CLASSIFICATION SYSTEMS, THEIR DIGITIZATION AND CONSEQUENCES FOR DATA-DRIVEN DECISION MAKING: UNDERSTANDING REPRESENTATIONAL QUALITY

Completed Research Paper

Mari-Klara Stein
Copenhagen Business School
Frederiksberg, Denmark
mst.itm@cbs.dk

Sue Newell
Bentley University, USA and
Warwick University, UK
snewell@bentley.edu

Robert D. Galliers
Bentley University
Waltham, MA, USA
rgalliers@bentley.edu

Erica L. Wagner
Portland State University
Portland, OR, USA
elwagner@pdx.edu

Abstract

Classification systems are foundational in many standardized software tools. This digitization of classification systems gives them a new 'materiality' that, jointly with the social practices of information producers/consumers, has significant consequences on the representational quality of such information systems. Based on a multi-site field study, we suggest that representational quality is achieved through four types of negotiations that human actors engage in when confronted with the materiality of a new IS. These negotiations are associated with three broad practices (instantiation, re-narration and meta-narration), and three different information production / consumption situations. We contribute to the relational theorization of representational quality and extend classification systems research by drawing explicit attention to the importance of 'materialization' of classification systems and the foundational role of representational quality in understanding the success and consequences of data-driven decision-making.

Keywords: Representational quality of IS, standardized packages, data-driven decision-making

Introduction

High-quality information is key to better decision-making and underlies the success of analytics, business intelligence and 'big data' projects, which are currently receiving considerable attention in business (Marchand and Peppard 2013; Watson and Wixom 2007)[1]. This suggests the importance of information systems (IS) research concerned with quality; that is, with the quality of the information captured by (and subsequently used from) an IS. However, both practically and theoretically, this issue of IS quality is often over-looked. For example, from a practical standpoint, many analytics projects in business tend to apply the traditional strategy of implementing a new IS on time, to plan and within budget (Marchand and Peppard 2013). However, an IS – especially one that is used as a basis for decision-making – cannot deliver its full benefits if the organization lacks an understanding of how people create and then use the information emanating from it. Theoretically, while IS quality has been recognized as an important antecedent to user satisfaction with – and/or acceptance and usage of – an IS (Nelson et al. 2005; DeLone and McLean 1992; Sabherwal et al. 2006), interest in IS quality in its own right, has waned over recent years (Nelson et al. 2005, Cordoba et al. 2012). Accordingly, in this paper, we examine IS quality with a particular focus on understanding the representational quality of information systems – that is, how well the system represents a particular social world.

We build on the premise that an IS is an "artifactual representation of real-world system (as perceived by someone)" (Wand and Weber 1995: 208). IS, therefore, are based on "sets of concepts humans use to organize knowledge about domains" (Parsons and Wand, 2008: 840). The more closely or faithfully an IS manages to model the real-world system it is intended to model, the better its representational quality. Typically, real-world systems (such as those related to a company's customers or the patients that a hospital admits) become manifested as classes or categories in an IS, so that the "selection of categories determines everything that is represented in a computer application" (ibid: 840, based on Sowa 2000: 51). Classes are so important in an IS because classification is a fundamental human activity that helps people construct useful abstractions; draw inferences about unobserved properties of an object, and reduce the complexity of dealing with individual instances of a phenomenon (ibid: 839-840).

Such digitized classification systems, thus, are foundational in many IS and become key in the achievement of representational quality. For example, electronic patient record (EPR) and customer relationship management (CRM) systems are both tools that are founded on digitized classification systems and are used heavily as a source for data-driven decision-making. In using such tools, people desire systems that "faithfully represent some domain, because they provide a more informed basis for action than unfaithful representations do" (Burton-Jones and Grange 2012: 5). Yet it is clear that the classification systems upon which such tools are based, and so the data stored in them, reflect *someone's* representation of a real-world phenomenon. Decisions based on these tools are, therefore, not neutral and have direct and significant consequences for the actors concerned.

Since the tool is often not neutral, it is important not to 'black box' the IT artifact (Baskerville, 2012; Benbasat and Zmud 2003; Orlikowski and Iacono 2001) but rather consider how well an artifact (and more specifically here the digitized classification system that is the focus of the empirical research) represents the real-world system it is intended to mirror (Milton 2007). The representational quality of an IS is not solely determined by the IT artifact, however, since the 'real-world' is subjectively known (Weick 1995). As many IS scholars have begun to recognize, there is an "indeterminate relationship between tasks and technologies such that a technology's fixed materiality could support multiple task structures depending upon people's desires and goals" (Leonardi 2012: 40). This indicates that, in order to understand representational quality, we need to examine both the materiality of the tool and the perceptions and practices of those using the tool.

The term materiality infers that there is some non-human agency involved in the representational quality of an IS; that is, 'stuff" (that does not necessarily have material substance, as with a classification system) can physically or cognitively constrain and/or enable human activities (Leonardi 2010; Markus and Silver

[1] See, for example, the detailed McKinsey consulting report, "Big data: The next frontier for innovation, competition and productivity" (Manyika et al. 2011)

2008). In order to narrow down non-human agency as it relates to representational quality, we elected to use the framework proposed by Wand and Weber (1995). This postulates three levels of system structure: surface, deep and physical. Surface structures refer broadly to the interface of the tool (e.g., the visual organization of information; format of reports available for users); deep structures refer to the inner logic or system of meaning (e.g., the rules about how transactions must be posted to ledgers in accounting systems), and physical structures refer to the hardware components of the system. For the purposes of this paper, we are particularly interested in how accurately and completely the surface and deep structures of the IT artifact represent the users' model of a real-world system.

Representational quality, as already indicated, is always understood in terms of users' interpretations and practices (Wand and Wang 1996). As human agents working with an IS are confronted with its surface and deep structures, they negotiate this materiality in practice in order to make sure the IS represents their social world in an acceptable manner. Returning to the example of an EPR, the tool (non-human agency) clearly imposes a particular classification of diseases, which constrains doctors' record-keeping practices. As well as including options for the checking-off of disease categories, however, the tool may also allow for the inclusion of free text. Doctors negotiate these constraints and affordances in practice by inputting important shorthand notes in fields that allow them to do so. The representational quality of this IS is achieved at the intersection of these material and human agencies, and doctors' decisions are likely to differ significantly with and without the presence of these additional notes (Newell et al. 2012).

In short, we argue for the examination of representational quality as something that is achieved (to a greater or lesser extent) in practice, rather than something that is given and static. Moreover, we view this practice as relational – representational quality is achieved through the interplay of human (e.g., user interpretations) and non-human (e.g., deep and surface structures of a software tool) agencies (cf. Wand and Weber 1995). More specifically we attempt to answer the question: *How do human agents negotiate the materiality of an IS in practice to achieve representational quality*? We explore this question in the context of the implementation and use of a particular packaged software tool – a tool that digitizes the classification of academic work - at two North American universities. We describe four types of negotiations that human actors engage in when confronted with the materiality of a new IS and through which representational quality of IS is achieved. We contribute to the theorization of representational quality as relational, and demonstrate how it is a key issue in understanding the success – or otherwise – of many standardized software tools that are founded on classification systems, and used to make decisions. We also contribute to research on classification systems by drawing explicit attention to the importance of the 'digitization' of such systems, as well as to the point that it is not classification systems per se that are of interest to information consumers/producers. Rather, the successful use of classification systems relies on the achievement of their representational quality in practice.

The rest of this paper is structured as follows: first, we introduce the theoretical foundations of our study, followed by a description of the chosen research methodology and research setting. We then summarize our findings, and end with a discussion of their theoretical and practical implications.

Theoretical background

This section outlines the theoretical foundations of this paper. We begin by introducing our conceptualization of representational quality of IS. We then introduce two common logics of organizing information that are relevant for our study – the narrative and database logics. Lastly, we discuss how the different elements of representational quality – the surface and deep structures of a particular system, the classifications embedded in these structures and the related social practices – work together.

Representational Quality of Information Systems

Our conceptualization of representational quality builds on the common definitions for information and system quality (Nelson et al. 2005) as well as representation theory (e.g., Wand and Weber 1995). Nelson et al. (2005) perceive both information and system quality to be multi-dimensional. Information quality comprises dimensions of accuracy; completeness; currency, and format, while system quality comprises accessibility; reliability; response time; flexibility, and integration. Generally speaking, most of these dimensions are considered contextual, in that different user populations may assess the completeness, currency and format of the information – and the flexibility and response time of the system – quite

differently. Wang and Strong (1996) define representational quality of data as the conciseness and consistency of its representation as well as its interpretability and ease of understanding (the format dimension as defined by Nelson et al. 2005). For the purposes of this paper, however, such an understanding of representational quality is too narrow. It is quite possible that an IS is concise, consistent and interpretable and yet does not represent users' social worlds in a manner that is agreeable to them. An unfaithful representation may be a product of the underlying deep structure lacking an element (e.g., a table) to capture some important aspect of the user's social world and/or a product of data being entered incorrectly into the system (Burton-Jones and Grange 2012).

Accordingly, we adopt a broader notion of representational quality in line with Wand and Weber (1995). Firstly, a 'good' IS (representing a real-world system) should be accurate and complete (Milton 2007, based on Wand and Weber 2002). That is, the system should present a complete and accurate picture of the user's social world and the many meanings contained in it (Table 1). However, considering that working classification systems are never fully consistent or complete (Bowker and Star 1998), we adopt a pragmatic approach towards assessing these representational quality criteria – focusing on whether the IS is 'good enough' to meet the user's needs (cf. Burton-Jones and Grange 2012; Wagner and Newell 2006). In addition, we argue that a 'good' representation is also easy to understand (well formatted), and current if necessary. Furthermore, system quality dimensions are also seen to be important, but are understood not to be independent from the information quality dimensions (Nelson et al. 2005). For example, the flexibility and integration dimensions of system quality are likely to be influential when it comes to the completeness and interpretability of information stored and captured in the system. In line with Leonardi (2012), we consider that these non-human agency aspects of an IS (which may be more or less adaptable to user needs) can support multiple representations (which are more or less compatible with the users' social world, depending upon their desires and goals). In line with our research question, we are interested in how information producers/consumers negotiate these non-human agency aspects of an IS to achieve representational quality.

Table 1. Dimensions of Representational Quality (Nelson et al. 2005; Milton 2007)

Dimension	Definition
Accuracy	Degree of agreement between the data values presented by an IS and the actual values the data represents in the real world
Completeness	Degree to which all possible states relevant to the user population are represented in the stored information
Currency	Degree to which the information is considered up-to-date
Format	Degree to which information is presented in a manner that is understandable and interpretable to the user, and thus aids in the completion of a task
Flexibility	Degree to which a system can adapt to a variety of user needs and to changing conditions
Integration	Degree to which a system facilitates the combination of information from various sources to support business decisions

In terms of thinking about the particular features of the IT tool that become part of such negotiations, we next introduce two common ways in which information is organized and represented in an IS: the narrative and the database logics (Manovich 1999).

Database and Narrative Logics

Broadly, database and narrative logics refer to two general forms of how humans organize their experiences in the world (cf. Hayles 2007; Manovich 1999; Snyder 2004). Database logic refers to a form of data organization that is characteristic of numerous new media (digital) objects, such as relational databases, but also to the World Wide Web more generally (Manovich 1999). Particular to the database logic is the organization of data into a collection of related categories on which the consumer can perform operations, such as querying, reporting and appending additional datasets in an effort to expand the

database's scope. The strength of the database logic is its ability to order large amounts of data, and then re-order that data based on query requests that present the data in the format most appropriate to the user, whether or not they are requesting data that reside within one, or hundreds, of different tables (Hayles 2007). Narrative logic, on the other hand, organizes data (human experience) into a cohesive story, with a beginning and an end, and some kind of thematic, causal pattern, in-between (Manovich 1999; Snyder 2004; Wagner 2003). New elements cannot be added to a story without making it a different story. Once told (given an explicit existence in a book, movie, etc.), a story is the only narrative available. If we start reordering someone's telling of events, we explicitly make it into a different story.

In essence, database and narrative logics characterize particular types of 'deep structure' that may be common to many different IS. Furthermore, the surface structures of different IS typically also reflect, to some extent at least, the underlying logic. For example, MS Word reflects a broadly narrative logic, in its surface structure (interface) as well as its deep structure (e.g., using customizable markup language). Conversely, an enterprise system like SAP ERP, reflects a broadly database logic – its interface is structured around transactions, fields and tables, and its deep structure is typically a relational database. The user experience of interacting with a collection of items (a database), thus, is quite different from the user experience of interacting with a narrative (e.g., a story or a movie). These logics present different models of what a world is like (Manovich 1999). In the narrative form, the words, sentences and storyline that make up the narrative are given to the reader and have a material existence on the page or the screen. Other elements, which form the world of a particular genre or style, and could have been chosen to appear in the narrative, are implicit and exist in the imagination of the narrative consumer. In the database form, this relationship is reversed – the elements of a particular 'world' are given an explicit existence, while the narrative is implicit and imagined by the information consumer (ibid.).

To use an illustration from our research setting – the world of academic CVs and annual reports – the narrative form is clearly present in each CV and annual report. These tell the story of a particular faculty member and her/his specific scholarly, service and teaching activities throughout their career and during the past year respectively. The report will tell us explicitly what kind of research the faculty member publishes (e.g., books), the kind of service activities in which (s)he engages (e.g., committee and editorial work), etc. However, the reader of the report (if at all familiar with academia) can also imagine the other, implicit, elements that are part of a particular academic discipline and the annual reporting genre, and that could have been chosen by the faculty member. Each annual reporting narrative, therefore, is a manifestation of a particular combination of elements making up this social world. If a different combination of elements were to be made explicit, it would constitute a different narrative.

In a database form, the elements of a particular social world are made explicit. Thus, the different types of scholarly, research and service activities undertaken by faculty members are first defined, and the resulting classifications embedded into the deep structure of the tool. Faculty members are still able to create their unique annual report by inputting and classifying their activities, but the database user is constrained by the finite number of categories provided to them. At the same time, by decoupling a cohesive story into different elements (e.g., entities and attributes – in this case, types of scholarly output), the database form makes it easier to re-couple the elements in various ways, so that other kinds of stories can be imagined. For example, database consumers (decision makers) have the freedom to construct narratives of comparison – for example, individual faculty members could be compared across departments, in terms of the number of publications they produce. Thus, a CV management tool based on database logic is designed to produce a representation of both individual faculty members' productivity as well as the state of the university as a whole, while a tool like MS Word, based on the narrative logic, is designed to produce a representation of an individual's story of their work activities.

This leads us back to the *quality* of this representation. Representational quality measures the degree to which an IS faithfully models a particular social world. In the particular case of digitized systems based on the database logic (such as EPRs, CV management systems), we run into the issue that the more or less faithful modeling of a social world is done through classification. Classification, in turn, is based on an imperfect domain of partially agreed-upon collective background knowledge, as no working classification system is fully consistent or complete (Bowker and Star 1998). Achieving representational quality in digitized classification systems in particular, therefore, becomes a question of practical negotiations between the human and non-human actors involved in getting the work done – as we discuss next.

Classification systems, logics, and the social practices of human agents

A classification system is generally defined as "a set of boxes, metaphorical or not, into which things can be put in order to then do some kind of work – bureaucratic or knowledge production" (Bowker and Star 1998: 233). In short, this means that an IS with a deep structure of a database is also typically a digitized classification system because of the way in which data must be structured into correlated entities or classes and described by a set of attributes (Hayles 2007)[2]. For example, a database of literature will most likely have the category 'author' characterized by attributes such as name, date of birth (and death), gender, nationality, etc. Another entity – 'book' – may be characterized by attributes such as title and genre. There are a number of classifications at work here that enable the description of a particular entity. For example, to fill out the nationality, gender and genre attributes, classification systems related to internationally recognized nation states, gender and literary genres are in use. These examples also point to the contested nature of classification systems, as new nation states, gender (e.g., transgender), and genre (e.g., Twitter fiction) categories come into existence through negotiations about their inclusion in common classification systems. Inclusion of a new category in a classification system affords the subject/object of that category legitimacy, while the exclusion of that category means the subject/object cannot be represented at all in a database. In essence, excluded categories do not exist in the model of the world presented by the classification system and the database (Bowker and Star 1998; Hayles 2007). Parsons and Wand (2000) refer to this problem as 'the tyranny of classes' and discuss how problems of inclusion/exclusion are the result of the traditional approach to information modeling, which is based on fixed, or "inherent" classifications, where predefined classes of things form the basis of system design.

Many packaged, off-the-shelf software tools used in work organizations are founded on database logic and pre-determined or inherent classifications (Parsons and Wand 2007). They offer structured and searchable containers for describing events and set up standard procedures (Bowker and Star 1998; Winman and Rystedt 2012). The popularity of such systems is the premise that their database logic and their standardized nature allows them to serve as efficient vehicles for representing different domains, such as the state of work processes in a business, the health condition of a person/nation, etc. (cf. Burton-Jones and Grange 2012). Through their database logic and inherent classifications, these systems place descriptive and interpretive demands on their users (White et al. 2009). They invite certain, organizationally acceptable, kinds of descriptions and categorizations of events, people and objects, which may obscure as much of a particular event, person or object as they reveal (ibid).

Our exploration is geared, in particular, towards those packaged software tools that are built on digitized classification systems. Such systems are not neutral media through which work practices are achieved, but rather actively reconfigure and transform professional practices (Pollock and D'Adderio 2012; White et al. 2009; Wagner et al. 2006). In particular, we choose to study a recently introduced tool, where a digitized classification system has not yet become institutionalized, i.e., gained taken-for-grantedness and legitimacy (cf. Colyvas and Powell 2006). This allows us to explore the negotiations that happen in relation to the materiality of the new system (see Figure 1 below), which we would not be able to see with an established tool where the negotiations had already been completed.

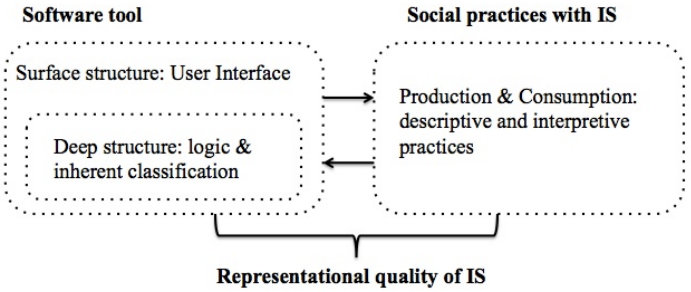

Figure 1. Proposed model for understanding the achievement of representational quality

[2] See Parsons and Wand (2000) for a discussion of system design not based on inherent classification

Method

A multi-site field study was chosen to investigate the practical achievement of representational quality in a Faculty Productivity software package (FP). The study was conducted across two sites in North America – a large state university ("State") and a small private university ("Private") – both of which had purchased and recently implemented FP in order to improve efficiency in gathering faculty activity data and other administrative functions. Overall, we conducted 47 semi-structured interviews across the two research settings over an 18-month period. Interviews were conducted with a range of stakeholders, including university administrators, faculty members and staff responsible for implementing FP. All interviews were tape recorded and transcribed. Additional data (e.g., meeting recordings; university-wide memos) were collected and examined. We also collected 17.5 hours of observational data, including documentation of faculty members using FP in filling out annual activity reports, used for performance evaluations.

Data Analysis

To understand how representational quality of a digitized classification system is/is not achieved in practice, we analyzed the interview and observational data in the following steps. First, we coded chunks of data, where dimensions of representational quality of the FP system were mentioned. Once relevant data were identified, we considered these chunks in relation to the conceptual scheme derived from prior literature (see Table 1). Next, we reviewed the chunks of data in context to identify how the social and the material were inter-related in the achievement of representational quality. From this step of our analysis we identified four different types of negotiations that occurred as faculty attempted to improve the achievement of representational quality (see findings section). Finally, we identified three different practices that were taking place in these episodes of negotiating representational quality that were undertaken by producers and/or consumers of the data (see findings section). Table 2 provides an example of our coding, relating to one type of negotiation – negotiating the elasticity of categories.

Table 2. Example of Data Coding

Data (chunks relevant for representational quality)	Dimensions of representational quality (Table 1)	Type of social-material negotiation	Practices
"*There is one thing I want to highlight* [in my annual report]. *It's this engagement that I've been doing at the [X-Russian institution] … so I'm just like, OK, it's non-credit instruction. It is kinda executive education, it's a seminar I gave, so… I'm just gonna go with it.*" (tenured faculty member, department chair, Private) "*Some of our biggest achievements with our presentations don't count as refereed, because it's a presentation … That's the trick with looking at the P&T guidelines – what do you get to put under that refereed box? So then [FP] comes along and the categories don't match up…. What I don't want to do is to make decisions different from my colleagues if it puts me at a disadvantage. So it's really important for me to know how other people are interpreting this information.*" (tenure-track faculty member/librarian, State)	*Accuracy / Completeness*: FP does not accurately or completely reflect the different types of activity (e.g., different types of conference proceedings, seminars) *Integration*: FP does not align with P&T guidelines *Format*: It is unclear how others will interpret data entered into FP	*Negotiating the elasticity of categories* (i.e., what activities can you put in what categories?) in relation to existing tools (e.g., P&T guidelines that are discipline-specific); dependent on power – tenured member 'goes for it' versus tenure-track member needs to know what others are doing before entering data.	Producers inputting personal activity information into FP (*instantiation*). Examples also show how *producers* are thinking about what account/story *consumers* will construct from this information (*re-narration*)

Research Setting

The FP software package enables the capturing of faculty activity data. Faculty input their research, teaching and service activities into FP through a web user interface and the data are stored on the vendor's cloud-based repository. The vendor first offered the package in 1999. Currently, there are approximately 3,000 FP adopters in over 25 countries. In essence, FP provides a standardized way to capture the information previously contained in faculty CVs and annual evaluation reports. Both of these artifacts are inherently a combination of database and narrative logics. They contain both ordered and unordered collections of items (e.g., lists of publications), to which it is usually possible to add a new element. Both also contain narrative elements – for example, in annual evaluation reports, faculty members are typically asked to explain their achievements in a short narrative. The explicit narrative fills in the blanks that cannot be inferred from looking at the items in isolation. Typically, CVs and annual reports have been created using various packaged software (e.g., MS Word and FP). FP is founded on a structured database logic, which requires each individual to create their CV / annual report according to the same deep structure/classification system, which is not amenable to end-user modifications (Table 3).

Table 3. Description of different artifacts used for CV maintenance and annual reporting

Artifact	Deep structure	Surface structure (UI)	Classification systems (examples)
FP	Relational database (database logic; embedded classification systems)	Digitized forms (sets of fields for data entry; some free text data entry possible) (database logic with few narrative elements)	Embedded in the surface & deep structure: - classification of academic work into areas related to teaching, service and scholarship - classification of service work into department, university, professional or public service; classification of professional service by pre-defined roles: editor, reviewer, chair, etc. - classification of scholarship into pre-defined types: journal articles, books, grants, etc.
MS Word	Customizable document markup language (narrative & database logics)	Text editor (narrative with few database elements)	User-defined (institutionally agreed upon + individually customized): - classification of academic work into areas related to teaching, service and scholarship - classification of service work into professional, etc. - classification of scholarship into types (peer-reviewed or not), etc.

In the universities under study, both MS Word and FP were in use for CV and annual reporting purposes. At "State", the decision to purchase FP was made by the Provost and was mainly driven by the need for some kind of central faculty vitae database that would allow for easier productivity reporting and would eventually feed into a performance-based budgeting approach. The goal was to replace idiosyncratic MS Word CVs with a standardized output from FP. At "Private", conversely, the decision to purchase FP was made by a committee, and was driven mainly by the need to collect up-to-date and accurate faculty activity data for accreditation reporting. In addition, FP replaced MS Word as the annual reporting tool and also fed standardized web profiles for all faculty members.

In sum, the setting – with its focus on biographical information, collected and stored in a newly introduced, highly structured IT tool (FP) – is particularly suitable for exploring the question of how representational quality of digitized classification systems is achieved in practice. Also, in our setting, information producers, consumers and 'objects' are often the same individuals – faculty members input their CV data, 'consume' other faculty members' CVs and are also the 'object' to which the information pertains. This makes faculty members a particularly suitable audience to study, as they are part of all the practices involved in negotiating the achievement of representational quality. Furthermore, the university setting makes these negotiations clearly visible, because change (new IT) typically cannot be forced upon faculty members (particularly those with tenure) by the administration.

Findings and Analysis

Our findings revealed four types of negotiations in which information producers/consumers engage in relation to the materiality of the new system: the "elasticity of categories"[3]; the revision of classifications; the legitimate meaning of categories, and the room for error in manipulating classified instances (see Table 4). These negotiations are associated with: a) three broad practices – instantiation, re-narration and meta-narration – common in digitized classification system use (see final row of Table 5), and b) three different information production/consumption situations – documenting information for oneself; for similar others, and for dissimilar others (cf. Markus 2001). Next, we will outline the details of these negotiation-practice-situation combinations (and present summaries in Tables 5 and 6 below).

Table 4. Different Types of Negotiation Aimed to Improve the Representational Quality of FP

Negotiating the "elasticity of categories"	Flexibly applying the categories in FP in order to improve how well the system represents a person's social world
Negotiating the revision of classifications	Changing the classification system so that a person's activities are not marginalized
Negotiating the legitimate meaning of categories	Changing the meaning of the classification categories to better represent a person's or group's interests and needs
Negotiating the room for error in manipulating classified instances	Establishing the acceptable level of accuracy in an aggregated data-set that the system flexibility affords

Types of Information Consumption/Production Situations

It is important to distinguish between different types of information production and consumption situations, because the constraints and affordances imposed by classification systems are likely to have differing significance across such situations. Thus, the negotiations in which information consumers and producers engage are also likely to vary. We find three different situations (overview in Table 5) that are relevant for how representational quality is achieved. We name these situations after a framework proposed by Markus (2001). The *first* describes a case where information is produced and consumed by the same individual (documenting information for oneself). We call such individuals' producer consumers. The *second* describes a case where information is produced and consumed by different individuals, who share similar work roles, backgrounds and interests (documenting information for similar others). The *third* describes a case where information is produced and consumed by different individuals, who do not share similar work roles, etc. (documenting information for dissimilar others). Each of these situations presents information consumers and producers with different challenges and opportunities. Thus, how representational quality of an IS (that is involved in the situation) is achieved differs as well. Next, we describe the three practices common in digitized classification system use and the different negotiations through which representational quality is achieved.

Table 5. Types of information production/consumption situations (adapted from Markus 2001) and the corresponding common practices

Documenting Information for Oneself	Documenting Information for Similar Others	Documenting Information for Dissimilar Others
Actors: - Producer consumers Information is produced and consumed by the same individual	Actors: - Producers - Consumers Different individuals who share similar roles/ background/interests	Actors: - Producers - 3rd party producers - 3rd party consumers Different individuals who do not share similar roles/background/interests

[3] We adopt this phrase from one of the field interviews, an excerpt of which is found later in the paper.

Example situation: Faculty input data into FP and then look at how this data presents them on their online profile	Example situation: Faculty input data into FP; other faculty look at this data to get an overview of colleagues' recent research	Example situation: Faculty input data into FP; administration aggregate this data or Secretaries input data for faculty; administration aggregate this data
Key practices in IS use: - Instantiation	Key practices in IS use: - Instantiation - Re-narration	Key practices in IS use: - Instantiation - Re-narration - Meta-narration
Instantiation: Producers' input in FP of concrete examples of activities	Re-narration: Consumers' construction of a story from an individual's information extracted from FP	Meta-narration: Consumers construction of a comprehensive explanation/story based on aggregated data extracted from the database

Practice of Instantiation: Converting Contextual Information into Standardized Instances

The practice of instantiation is related to the conversion or translation that must happen when contextual information is entered into a database-type tool that embeds a rigid classification system in its deep structure and, to some extent, in its surface structure. While entering data into a system like FP, information producers also consider how potential consumers may interpret those data, closely linking instantiation and re-narration. When documenting information for themselves, this re-narration is least problematic: if the producer and the consumer are the same person, that person is likely to have few problems in interpreting their own records. In such cases, representational quality of an IS is mostly related to the successful practice of instantiation.

To understand better how humans negotiate the materiality of the new IS, we can separate the human and non-human agency in this practice for analytical purposes (Leonardi and Barley 2010). One particularly relevant material element is *integration*: the degree to which a system facilitates the combination of information from various sources to support business decisions (Nelson et al. 2005). As a new IS is implemented, its legitimacy is not yet established and users naturally look to how well it aligns with existing tools and procedures used for the same purpose. In the case of FP, faculty members expect the system to conform to their established Promotion and Tenure (P&T) guidelines and existing annual reporting forms. These expectations were not completely satisfied in our settings. Confronted with the materiality of FP, information producers engaged in three types of negotiations (Tables 4 and 6): negotiating the "elasticity of categories"; the revision of the classifications, and the legitimate meaning of classifications.

Negotiating the "elasticity of categories" is relevant for the achievement of representational quality in all information production/consumption situations. It allows producer consumers to flexibly apply the classification to their data, thereby increasing the *accuracy* of the representation for themselves. For producers who are documenting information for others, this flexible application of classification allows them to tweak how the system is representing their social world and to take some control over how the standardized data will be interpreted (re-narration) by consumers (improving *accuracy* and *format*). Negotiating elasticity of categories becomes easier the more powerful the information producer is. Below are two illustrative quotes from more and less powerful faculty members in the situation of filling out their annual reports (documenting information for others):

"*There is one thing I want to highlight* [in my annual report]. *It's this engagement that I've been doing at the [X-Russian institution]. So I want a record of it, coz I'm gonna also talk about it in my report* [also recorded in FP]*, so I want a record of it here. Now, I don't know how the Russian education system works… it's supposed to be for credit, but not everyone in there is getting credit … not every attendant was enrolled in the class, so I'm just like, OK, it's non-credit instruction. It is kinda executive education, it's a seminar I gave, so… I'm just gonna go with it.*" (tenured faculty member, department chair, Private)

"*So some of our biggest achievements with our presentations don't count as refereed, because it's a presentation? But our papers, which come out of those presentations go under refereed? That's the trick with just looking at the P&T guidelines – what do you get to put under that refereed box? So then [FP] comes along and the categories don't match up …. What I don't want to do is to make decisions different from my colleagues if it puts me at a disadvantage. So it's really important for me to know how other people are interpreting this information*." (tenure-track faculty member/librarian, State)

We can see that the tenured faculty member is more successful in his negotiation and flexibly applies the category of non-credit instruction to achieve what he considers an accurate representation of his work. The tenure-track faculty member, conversely, is less successful, because his room for negotiation is much narrower – how flexibly the categories can be applied depends heavily on the collective understanding of the legitimate meaning of categories (see more below). If such meanings have not been agreed upon, less powerful actors are left to struggle in an attempt to achieve the 'right' level of accuracy and format. While negotiating the elasticity of categories is undertaken, to some extent, by all information producers, ***negotiating the revision of the classification systems*** themselves is generally taking place at the margins (cf. Star and Griesemer 1989). People who do not (relatively) neatly fit into the classification system embedded in a digital tool will more likely seek to renegotiate the classifications. As exemplified by the quote below, faculty members clearly feel marginalized and seek revisions of the existing classification system when they perceive that how they instantiate their 'academic narrative' in FP becomes likely to be interpreted and re-narrated as that of an 'other' (an *inaccurate* and *incomplete* representation):

"*If ninety percent of your work is 'Other', obviously you don't matter, because they're telling you that what you're doing, while it may be crucial to the institution – and, I mean, literally crucial – they're not valuing it. […] A part of this is defining exactly what the categories are. There certainly haven't been discussions about any of these kinds of things*." (tenured faculty member, department chair, State).

As already noted, information producers also consider how potential consumers may interpret the data they enter, closely linking instantiation and re-narration. When documenting information for others, in particular, producers will want to instantiate their records in such a way that they will be interpreted by others in line with what is considered generally accepted in their social world (e.g., academic discipline). This involves not only the negotiation of "elasticity of categories", but also ***negotiating the legitimate meaning of classifications***. Again, we find that actors with less power, and those in the minority, have a bigger stake in this negotiation. Existing classifications are typically based on the interests and needs of the majority (the powerful), requiring little negotiation on their part. The quote below demonstrates the kinds of negotiations undertaken by Natural Sciences faculty at Private (with a focus on business). It shows how the classifications taken-for-granted by the majority (e.g., business faculty members) can hinder the achievement of *accuracy, completeness* and 'good' *format* of the system, for the minority:

"*[In Natural Sciences], we have a huge number of sub-specialty journals, and these have smaller distribution, lower Impact Factor, but higher prestige – because there, you're really talking about the nitty-gritty of your field. As opposed to "Science" and "Nature", which have huge Impact Factors, but you also know half the stuff in there isn't true … if someone only publishes in "Science" and "Nature" and there isn't a follow-on paper in a couple of years in "Journal of Biological Chemistry" or … one of these 'everything-better-have-three-decimal-points-after-it' type of journals, then you worry about the original thing. And we've dealt with that in P&T, but not in [FP]. I mean there's something in [FP] about the quality of journals, but we do not have such a thing as A, B and C journals. So [FP] has no flexibility beyond listing things and a lot of Arts & Sciences faculty feel it's being foisted upon them by business needs*" (tenured faculty member, former department chair, Private)

Practice of Re-narration: Interpreting Standardized Instances

The second practice common in digitized classification systems use, especially in situations where information is documented for similar and dissimilar others (i.e., producer and consumer are not the same individual) is related to the interpretation that must happen when consumers place an already-classified instance back into a broader context and story. Because instantiation and re-narration are so closely linked, the material constraints and affordances are the same. Thus, the material element particularly relevant is *integration*. The practices of negotiating the "elasticity of categories" and the legitimate meaning of classifications allow not only for information producers to satisfactorily instantiate their records, while constrained and supported by the materiality of the classification system, they also

allow for information consumers to interpret or re-narrate these instances in a collectively acceptable manner (see Tables 4 and 6). Research in the context of other digitized classification systems has demonstrated that decision making within particular institutional settings presupposes extensive knowledge of the relevant classification systems – which originates in the participants' shared institutional history (Winman and Rystedt 2012). Our findings also suggest that reaching collectively acceptable interpretations (including reaching a joint understanding of the legitimate meaning of categories and how to apply categories flexibly) is only possible with some level of shared background understanding present:

"*I sort of understand the 'why' part, like, OK, it's for planning … these are metrics they're going to do something with. That's all kind of hocus-pocus to those of us who are faculty, even chair level administrators, but we get that idea. The point that I still don't understand is, which of these categories matter. And what would be really useful [would be] to know why are these the ones that matter. Is it about presidential priorities, university commitments … improving the quality of the institution in the "Newsweek" rankings, whatever it is*?" (tenured faculty member, department head, State)

Without a shared background understanding, collective interpretation becomes problematic (good *format* will not be achieved for information consumers). In our setting, such breaks can be seen between different academic disciplines, who are all required to classify their work using the standard system embedded into FP. This then also influences the conversion practice we described above. As indicated in one of the quotes above, it is important for information producers (especially for those on the margins or those less powerful) to control both the instantiation and the re-narration, so they can make sure they are not at a disadvantage. More powerful actors, conversely, can negotiate the elasticity of categories more easily, without worrying too much how others will re-narrate the instances they create.

Practice of Meta-narration: Manipulation of Standardized Instances

One of the benefits of a digitized classification system is that it allows for the comparison, aggregation (and other manipulations) of the individual instances. This is typically not possible with contextual narrative versions of these instances (e.g., aggregation across multiple annual reports in MS Word is labor intensive and would require, first, the data to be converted into some kind of tabular format). The right for this kind of manipulation is typically given to particular actors in an organizational hierarchy – in our cases, only higher administration has access to FP reports that enable aggregation of data. This allows higher administration to decouple various elements of an individual annual report and then compare all faculty members on one particular element, such as the number of publications. This decoupling and recombination generally results in the loss of the little contextual data left in standardized instances. Third party consumers, then, in recombining data and interpreting these new collections of instances, are, in essence, creating new stories. Thus, we call this practice meta-narration (to differentiate it from re-narration, which refers to the interpretation of instances as created by producers, without decoupling and recombining them in different ways). Because of this manipulability of data, meta-narration introduces further difficulties in achieving representational quality.

By analytically separating the human and non-human agency at work in this practice, we see that humans are confronted by the level of flexibility of the digital tool, some of which is afforded by the database logic itself (capacity for decoupling, recombination and aggregation). The more easily the tool can adapt to a variety of user needs and to changing conditions (Nelson et al. 2005), the easier it is for information consumers to seamlessly recombine individual and/or aggregate instances, as needed for a particular task. For example, third party consumers can focus on aggregating only those instances that can easily be compared across individuals (e.g., the number of peer-reviewed journal articles), while ignoring the harder to compare instances (e.g., faculty descriptions of why they produced certain kinds of research during the year). When looking at instances produced by one individual only (e.g., an annual report), third party consumers can decouple these, select only a few of them and recombine them to fit their needs (e.g., to produce a shorter overview of an annual report containing only key metrics). What becomes particularly important, then, in terms of achieving representational quality for both information producers and consumers, is ***establishing a collectively acceptable level of 'room for error or fuzziness'*** in this kind of meta-narration (Table 6).

Table 6. Overview of the Practices of Instantiation, Re-narration & Meta-narration			
Practice	Human Agency / Non-human Agency		
	Situation I: For oneself	Situation II: For similar others	Situation III: For dissimilar others
1. Practice of Instantiation 2. Practice of Re-narration	*Material:* level of *integration* of the classification system *Social:* negotiating the "elasticity of categories"	*Material:* level of *integration* of the classification system *Social:* negotiating the "elasticity of categories"; negotiating the revision of the classifications themselves; negotiating the legitimate meaning of classifications (acceptable interpretations)	
3. Practice of Meta-narration	Not relevant	Not relevant	*Material:* level of flexibility of the digital tool; database logic itself *Social:* negotiating the level of 'fuzziness' or 'room for error' in meta-narration

For example, when dealing with aggregate level instances, there is generally more fuzziness allowed to achieve acceptable representational quality (*accuracy*, in particular) of the IS:

"*When I first looked at the modified form from [FP], it looks pretty close to P&T guidelines. But then, it's in different disciplines, and individual faculty think in different ways. So, when you're throwing them all in a database and then chunking out a report, it's gonna be a bunch of noise and it probably doesn't matter at that level. At the aggregate level, it doesn't make any difference if there's three in a column that should have been in a different column. If it's going down to the level of promotion and tenure, then the stakes are much higher*." (implementation team head/faculty member, State)

However, when dealing with meta-narration on the level of individual instances (e.g., administration picking and choosing which elements of an individual annual report to look at), there is much less room for 'fuzziness' or error (good representation needs a higher level of *accuracy*, *completeness* and good *formatting*). When information producers cannot negotiate an acceptable level of fuzziness, thus not achieving a good representation, they may opt out of using the system completely:

"*I sort of opted out after I realized that it's not going to work for me without a lot of input. I can't use it as my CV. I can't print it out. I can't use it on the Web because it doesn't really reflect my story. It's sort of like 'the man' is asking questions that he wants to ask, but it doesn't tell my story. In your promotion and tenure document, I think the narrative is so important because it gives the spirit of what you're doing and you're making the connections and the cognitive jumps between … this* [FP] *is asking just for the bones. It's not asking for the person.*" (tenured faculty member/librarian, State)

We summarize the implications for achieving representational quality under different information production/consumption situations in Table 7 below.

Discussion

This paper has explored the question of how representational quality of IS (digitized classification systems, in particular) is achieved. Based on our empirical evidence, we contend that representational quality is not a static perception of the degree to which the IS is a 'good enough' (sufficiently accurate, complete, understandable, flexible, etc.) representation of all users' social worlds (cf. Wand and Weber 1995; Nelson et al. 2005). Rather, *representational quality is continuously achieved in practice through information consumers and producers negotiating the materiality of classification systems* embedded into the surface and deep structures of the software. We outlined four types of negotiations that information producers/consumers engage in: negotiating the "elasticity of categories"; the revision of classifications; the legitimate meaning of categories, and the room for error in manipulating classified instances. These negotiations are associated with three broad practices (instantiation, re-narration and meta-narration), and three different information production/consumption situations (documenting

information for oneself; for similar others, and for dissimilar others) (Markus 2001). We will now discuss the theoretical and practical implications of our findings for representational quality of digitized classification systems.

Table 7. Implications for Achieving Representational Quality (RQ) in Different Information Production/Consumption Situations

Situation I: Documenting information for oneself	Situation II: Documenting information for similar others	Situation III: Documenting information for dissimilar others
Information producers / consumers negotiate the *elasticity of categories* in relation to the integration the system affords. This negotiation focuses particularly on the achievement of the *accuracy* dimension of RQ.	Information producers / consumers negotiate the *elasticity of categories*; the *revision of classifications*; the *legitimate meaning of classifications* (overall, a *collectively acceptable interpretation* of the information) in relation to the integration the system affords. These negotiations focus particularly on the achievement of the *accuracy, completeness* and *format* dimensions of RQ.	Information producers / consumers negotiate the *elasticity of categories*; the *revision of classifications*; the *legitimate meaning of classifications* (overall, a *collectively acceptable interpretation* of the information) in relation to the integration the system affords. Information producers/consumers negotiate the *acceptable level of 'fuzziness'* (or error) in relation to the flexibility the system affords. These negotiations focus particularly on the achievement of the *accuracy, completeness* and *format* dimensions of RQ.

Theoretical Implications

First, our study takes a step towards explicitly theorizing representational quality as relational – both consumer/producer social practices and the material characteristics of the digitized classification system are given their due consideration. Prior research has commented on the fact that representational quality is contextual, i.e., it cannot be assessed independently of human agency or user needs and interpretations (cf. Nelson et al. 2005). Other research has suggested formal models and rules for creating classifications and building IS – adapting the non-human agency – so that it will provide a good representation (Parsons and Wand 2008; Wand and Weber 1995; Weber 2003). Our research contributes to these existing considerations of representational quality by demonstrating how it is achieved *at the nexus of the human and non-human agencies*. Through the unfolding of the four types of negotiations outlined above, a more or less 'good' representation of the real-world phenomena captured in the system is achieved. Human actors can be more or less successful in these negotiations. The more powerful and central actors are typically more successful in their various negotiations, because a) they have more freedom to negotiate and b) they have less to negotiate. Conversely, actors on the margins (in the minority or with less power) are typically less successful and often have more to negotiate. As shown above, often such actors engage in negotiation, but find the achievement of representational quality difficult, because other kinds of negotiations have not yet been undertaken or have not been successful. As a result, in some cases, these actors may stop negotiating and simply stop using the system (opting out, as one of our quotes suggests). This, in turn, will impact the achievement of representational quality for other actors. In sum, depending on the particular way the material and social elements come together for a particular actor, the information production and consumption for that actor is more or less meaningful (cf. Weber 2003), and representational quality is more or less achieved.

Second, we contribute to the research on standardized classification systems. Typical studies in this area focus on the performativity and transformativity of classification devices and the, often unforeseen, consequences of this (e.g., Pollock and D'Adderio 2012; White et al. 2009; Winman and Rystedt 2012). For example, the Magic Quadrant from Gartner not only describes a market and classifies the players in it, but it also creates the market (becomes a force that "moves the dots") (Pollock and D'Adderio 2012). The Common Assessment Framework (CAF), used in social work in the UK, transforms the practice of the

social workers, but ultimately remains unused as a true common language, as local practices and idiosyncrasies, in effect, remove the "common" part from CAF (White et al. 2009). Furthermore, when classifications are used for biographical information (and they do not fit with individuals' personal understanding of themselves), it can create what W.E.B. Du Bois called "double consciousness", causing feelings of incongruence and stress (Rawls 2000). While these studies offer valuable insights into the work classification devices do, they usually do not draw explicit attention to the fact that: a) classification systems are becoming more and more embedded in digital tools and this new kind of 'materialization' (e.g., the move from paper to digital software) has consequences on the interpretive and descriptive demands that these IS place upon information consumers and producers, and b) it is not classification systems per se that are of interest to information consumers/producers and shape their successful use of these IS; rather it is the representational quality of such IS. The unfortunate, in our view, side effect of information and system quality topics losing some of their appeal over recent years (cf. Nelson et al. 2005; Cordoba et al. 2012) is that these concepts are not explicitly addressed in situations where they deserve attention. For example, as information consumers, we often have no idea "what gets edited out" when we query a database or conduct an Internet search. The result is that instead of a balanced "information diet" we can end up consuming "information junk food" (Pariser 2011). Our aim in studying representational quality and in theorizing it from a practical, jointly human and non-human achievement perspective, is to renew the interest in attempts to understand the fundamental importance of representation in IS, while also expanding this topic to include more than studies of user perceptions (e.g., Nelson et al. 2005) or of ontology, conceptual modeling and deep structures (cf. Wand and Weber 2002; Parsons and Wand 2008).

Third, we raise a critical point of discussion in relation to the current 'hot' topic of "big data" (Marchand and Peppard 2013) and data-driven decision-making. Digitized classification systems or IS that represent concepts as classes or types more generally (Parsons and Wand 2008) capture, structure and present information in particular ways. This helps to reduce the complexity of dealing with individual instances of phenomena and provides a way to draw inferences about future instances, other similar phenomena, etc. (ibid.). It is unsurprising, therefore, that digitized classification systems are considered efficient vehicles for representing many different real-world phenomena (Burton-Jones and Grange 2012) and are expected to support decision-making in relation to these phenomena (Winman and Rystedt 2012). Many decisions that impact our daily lives as academics, customers, patients and citizens are, thus, based on various digitized classification systems. However, all working classification systems are to some extent incomplete and inconsistent (Bowker and Star 1998) and, as we have shown, in practice, they may be more or less familiar to the information producers and consumers, and more or less legitimate representations of their social worlds. Further, in situations where information producers are documenting information for dissimilar others (third party consumers, decision makers), they typically have the ability to decouple, recombine and aggregate data in ways about which information producers may be unaware. These meta-narratives are actually used as a basis for many decisions (e.g., budgeting, patient care, citizen services). Thus, understanding how representational quality is achieved is key in designing and implementing IS in such a way as to be aware of relevant information practices (Marchand and Peppard 2013), thereby ensuring that data-driven decision making is based on faithful and meaningful, rather than unfaithful, representations of real-world phenomena.

Practical Implications

From a practical perspective, by seeing representational quality as dynamic and achieved in practice, we advocate that tool and information designers must attend to the nuances of these practices, the audiences, and the surface and deep structures of the tool itself. While our study does *not* provide prescriptive guidelines on how to achieve representational quality of digitized classification systems, it *does* suggest some important points of intersection, where particular attention is warranted. The first intersection is at the social practices of negotiating legitimate meanings of classifications, the elasticity of categories (looser or more strict information structures), and complete revisions of classifications. These negotiations happen in relation to the level of integration afforded by the classification system. For practitioners, these negotiations represent a valuable opportunity to better understand the representational quality of their tool for different human actors, and to implement changes in its surface or deep structure.

The second intersection is at the flexibility of the digital tool and negotiating the room for 'fuzziness' in generating meta-narratives. When a digitized classification system is aimed at supporting instantiation,

re-narration and both individual- and aggregate-level meta-narration, the tool creates a temporary intertwining and transformation of different "trajectories" (Timmermans and Berg 1997). That is, for faculty members, a performance evaluation is part of their life trajectory; for administration, aggregated performance evaluations are part of the budget allocation trajectory. In FP these trajectories are forced to meet, and tensions arise at the intersection. Even if faculty members understand the value of the tool for aggregation purposes, they may not wish to classify their work into standardized instances, because the information would in addition be used for evaluating them personally, thereby not only requiring a re-narration of their story, but also the generation of a new, meta-narrative, of their story. Eliminating or cutting the performance evaluation trajectory short would, however, also disable the accreditation and budgeting trajectories, because the relevant information for both is produced on the individual level. The outcome would be that FP would not achieve representational quality in any of the information production/consumption situations. For practitioners, awareness of these 'trajectories', brought together in the tool, represents an opportunity to better understand the kinds of tensions generated at the meeting points. Based on this understanding, one option would be to implement changes in the tool, but, in this case, a more fruitful avenue may be to reconsider not the tool, but the standard procedures that various user groups follow when using the tool. For example, if faculty members have a detailed understanding of the procedures that administration follow to create meta-narratives of their annual reports, they are more likely to be willing to record their data in FP and have it be used by third party consumers.

Limitations and Future Work

Our study points to numerous further research avenues that can address the limitations of this work. First, we have examined only one particular digitized classification system in our empirical research. Our findings are, therefore, generalizable only in the theoretical sense – as conceptual arguments. Future studies of various digitized classification systems could expand and flesh out the nuances of our arguments. Second, as our effort was to describe how representational quality is achieved, our practical implications are relatively abstract. Future studies are needed to better understand how our conceptualization of representational quality could influence conceptual modeling and system development. Notwithstanding these limitations, we contend that our study has contributed to the relational theorization of representational quality and to critical discussions around the consequences of digitized classification systems on the creation of 'big data' and data-driven decision-making.

References

Baskerville, R. 2012. "Reviving the IT in the IS," *European Journal of Information Sys*tems (21:6), pp. 587–591.

Benbasat, I., and Zmud, R. W. 2003. "The Identity Crisis within the IS Discipline: Defining and Communicating the Discipline's Core Properties," *MIS Quarterly* (27:2), pp. 183-194.

Bowker, G. C., and Star, S. L. 1998. "Building Information Infrastructures for Social Worlds – The Role of Classifications and Standards," in *Community Computing and Support Systems* (LNCS 1519), T. Ishida (ed.), Berlin / Heidelberg: Springer-Verlag, pp. 231-248.

Burton-Jones, A., and Grange, C. 2012. "From Use to Effective Use: A Representation Theory Perspective," *Information Systems Research*, Articles in Advance, pp. 1–27.

Colyvas, J. A., and Powell, W. W. 2006. "Roads to institutionalization: The remaking of boundaries between public and private science," *Research in Organizational Behavior: An Annual Series of Analytical Essays and Critical Reviews* (27), pp. 305-353.

Cordoba, J-R., Pilkington, A., and Bernroider, E. W. N. 2012. "Information systems as a discipline in the making: comparing EJIS and MISQ between 1995 and 2008," *European Journal of Information Systems* (21:5), pp. 479–495.

DeLone, W. H., and McLean, E. R. 1992. "Information Systems Success: The Quest for the Dependent Variable," *Information Systems Research* (3:1), pp. 60-95.

Hayles, N. 2007. "Narrative and Database: Natural Symbionts," *PMLA* (122:5), pp. 1603–1608.

Leonardi, P. M. 2010. "Digital Materiality? How artifacts without matter, matter," *First Monday* (15:6).

Leonardi, P. M. 2012. "Materiality, Sociomateriality, and Socio-Technical Systems: What Do These Terms Mean? How Are They Related? Do We Need Them?" in *Materiality and Organizing: Social Interaction in a Technological World,* P. M. Leonardi, B. A. Nardi and J. Kallinikos (eds.), Oxford: Oxford University Press, pp. 25-48.

Leonardi, P., and Barley, S. 2010. "What's Under Construction Here? Social Action, Materiality, and Power in Constructivist Studies of Technology and Organizing," *The Academy of Management Annals* (4:1), pp. 1-51.

Manovich, L. 1999. "Database as Symbolic Form," Convergence: The International Journal of Research into New Media Technologies (5:2), pp. 80–99.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Byers, A. H. 2011. "Big data: The next frontier for innovation, competition, and productivity," McKinsey Global Institute Report, May 2011.

Marchand, D. A, and Peppard, J. 2013. "Why IT Fumbles Analytics," *Harvard Business Review* (2013:1), pp. 1–9.

Markus, M. L. 2001. "Toward A Theory of Knowledge Reuse: Types of Knowledge Reuse Situations and Factors in Reuse Success," *Journal of Management Information Systems* (18:1), pp. 57-93.

Markus, M. L. and Silver, M. S. 2008. "A Foundation for the Study of IT Effects: A New Look at DeSanctis and Poole's Concepts of Structural Features and Spirit," *Journal of the Association for Information Systems* (9:10), pp. 609-632

Milton, S. K. 2007. "Ontological Foundations of Representational Information Systems: An Australian Perspective," *Scandinavian Journal of Information Systems* (19:1), pp. 109–134.

Nelson, R. R., Todd, P. A. & Wixom, B. H. 2005. "Antecedents of Information and System Quality: An Empirical Examination Within the Context of Data Warehousing," *Journal of Management Information Systems* (21:4), pp. 199–235.

Newell, S., David, G., and Galliers, R. D. 2012. "Challenges of Coordination Using Electronic Health Records: A Genre Analysis," in *Proceedings of the 20th European Conference on Information Systems*, Barcelona, Spain, Paper 4.

Orlikowski, W. J., and Iacono, C. S. 2001. "Desperately Seeking the 'IT' in IT Research—A Call to Theorizing the IT Artifact," *Information Systems Research* (12:2), pp. 121-134.

Pariser, E. 2011. "The Filter Bubble," TED Talk, Long Beach, California. Available at: http://www.thefilterbubble.com/ted-talk.

Parsons, J., and Wand, Y. 2000. "Emancipating Instances from the Tyranny of Classes in Information Modelling", *ACM Transactions on Database Systems* (5:2), pp. 228-268

Parsons, J., and Wand, Y. 2008. "Using Cognitive Principles to Guide Classification in Information Systems Modeling," *MIS Quarterly* (32:4), pp. 839–868.

Pollock, N., and D'Adderio, L. D. 2012. "Give me a two-by-two matrix and I will create the market: Rankings, graphic visualisations and sociomateriality," *Accounting, Organizations and Society* (37:8), pp. 565-586.

Rawls, A. W. 2000. ""Race" as an Interaction Order Phenomenon: W.E.B. Du Bois's "Double Consciousness" Thesis Revisited," *Sociological Theory* (18:2), pp. 241-274.

Sabherwal, R., Jeyaraj, A., and Chowa, C. 2006. "Information System Success: Individual and Organizational Determinants," *Management Science* (52:12), pp. 1849–1864.

Snyder, I. 2004. "New media and cultural form: narrative versus database," in *Teaching English with ICT,* A. Adams and S. Brindley (eds.), London: Open University Press & McGraw Hill.

Sowa, J. 2000. Knowledge Representation: Logical, Philosophical, and Computational Foundations, Pacific Grove, CA: Brooks/Cole.

Timmermans, S., and Berg, M. 1997. "Standardization in Action: Achieving Local Universality through Medical Protocols," *Social Studies of Science* (27:2), pp. 273–305.

Wagner, E. L. 2003. "Inter-Connecting information systems narrative research: An end-to-end approach for process-oriented field studies," in *Global and organizational discourse about information technology,* Wynn, Whitley, Myers, De Gross (eds.), Boston, MA: Klewer, pp. 419-436.

Wagner, E. L., Scott, S. V. and Galliers, R. D. 2006. "The creation of 'best practice' software: Myth, reality and ethics," *Information and Organization* (16:3), pp. 251-275.

Wagner, E., & Newell, S. 2006 "Repairing ERP: Producing social order to create a working information system," *Journal of Applied Behavioral Science* (42:1), pp. 40-57.

Wand, Y., and Wang, R. 1996. "Anchoring Data Quality Dimensions in Ontological Foundations," *Communications of the ACM* (39:11), pp. 86-95.

Wand, Y., and Weber, R. 1995. "On the deep structure of information systems," *Information Systems Journal* (5:3), pp. 203-223.

Wand, Y., and Weber, R. 2002. "Research Commentary: Information Systems and Conceptual Modeling - A Research Agenda," *Information Systems Research* (13:4), pp. 363-376.

Wang, R., and Strong, D. 1996. “Beyond accuracy: What data quality means to data consumers,” *Journal of Management Information Systems* (12:4), pp. 5–34.
Watson, H., and Wixom, B. 2007. “The Current State of Business Intelligence,” *Computer*, 96–99.
Weber, R. 2003. “Conceptual Modelling and Ontology: Possibilities and Pitfalls,” *Journal of Database Management* (14:3), pp. 1-20.
Weick, K. E. 1995. *Sensemaking in organizations*, Thousand Oaks, CA: Sage.
White, S., Hall, C., and Peckover, S. 2009. “The Descriptive Tyranny of the Common Assessment Framework: Technologies of Categorization and Professional Practice in Child Welfare,” *British Journal of Social Work* (39:7), pp. 1197–1217.
Winman, T., and Rystedt, H. 2012. "Electronic Patient Records in Interprofessional Decision Making: Standardized Categories and Local Use,” *An Interdisciplinary Journal on Humans in ICT Environments* (8:1), pp. 46–64.