Engineering Design* **24** (6), 397–428; doi:10.1080/09544828.2012.757693.

Boisseau, É., **Omhover, J.-F.** & **Bouchard, C.** 2018 Open-design: a state of the art review. *Design Science* **4** (e3), 1–44; doi:10.1017/dsj.2017.25.

Boujut J.F., **Pourroy, F.**, **Marin P.**, **Dai X.**, **Richardot G.** 2019, "Opensource hardware communities: investigating participation in design activities", In: *Proceedings of the 22nd International Conference on Engineering Design (ICED19)*, Delft, The Netherland.

Bonvoisin, J., **Buchert, T.**, **Preidel, M.** & **Stark, R. G.** 2018 How participative is open source hardware? Insights from online repository mining. *Design Science* **4**, e19; doi:10.1017/dsj.2018.15.

Bonvoisin, J., **Mies, R.**, **Boujut, J.-F.** & **Stark, R.** 2017 What is the "source" of open source hardware? *Journal of Open Hardware* **1** (1), 1–18; doi:10.5334/joh.7.

Bonvoisin, J., **Thomas, L.**, **Mies, R.**, **Gros, C.**, **Stark, R.**, **Samuel, K.**, **Jochem, R.** & **Boujut, J. F.** 2017 Current state of practices in open source product development. *Proceedings of the International Conference on Engineering Design ICED 17*, **2**, 111–120.

Bracewell, R., **Wallace, K.**, **Moss, M.** & **Knott, D.** 2009 Capturing design rational, 1–3.

Bracewell, R., **Wallace, K.**, **Moss, M.** & **Knott, D.** 2009 Capturing design rationale. *Computer-Aided Design* **41** (3), 173–186; doi:10.1016/j.cad.2008.10.005.

Chandrasegaran, S. K., **Ramani, K.**, **Sriram, R. D.**, **Horváth, I.**, **Bernard, A.**, **Harik, R. F.** & **Gao, W.** 2013 The evolution, challenges, and future of knowledge representation in product design systems. *Computer-Aided Design* **45** (2), 204–228; doi:10.1016/j.cad.2012.08.006.

Chen, Y.-J., **Chen, Y.-M.** & **Chu, H.-C.** 2008 Enabling collaborative product design through distributed engineering knowledge management. *Computers in Industry* **59** (4), 395–409; doi:10.1016/j.compind.2007.10.001.

Eng, N., **Marfisi, E.** & **Aurisicchio, M.** 2011 Adapting aerospace design rationale mapping to civil engineering: A preliminary study. ICED 11 – 18th International Conference on Engineering Design - Impacting Society Through Engineering Design, 6(August), 468–479.

Fjeldsted, A., **Adalsteinsdottir, G.**, **Howard, T. J.** & **McAloone, T.** 2012 Open source development of tangible products. In *DS 71: Proceedings of NordDesign 2012, the 9th NordDesign Conference,* Aarlborg University, Denmark. August 22–24, 2012.

Foucault, M., **Palyart, M.**, **Blanc, X.**, **Murphy, G. C.** & **Falleri, J.-R.** 2015 Impact of Developer Turnover on Quality in Open-source Software. In *Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering*, pp. 829–841. ACM; doi:10.1145/2786805.2786870.

Giovan Francesco, L. & **Michle, M.** 2003 The knowledge ecology of open-source software projects. *Proceedings of 19EGOS Colloquium Copenhagen*, 3–5.

Glaser, B. G., **Strauss, A. L.** & **Strutzel, E.** 1968 The discovery of grounded theory; strategies for qualitative research. *Nursing Research* **17** (4), 364.

Hansen, M. T., **Nohria, N.** & **Tierney, T.** 1999 What's your strategy for managing knowledge? *Harvard Business Review* **77** (2), 106–116, 187. http://www.ncbi.nlm.nih.gov/pubmed/10387767.

Howard, T. J., **Achiche, S.**, **Özkil, A.** & **Mcaloone, T. C.** 2012 Open design and crowd-sourcing : maturity , methodology and business models. In *Design Conference.* Dubrovnik.

Huang A. 2008 *Proceedings of the sixth new zealand computer science research student conference (NZCSRSC2008)*, Christchurch, New Zealand.

Huysman, M. & **Wulf, V.** 2006 IT to support knowledge sharing in communities, towards a social capital analysis. *Journal of Information Technology* **21** (1), 40–51; doi:10.1057/palgrave.jit.2000053.

Klein, M. 1993 Capturing design rationale in concurrent engineering teams. *IEEE Computer Journal* **26** (1), 1–17; doi:10.1109/2.179154.

Kohlbacher, F. 2006 The use of qualitative content analysis in case study research. *Forum: Qualitative Social Research* **1** (7), 1–30; doi:10.1007/s10551-011-0925-7.

Lee, J. 1997 Design rationale systems: understanding the issues. *IEEE Expert* **12** (3), 78–85; doi:10.1109/64.592267.

Lethbridge, T. C., **Sim, S. E.** & **Singer, J.** 2005 Studying software engineers: data collection techniques for software field studies. *Empirical Software Engineering* **10** (3), 311–341; doi:10.1007/s10664-005-1290-x.

Licorish, S. A. & **MacDonell, S. G.** 2014 Understanding the attitudes, knowledge sharing behaviors and task performance of core developers: a longitudinal study. *Information and Software Technology* **56** (12), 1578–1596; doi:10.1016/j.infsof.2014.02.004.

Maher, M. L., **Paulini, M.** & **Murty, P.** 2010 Scaling up: from individual design to collaborative design to collective design. In *Design Computing and Cognition* (ed. J. S. Gero), pp. 581–599. Springer.

McMahon, C., **Lowe, A.** & **Culley, S.** 2004 Knowledge management in engineering design: personalization and codification. *Journal of Engineering Design* **15** (4), 307–325; doi:10.1080/09544820410001697154.

Müller-Seitz, G. & **Reger, G.** 2010 Networking beyond the software code? An explorative examination of the development of an open source car project. *Technovation* **30** (11–12), 627–634; doi:10.1016/j.technovation.2010.07.006.

Ouertani, M. Z., **Baïna, S.**, **Gzara, L.** & **Morel, G.** 2011 Traceability and management of dispersed product knowledge during design and manufacturing. *CAD Computer Aided Design* **43** (5), 546–562; doi:10.1016/j.cad.2010.03.006.

Patton, M. Q. 1990. *Qualitative Evaluation and Research Methods*. SAGE Publications, Inc.

Pawlowski, J. M. & **Bick, M.** 2015. The global knowledge management framework: towards a theory for knowledge management in globally distributed settings. *Leading Issues in Knowledge Management* **2**, 134.

Pope, C., **Ziebland, S.**, **Mays, N.** & **Mays, N.** 2000 Analysing qualitative data. *BMJ 2000*; **320** (7227), 114–116 doi:10.1136/bmj.320.7227.114.

Powell, A. B. 2015. Open culture and innovation: integrating knowledge across boundaries. *Media, Culture and Society* **37**(3), 376–393; doi:10.1177/0163443714567169.

Raasch, C. 2011 Product development in open design communities: a process perspective. *International Journal of Innovation and Technology Management* **08** (04), 557–575; doi:10.1142/S021987701100260X.

Raasch, C., **Herstatt, C.** & **Balka, K.** 2009 On the open design of tangible goods. *R and D Management* **39** (4), 382–393; doi:10.1111/j.1467-9310.2009.00567.x.

Rashid, M., **Clarke, P. M.** & **O'Connor, R. V.** 2017 Exploring knowledge loss in open source software (OSS) projects. *Communications in Computer and Information Science* **770**, 481–495; doi:10.1007/978-3-319-67383-7_35.

Riege, A. 2005 Three-dozen knowledge-sharing barriers managers must consider. *Journal of Knowledge Management* **9** (3), 18–35; doi:10.1108/13673270510602746.

Rigby, P. C., **Zhu, Y. C.**, **Donadelli, S. M.** & **Mockus, A.** 2016 Quantifying and mitigating turnover-induced knowledge loss: case studies of chrome and a project at Avaya. *2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE)* 1006–1016; doi:10.1145/2884781.2884851.

Rohde, D., **Štorga, M.** & **Marjanović, D.** 2015 Design rationale capturing model for use during the embodiment phase of the product design. *Transactions of FAMENA* **39** (1), 27–42.

Sandelowski, M. 1995 Qualitative analysis: What it is and how to begin. *Research in Nursing & Health* **18** (4), 371–375.

Schwandt, T. A., **Lincoln, Y. S.** & **Guba, E. G.** 2007 Judging interpretations: but is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. *New Directions for Evaluation* **2007** (114), 15–25.

Smith, J. & **Firth, J.** 2011 Qualitative data analysis: the framework approach. *Nurse Researcher* **18** (2), 52–62; doi:10.7748/nr2011.01.18.2.52.c8284.

Sowe, S. K., **Stamelos, I.** & **Angelis, L.** 2008 Understanding knowledge sharing activities in free/open source software projects: an empirical study. *Journal of Systems and Software* **81** (3), 431–446; doi:10.1016/j.jss.2007.03.086.

Stuwart, T. & **Ruckdeschel, C.** 1998 Intellectual capital: the new wealth of organizations. *Performance Improvement* **37** (7), 56–59; doi:10.1002/pfi.4140370713.

Treude, C. & **Storey, M.** 2010 Awareness 2.0: staying aware of projects , developers and tasks using dashboards and feeds. *International Conference on Software Engineering (ICSE)*, 365–374; doi:10.1145/1806799.1806854.

Vaismoradi, M., **Turunen, H.**, **Bondas, T.** 2013 Content analysis and thematic analysis : Implications for conducting a qualitative descriptive study. *Nursing and Health Sciences* **15**, 398–405; doi:10.1111/nhs.12048.

van der Ven, J. S., **Jansen, A. G. J.**, **Nijhuis, J. A. G.** & **Bosch, J.** 2006 Design decisions: the bridge between rationale and architecture. In *Rationale Management in Software Engineering*, pp. 329–348. Springer.

van Schaik, J., **Scanlan, J.**, **Keane, A.**, **Takeda, K.** & **Gorissen, D.** 2011 Improving design rationale capture during embodiment design.

Voigt, C., **Montero, C. S.** & **Menichinelli, M.** 2016 An empirically informed taxonomy for the maker movement. In: Bagnoli F. et al. (eds) Internet Science. INSCI 2016. Lecture Notes in Computer Science, vol 9934. Springer, Cham.

Vuletic, T., **Duffy, A.**, **Hay, L.**, **McTeague, C.**, **Pidgeon, L.** & **Grealy, M.** 2018 The challenges in computer supported conceptual engineering design. *Computers in Industry* **95**, 22–37; 10.1016/j.compind.2017.11.003.

Welbers, K., **Van Atteveldt, W.** & **Benoit, K.** 2017 Text analysis in R. *Communication Methods and Measures* **11** (4), 245–265.

Zhang, Y. & **Wildemuth, B. M.** 2016 Qualitative analysis of content. *Applications of Social Research Methods to Questions in Information and Library Science* second edition, Libraries United, 318–329.

Zhuoxuan, L., **Seering, W.**, **Ramos, J. D.**, **Yang, M.** & **Robert, W. D.** 2017 Why open source? Exploring the motivations of using an open model for hardware. In *Proceedings of the ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference.* Cleveland, Ohio, USA, © 2017 by ASME http://dx.doi.org/10.1115/DETC2017–68195.