A Taxonomy of Classification Approaches in IS Research

Full Paper

Aurona Gerber
Department of Informatics, University of Pretoria and CAIR, Pretoria South Africa
aurona.gerber@up.ac.za

Richard Baskerville
Georgia State University, Atlanta USA and Curtin University, Perth Australia
baskerville@acm.org

Alta van der Merwe
Department of Informatics, University of Pretoria, Pretoria South Africa
alta@up.ac.za

Abstract

Even though the word *classification* appears in a number of publications in high ranking IS journals, few discussions on the fundamental aspects regarding classification could be found. Most information systems (IS) scholars intuitively embrace some classification approach as a fundamental activity in their research but without considering what classification entails. This paper reports on an investigation into classification, how classification is used within science and disciplines related to IS, as well as how it is approached within IS research itself. The main contribution of the paper is a proposed *taxonomy of classification approaches* (ToCA*)* that was validated by classifying approaches in relevant publications in three IS journals. ToCA provides a language for scholars to describe and comment, as well as understand the impact of the diverse adoption of classification approaches within IS research.

Keywords

IS Research, Classification, Taxonomy.

Introduction

The notion of classification has been embraced intuitively by scholars throughout the ages when they were confronted with any seemingly disorganized collection of *things* (abstract or real). Our need to understand such a collection compels us to set about organizing it in some way. Classifying things often seems intuitive and obvious. This intuition sometimes assumes that all classifications are somehow made in the same obvious and fundamental way, but classifications are actually made in heterogeneous ways and approached from different fundamental assumptions. Discussions about what exactly classification entails are characterized by rigorous debate from different perspectives because the notion of classification is not universal and several different definitions, theories and approaches exist within different disciplines (Parrochia and Neuville 2013). For example, classification artifacts resulting from classification often constitute the central core of research for some disciplines such as systematics (in the life sciences) or classification systems (in the information sciences), whilst the process of, or cognitive functions involved in, classification form a central object of study for disciplines such as cognitive psychology or philosophy (Bowker and Star 1999; Cohen and Lefebvre 2005).

So how do scholars in the discipline of information systems (IS) approach classification? Our interest in classification begun when we initiated research that included the development of a classification schema for IS research results. The development of such a schema would include classification but an investigation of literature revealed a lack of applicable classification discussions within IS research. This result was unexpected given that a search for the words ‘classification’, ‘taxonomy’ or ‘typology’ in the title in three of the top IS journals[1] returned more than 20 results implying that classification featured in the

[1] MIS Quarterly, Information Systems Research (ISR) and the Journal of Management Information Systems (JMIS)

main contribution of these papers. The word 'classification' within text of publications in these three journals appeared in more than 400 publications. Some form of classification is thus widely used within IS Research, but few thorough discussions about what exactly classification entails or how it is approached, could be found.

The purpose of this paper is to explore the major ways in which IS researchers approach classification. We initiated the investigation by first considering what classification entails given the history of classification. We furthermore identified classification approaches within disciplines related to IS before analysing classification within IS itself. Based on our results we propose a *taxonomy of classification approaches (ToCA)* in IS, and validated the taxonomy using document analysis by classifying classification approaches within contributions of three IS journals.

Background

On each page, your material (not including the header and footer) should fit within a rectangle of 18 x 23.5 cm (7 x 9.25 in.), centered on a US letter page, beginning 1.9 cm (.75 in.) from the top of the page. Please adhere to the US letter size only (hopefully Word or other word processors can help you with it). If you cannot do so, please contact the review coordinator for assistance. All final publications will be formatted and displayed in US letter size. Right margins should be justified, not ragged. All margins must measure 1" (2.5 cm) around. Beware, especially when using this template on a Macintosh, Word may change these dimensions in unexpected ways.

This section presents the general notions of classification by first defining classification, followed by a discussion of the use of classification within disciplines related to IS, and finally IS itself.

Several definitions of the term *classification* exist in research literature navigating across several different disciplines. Scholars describe classification as one of the basic cognitive functions of humans that is intuitively recognized (Lakoff 1987). However, exactly what classification entails and how the human mind does classification were topics of rigorous scholarly debate through the centuries. A general definition describes *classification* as the activity of categorizing or organizing objects into groups according to their similarities and differences, or their relation to a predetermined set of criteria (Parrochia and Neuville 2013). Closely associated with classification are terms such as *categorization* and *conceptualization*. *Categorization* is broadly defined as the process that sorts things, objects (or ideas) into *categories* that can be used to recognize, differentiate, order, understand and explain the world, whilst *conceptualization* is the process that define concepts where a *concept* is a term mostly used to describe an abstract category to which things can be assigned (Cohen and Lefebvre 2005; Jacob 2004).

The artifacts resulting from classification such as classifications (noun) or classification schemas, taxonomies and typologies are quite similar in purpose in that they organize objects, things and concepts into categories for ease of access and comparison. Classifications, typologies and taxonomies are often treated as synonyms (Earl 2001; Nickerson et al. 2013). Some authorities do distinguish between these artifacts where taxonomies are seen as exhaustive, hierarchical lists describing relationships between items, and classifications or typologies are non-exhaustive and can group items according to their similarities (Bailey 1994; Fiedler et al. 1996). In our study we used the terms *classification*, *taxonomy* and *typology* in our search, and we included *classification artifacts* in our analysis and proposed taxonomy.

Classification is a core activity in the life and information sciences. In the life sciences *systematics* and *taxonomies* deal with classification and nomenclature, developing rigorous classification structures for life forms (Cotterill 1995; Pavlinov 2013). The Linneaen Taxonomy groups living organisms into a series of increasingly specialized taxa due to similarities in observed characteristics that indicate commonality (Carolus Linnaeus 1758). The Linneaen taxonomy is a classification artifact that directs the way we deal with the natural world in that we describe any living organism using the notion of, for example, *species*.

In information sciences the fundamental problem of indexing and retrieving information in libraries resulted in a significant body of research on classification (Batley 2014; Jacob 2004). Key classification systems include the Dewey Decimal Classification (DDC) system (Dewey 1876; OCLC 2016), the Library of Congress Classification system (LCC) (LC 2016) and the Universal Decimal Classification (UDC) (UDC 2017). Related to this approach towards classification is the ACM Computing Classification System (CCS), one of the classification schemes most computing authors are familiar with. The ACM CCS is a 4 level classification tree authors of ACM publications use to index and retrieve publications by ACM Press (ACM 2015).

Disciplines that actively study classification are philosophy, cognitive science, cognitive psychology, mathematics and language studies (Baele 2010; Jacob 2004; Parrochia and Neuville 2013). Philosophy is generally concerned with the nature of classification, whilst cognitive science investigates the kinds of knowledge and processes that underlie human cognition in classification (Cohen and Lefebvre 2005). In these disciplines several classification theories were developed (Elliott 2010; Parrochia and Neuville 2013), which influenced branches within mathematics, statistics and computer science (Aisbett and Gibbon 1999; Baldwin 1988; Elliott 2010). There is renewed interest from neuroscience and cognitive psychology in classification, categorization and organization of information given modern technologies and phenomena such as tagging or the use of hashtags in social media (Levitin 2002, 2014).

Baily published a comprehensive overview of typologies and taxonomies as applied to the social sciences, defining classification as *grouping based on similarity* (Bailey 1994). Classification is included in methodologies for social research, for instance by Marradi (1990).

Within organizational research, work on the classification of organizations include the earlier work by Pinder and Moore (1979) that argued that taxonomy could aid theory development for organizational behaviour. McKelvey developed *organizational systematics* by applying the principles of systematics in biology, arguing that there exist natuarally occurring groupings of organizations and the work of the organizational systematist is to develop a classification reflecting the groupings using empirical methods (McKelvey 1982, 1987). Rich (1992) extended some of this thinking and developed an organizational taxonomy from empirical, theoretical and evolutionary perspectives.

Within computing classification is relevant because of digitized information in computer systems (Amgoud and Serrurier 2007). However, definitions of classification as well as approaches towards classifcation, differs. Classification in artificial intelligence (AI) include the development and application of algorithms that analyse or mine different data sources, including text, to detect and identify similarities, patterns or concepts (Fayyad et al. 1996). Aisbett and Gibbon (1999) state that classification could be described as a special case of common-sense reasoning. In data mining, classifier algorithms are used to link a set of independent variables (that represent an observation) to a dependent variable that classifies the observation (Spangler et al. 1999). Conceptual paradigms within software engineering such as object-orientation resulted in research on classification (Simons 2002, 2006; Sowa 2006), including, for instance, work on the detection and definitions of suitable classes (Parsons and Wand 2008).

The development of the Semantic Web, metadata languages and the associated tagging of information resources use classification (Bizer et al. 2009). For example, the Dublin Core Metadata Initiative (DMCI) focus on the generation of a 'library card catalogue' for the Web (DMCI 2016). Depending on the use and type of artefact, such socially constructed, collaboratively generated and open-ended classifications are for instance called tagging schemas, folksonomies or ontologies (Specia and Motta 2007).

Within IS specifically, various different classification approaches and artifacts are adopted, however, few discussions on the nature of classification exist. Nickerson et. al (2009, 2013) discuss classification as well as the lack of a method for taxonomy development. They subsequently develop two methods to construct taxonomies (Nickerson et al. 2009, 2013). Whilst this work supports the discussion on classification within IS, it lacks a comprehensive survey of the kinds of classifications and classification artifacts encountered such as meta-data tagging schemes and folksonomies (Specia and Motta 2007), supporting the motivation for this study.

A Brief History of Classification

Within this section we introduce a brief history of classification with the aim of identifying the major classification approaches through the ages.

Classical Classification (Antiquity to the 18th Century)

The typical topics of discussion in antique Greek philosophy, for instance by Aristotle and Plato, was characterized by trying to order our complex reality by, for instance, defining clear and precise logical categories and *classifying objects* in the observed world into these categories. The writings in Plato's Dialogues are evidence of this attempt to simplify the world by creating entities (archetypal forms) in a perfect reality that is manifested by a messy world empirically presented (Plato 427AD). Plato used these archetypal forms to classify 'things' in a simple, dichotomous hierarchy based on the 'distance' that separates them from the ideal form. Aristotle rejected these Platonic forms and developed his own classifications, attempting to identify the most general categories into which entities in the world divide, that are mutually exclusive and totally exhaustive, because, according to Aristotle, there could not exist cross-classifications in nature (Aristotle 384AD; Parrochia and Neuville 2013). The thinking of Aristotle was extended by Porphyry who is known for his *Introduction (to Categories)*, which is a commentary on Aristotle's categories (Porphyry 2006). In summary, Porphyry proposed a classification of substance into the five categories of being namely *genus*, *species*, *difference*, *property*, and *accident* (often depicted in medieval texts as the Porphyrian Tree). Porphyry's thinking is discernable in modern taxonomy, and also influenced later philosophical discussions on *universals* and *particulars*, thinking that are evident in modern developments such as object-orientation within computer science.

Classification in the Natural Sciences (18th Century)

The Aristotelian approach to classification remained fundamentally unaltered until the 18th century with the advent of the Victorian age and the developments in natural science that required the identification, naming and grouping of living organisms by naturalists such as Carl Linnaeus (C. Linnaeus 1758). Kant, in his Logic, summarised the most important developments regarding classifications of the time and discussed the terms and rules that the naturalist scientists empirically applied (Kant 1992). Kant used classification such that 1) members of concepts are mutually exclusive, 2) their union restores the sphere of the divided concept (i.e. all member concepts together constructs the original concept), and 3) each member can be itself divided into a subdivision (Kant 1992; Parrochia and Neuville 2013).

An essay on classification by Louis Agassiz, a biologist, geologist and natural history scholar, summarises most of the fundamental thinking regarding classification of the time and included discussions on taxonomic concepts, categories and divisions, as well as the prominent classification systems (Agassiz 1859). In general, the classification structure of the Victorian age resembles a tree, and several such tree-like classification schemes are discernible today. A famous example of this classification scheme is Charles Darwin's Tree of Life sketch in his 1837 *On the Origin of Species* notebook that shows how a genus of related species might originate by divergence from a starting point (Darwin 1876). Tree-like classification structures have acknowledged limitations since any single object can only belong to one category and the fundamental division of the main higher level concepts are crucial and reflects the adoption of a specific philosophical approach. Modern systematics continue with the refinement of this classification theory (Mayr 1999; Pavlinov 2013).

At the end of the Victorian age, *combinatorial classification* evolved, mostly due to developments in Chemistry where an object can be classified according to two (or multiple) dimensions, often visually depicted as a table, for instance the Mendeleyev or Periodic table where chemical elements are ordered using their atomic numbers, electron configurations and chemical properties (RSC 2004).

Classification of Knowledge (Modern Developments)

Information Science has undoubtedly influenced the thinking and theory of classification fundamentally (Jacob 2004). At the end of the 19th century, the developments in scientific research and the associated publications that were stored in libraries resulted in the questions associated with the storage and retrieval of information from extensive library catalogues, as well as the maintenance of these catalogues of knowledge that grow indefinitely. The subsequent classification systems include the Dewey Decimal Classification (DDC) system (Dewey 1876) and the Library of Congress Classification system (LCC) (LC 2016). The fundamental assumptions in these classification systems, similar to the natural sciences, still included completeness (everything is classified), exclusiveness (everything belong to only one category alone) and all categories have members (no empty category exists). These assumptions stem from the fact that a physical object (typically a book) that exists, needs to be stored at a single location on a shelf, and it should subsequently be possible to retrieve the object from this location whenever necessary.

Growing library collections resulted in work on faceted classification (Hjørland 2013). Ranganathan developed a classification system based on 5 facets (PMEST) that evolved into the Universal Decimal Classification (UDC) system (Ranganathan 1946). Different from previous classification systems, a main facet is divided into subfacets using specific characteristics, which are then combined as necessary to form subcategories (UDC 2017).

Information classification approaches evolved significantly with the development of digitized information. Digital information objects are not restricted by physical locations and therefore could be classified to belong to several categories. Furthermore, categories have do not need to be exhaustive or distinct i.e. categories can overlap and abstract categories without members may exist. This approach to classification is for instance represented by the vision of metadata and tagging, linked data and the Semantic Web (Berners-Lee et al. 2001; Bizer et al. 2009). Metadata categories are often represented using the set of language standards maintained by the World Wide Web Consortium (W3C) (W3C 2017).

The Taxonomy of Classification Approaches (ToCA)

In the history of classification two major approaches to classification can be distinguished, namely the conceptual approach (which is represented by the approaches in the classical and information sciences periods) and the empirical approach (which is represented by periods dominated by the natural sciences). These two approaches are also identified by Marradi (1990), who termed the conceptual approach *intentional classification*, and the empirical approach, *extensional classification* (Marradi 1990). The two major classification approaches are furthermore supported by several other authors on classification related to IS such as Rich (1992), McKelvey (1982), (Fiedler et al. 1996), (Pinder and Moore 1979) and (Nickerson et al. 2013). Borrowing from these authors we refine the description of the two major classification approaches as follows:

Intensional classification (IC) is a top-down approach where a general class is subdivided. The approach is also termed a theoretical approach using *a priori* classes and heuristics to sort objects into these classes, or a deductive, *conceptual to empirical* approach.

Intensional classification artifacts: Some authors claim that the intensional classification approach develops typologies, which is defined as a structure that classify subjects by forcing deductive assignment to *a priori* predefined groups or types (Bailey 1994), but Nickerson et al. (2013) differ from this view because of the adoption of the terms *typology* and *taxonomy* as synonyms within IS research. They prefer to call all classification artifacts taxonomies. Relevant examples associated with IS research of the intensional approach would include the ACM CCS where publications are assigned predefined keywords from a 4 level classification system (ACM 2015).

Extensional classification (EC) is a bottom-up approach where the characteristics of individuals are analyzed and grouped into classes from specific to general based on perceived similarities. This classification approach is also termed an *empirical to conceptual* approach, sorting things on the basis of similarity or contrast into a hierarchy of classes. Extensional classification thus aligns with the approaches followed by naturalist systematics.

Extensional classification artifacts: Several authors claim that a taxonomy is the product of extensional classification. Taxonomies determine membership into *a posteriori* categories that emerge from empirical analysis inductively (Fiedler et al. 1996). Taxonomies so developed are usually mutually exclusive, exhaustive and covering, meaning that an individual cannot belong to more than one category, and all individuals can be categorised given the taxonomy. An example of this approach associated with IS research would be the organizational systematics by McKelvey who applied the principles of systematics in biology to organizations (refer to discussion in the background section) (McKelvey 1982, 1987).

In addition to the two main classification approaches, we identified a third dimension that resonates with a cognitive science perspective where humans intuitively categorize objects in the world in order to make sense. Classification authors identify this third approach, however, it is often vague, for instance described as *common sense classification* (Rich 1992) or *categorical assignment* (Pinder and Moore 1979). The *classing* of Marradi (1990) is described as the 'assigning of objects or events to classes or types which have been previously defined' and the *intuitive* approach of Nickerson et. al (2009) is described as *essentially* *ad hoc* without an explicit method where the researcher uses 'his or her understanding of the objects to be classified to propose a taxonomy based on the researcher's perceptions of what makes sense'.

We identified this intuitive approach within both the intensional and extensional approaches whenever human intuition was applied. We therefore propose a distinction in the classification approaches taxonomy between *formal intensional* and *intuitive intensional* as well as *formal extensional* and *intuitive extensional*. A *formal classification approach* (extensional or intensional) is rigorous and the resulting classification artifact is usually a structured, multi-level, hierarchical, exhaustive and exclusive taxonomy, and would include an approach that stipulates an automated technique such as cluster analysis. An intuitive classification approach (intensional or extensional) is intuitive. The resulting artifact is often multi-dimensional, non-exclusive without a strict hierarchy. Individuals could often belong to more than one category and the categories are not exhaustive i.e. new categories could be added as needed. ToCA is presented in Table 1 depicting a high-level division between intensional and extensional approaches, each then further subdivided into formal or intuitive approaches, as well as possible classification artifacts.

Classification Approaches		**Possible Classification Artifact**
Intensional Classification Approach (IC) Top-down, conceptual, deductive.	**Formal Intensional Classification (FIC)** Explicit top-down deductive method.	Multi-level or hierarchical taxonomy, often developed using formal classification methods, e.g. Dewey DC system (Dewey 1876).
	Intuitive Intensional Classification (IIC) Intuitive top-down deductive method, not explicitly described.	A taxonomy developed by identifying top-level concepts, often from literature. Examples include folksonomies or metadata tagging schemas
Extensional Classification Approach (EC) Bottom-up, empirical, inductive.	**Formal Extensional Classification (FEC)** Explicit bottom-up inductive method.	Often a multi-level taxonomy, or quantifiable two dimensional classification schema e.g. Organizational Taxonomy (McKelvey 1982, 1987).
	Intuitive Extensional Classification (IEC) Intuitive bottom-up inductive method, not explicitly described.	Both a multi-level / hierarchical taxonomy or a multi-dimensional (tabular) classification schema intuitively developed from empirical evidence. E.g. the framework for MIS development projects (Moore 1979).

Table 1: Taxonomy of Classification Approaches (ToCA)

For the purpose of validating ToCA, we analyzed classification approaches in IS research. We collected all papers from three of the top IS journals (MIS Quarterly, Information Systems Research (ISR) and the Journal of Management Information Systems (JMIS)) with the words 'classification', 'taxonomy' or 'typology' in the title. We collected 21 papers, but since Barki's 2 publications on an IS keyword classification scheme are related, we combined the two papers as one entry (Barki et al. 1988, 1993). We furthermore removed four papers from the data set because these do not develop a classification system:

- Spangler et al. (1999) present the choosing of a method for multiple classification in data mining;
- Zhou et al. (2004) experimentally compare algorithmic classification methods for the automatic prediction of deception;
- Zurada (2012) investigates the impact of removing and replacing missing values in large data sets;
- Parsons and Wand (2008) formalize ideas underlying effective classification.

Using document analysis, a systematic procedure for interpreting document content using predetermined terms (Bowen 2009) (in this case extracted from ToCA) the remainder of the collected papers were analysed and the primary classification approach and artifact was identified (depicted in Table 2).

The results of our analysis indicate that the intensional classification approach is more prevalent than the extensional approach in contributions in IS research (9 vs. 6). Furthermore, formal and intuitive classification approaches are adopted nearly equally overall. An interesting result is that the intuitive approach is less frequent within extensional classification than the formal approach (2 vs. 4) possibly because IS reviewers and authors have a higher regard for the formal empirical methods that is inherent in extensional classifications. With regards to the second level distinction in the taxonomy the analysis found that formal approaches (both extensional or intensional) adopted a rigorous method (often supported by automation) and this formal method usually provides key evidence to support the research contribution. On the other hand intuitive approaches adopt informal arguments often using literature or existing models to motivate the classification. In these examples the proof of the classification contribution is not the method of construction but the greater application utility of the contribution. Demonstrating the usefulness and effectiveness of the classification system in practical operation typically shows this "greater application utility". In the results the classification artifacts (third distinction in the taxonomy) are more varied than anticipated. We expected rigorous, hierarchical classification artifacts from the extensional approaches due to the empirical inductive nature, and multi-dimensional, non-exhaustive classification artifacts from intensional approaches due to the deductive and conceptual nature. The results do indicate a tendency towards more structure in the classification artifacts of extensional approaches, but the evidence is not conclusive, possibly because the paper sample is quite limited.

Source	ToCA Approach		Classification Artifact
Belardo and Pazer (1985)	IIC	Build on previous work and identify further dimensions using literature.	2 Dimensional Scope/ Complexity Framework (diagram) depicted as a graph, subdivisions (breadth, depth, uncertainty, variety).
Barki et al.(1988, 1993)	FIC	Start with op-level categories, subdivide with a formal process.	Keyword classification scheme. Rigorous tree-structure taxonomy.
Tan and Benbasat (1990)	IEC	Review literature and intuitively develop two taxonomies.	Develop two task taxonomies (two-dimensional) for classifying tasks.
Ein-Dor and Segev (1993)	IEC	Seventeen major types of information systems, identified from literature, assign metrics.	A two-dimensional classification scheme for Information Systems, subdivisions on dimensions. ISs are classified by numerical methods.
Sabherwal and King (1995)	FEC	Empirically identify categories and then do empirical cluster analysis to divide.	Empirical taxonomy of strategic IS decision processes.
Fiedler et al. (1996)	FEC	Empirically develops a taxonomy using cluster analysis.	Taxonomy that match IT to organizational structure.
Earl (2001)	IIC	Develop typology based on descriptive and deductive research.	Typology of 'schools' of knowledge management". Claims that schools (as top categories) are not mutually exclusive.
McKnight et al.(2002)	IIC	Typology: high-level concepts identified and verified by analysis.	Integrative Typology: 4 high-level constructs, sub-divided with associated measures.
Larsen (2003)	FEC	Clearly articulated extensional approach, include cluster analysis.	Taxonomy of 12 general categories is created, and existing research within each category is examined.
Son and Kim (2008)	IIC	No clear method for the 'taxonomy' articulated, extended into a model	A taxonomy of information privacy-protective responses (IPPR): 6 types of behavioral responses, 3 categories.
Li et al.(2009)	FIC	Formal method/algorithm to do classification of patent documents.	A patent classification schema enhanced with meta-data on knowledge evolution.
McKinney Jr and Yoos (2010)	IIC	Method not articulated. Taxonomy substantiated with literature as evidence.	A two-dimensional tabular taxonomy of four top-level views on information.
Vlas and Robinson (2012)	FIC	Use existing categories. Requirements are classified with NLP.	A classifier that is a kind of weak ontology-based information extraction (OBIE) system.
Tsang and Williams (2012)	IIC	No method, extends existing work.	A classification of induction: 5 types of generalization.
Posey et al.(2013)	FEC	Use a systematics approach.	A multi-dimensional taxonomy adversity for protection-motivated behaviors (PMBs).
Prat et al.(2015)	FIC	Use described method of Nickerson.	A taxonomy of evaluation methods for IS artifacts
Totals: FIC (4), IIC (5), FEC (4), IEC (2). Intensional (9), Extensional (6), Formal (8), Intuitive (7).			

Table 2: Using ToCA to Classify ISR Classification Approaches.

Contribution, Discussion and Conclusion

Our work is concerned with the question of how scholars in the IS discipline approach classification. Forms of classification are evident in a large number of prominent research contributions within IS, however, we found little discussion about applicable IS classification approaches and artifacts. Our work is anchored in the history of classification that suggested two distinct classification approaches namely *intensional*, which is an top-down, deductive and conceptual classification approach working from the general towards the specific (evident in the classical and information sciences eras), and *extensional*, which is an bottom-up, empirical, inductive classification approach analysing similarities and differences of objects to establish common categories working from the specific to the general (evident in the eras dominated by natural sciences). We extend the work on classification by defining sub-categories namely *intuitive* and *formal*. The combined results that embody our *taxonomy of classification approaches* (ToCA) are depicted in Table 1. We adopt the position of Nickerson et al. (2013) who call all IS classification artefacts taxonomies. Our classification approach could be described as an intuitive intensional classification approach and our contribution is a taxonomy or a structure that classify objects by deductive assignment to predefined types. Using our taxonomy, we analysed a set of papers representing IS research collected from three top IS journals (MIS Quarterly, ISR and JMIS) with the words 'classification', 'taxonomy' or 'typology' in the title. We found that:

- IS scholars adopt both intensional and extensional approaches but with a bias towards intentional approaches.
- The formal and intuitive intensional classification approaches were adopted nearly equally, whilst there were significantly more formal than intuitive extensional approaches in our sample.
- Formal approaches often provide supporting evidence for the classification artifact, whilst intuitive approaches support a classification artifact contribution with evidence of its practical usefulness and application utility.
- Classification artifacts identified are varied with a tendency towards more structure in the classification artifacts of extensional approaches. The evidence is not conclusive, possibly because the paper sample is limited, and warrants further research.

The adoption of both intensional and extensional approaches by IS scholars is arguably because of the heritage of IS that borrows from several disciplines. We initially expected a larger focus on the extensional classification approaches because of its foundations in empirical methods. To our surprise, the paper sample actually yielded a larger number of contributions using the intensional approach that start with analytical concepts. A possible reason for this might be the intangible nature of IS and it's associated artifacts that necessitate that any examination of IS is initiated by presenting a conceptual model. This rationale is supported by the fact that the most common approach is the intuitive intensional classification approach. However, the approach that is the least evident is intuitive extensional, possibly because of the rigorous nature of extensional approaches to limit researcher bias and exclude intuition. The bias toward intensional approaches may also reflect the foundations of IS in Systems Science, which is a perspective with a tendency to regard phenomena as systems that might be decomposed into subsystems as a means toward understanding its nature. In keeping with the mainstream tenets of Systems Science, IS research typically attacks complexity by analysis and decomposition, or a top-down approach with an upper concept divided into components. A further consequence of the higher frequency of intensional approaches could suggest that IS scholars may have an innate bias toward certain kinds of classifications. For example, for our field, it may be less convincing to justify classifications based on extensional principles, especially intuitive extensional approaches. While this bias certainly does not exclude extensional classifications, it may mean that examination boards, peer publication reviewers, editors, and the like, will struggle a bit more in accepting contributions resulting from such classifications, and being cognizant of the differences in classification approaches as presented in our taxonomy, the negative impact of such biases could be minimized.

ToCA contributes a language that is useful for characterizing and distinguishing different ways IS scholars approach the development of classification contributions. It diminishes the notion that classifications are a homogenous phenomenon, and helps us to approach different types of classification accordingly. This taxonomy helps us understand how to position future work convincingly and to appraise different types of classifications differently. Formal intensional or extensional approaches typically need to be rigorous in the development and description of the classification method, whilst an intuitive approach would need convincing evidence of the application and use of the classification artifact. This taxonomy also guides us away from possible pitfalls, such as failing to recognize the validity of both intensional and extensional classification approaches within IS research.

In conclusion, we developed a *taxonomy of classification approaches* (ToCA) in order to understand and comment on the diverse adoption of classification approaches within IS research, which was validated using document analysis on IS research papers. The goal of ToCA is to provide a language for describing different kinds of classification within IS research, and the possible impact of the adopted approaches. Using ToCA as starting point, further research may refine ToCA by considering classification criteria and evaluation methods for classification (Mayr 1999). Refinement of ToCA can improve the use of classification more effectively within IS research.

REFERENCES

ACM. 2015. "ACM Computing Classification System," (available at http://www.acm.org/about/class; retrieved July 14, 2016).

Agassiz, L. 1859. *An Essay on Classification*, The Bavarian State Library: Longman, Brown, Green, Longmans & Roberts & Trübner.

Aisbett, J., and Gibbon, G. 1999. "Cognitive Classification," in *AAAI-99 Proceedings*, pp. 100–107.

Amgoud, L., and Serrurier, M. 2007. "Arguing and Explaining Classifications," Presented at the AAMAS'07, Honolulu, Hawai'i, USA.: ACM, May 14.

Aristotle. 384AD. "Internet Sacred Text Archive: Aristotle," *Internet Sacred Texts Archive*, 322 BC (available at http://www.sacred-texts.com/cla/ari/index.htm; retrieved November 30, 2016).

Baele, S. 2010. "Interdisciplinary Introduction to Categorization: Interview with Amie Thomasson (philosophy), Dvora Yanov (political sciences), and Thomas Serre (artificial intelligence)," (available at http://www.revue-emulations.net/archives/n8/categentretien).

Bailey, K. D. 1994. *Typologies and taxonomies: an introduction to classification techniques.*, Quantitative applications in the social sciences, Thousand Oaks, California: Sage Publications.

Baldwin, J. T. (Ed.). 1988. *Classification Theory.*, Lecture Notes in Mathematics, Springer.

Barki, H., Rivard, S., and Talbot, J. 1988. "An Information Systems Keyword Classification Scheme.," *MIS Quarterly*, (12:2), p. 299.

Barki, H., Rivard, S., and Talbot, J. 1993. "A Keyword Classification Scheme for IS Research Literature: An Update.," *MIS Quarterly*, (17:2), p. 209.

Batley, S. 2014. *Classification in theory and practice.*, Chandos information professional series, (2nd edition.), Waltham, MA: Chandos Pub., an imprint of Woodhead Pub.

Belardo, S., and Pazer, H. L. 1985. "Scope/Complexity: A framework for the classification and analysis of information-decision systems," *Journal of Management Information Systems*, (2:2), pp. 55–72.

Berners-Lee, T. J., Hendler, J., Lassila, O., and others. 2001. "The Semantic Web," *Scientific American*, (284:5), pp. 28–37.

Bizer, C., Heath, T., and Berners-Lee, T. 2009. "Linked Data - The Story So Far," *International Journal on Semantic Web and Information Systems*, (5:3), pp. 1–22.

Bowen, G. A. 2009. "Document Analysis as a Qualitative Research Method," *Qualitative Research Journal*, (9:2), pp. 27–40 (doi: 10.3316/QRJ0902027).

Bowker, G. C., and Star, S. L. 1999. *Sorting things out: classification and its consequences*, Inside technology, Cambridge, Mass: MIT Press.

Cohen, H., and Lefebvre, C. 2005. *Handbook of Categorization in Cognitive Science - 1st Edition*, (1st ed.), Elsevier Science.

Cotterill, F. P. D. 1995. "Systematics, biological knowledge and environmental conservation," *Biodiversity & Conservation*, (4:2), pp. 183–205.

Darwin, C. 1876. *The Origin of Species By Means of Natural Selection*, Cambridge Library Collection, Cambridge University Press.

Dewey, M. 1876. *Classification and Subject Index for Cataloguing and Arranging the Books and Pamphlets of a Library*, Project Gutenberg.

DMCI. 2016. "Dublin Core Metadata Initiative," *Metadata Innovation* (available at http://dublincore.org/; retrieved January 25, 2017).

Earl, M. 2001. "Knowledge Management Strategies: Toward a Taxonomy," *Journal of Management Information Systems*, (18:1), p. 215.

Ein-Dor, P., and Segev, E. 1993. "A classification of information systems: Analysis and interpretation," *Information Systems Research*, (4:2), pp. 166–204.

Elliott, G. A. 2010. "Towards a theory of classification," *Advances in Mathematics*, (223), pp. 30–48.

Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. 1996. *Advances in knowledge discovery and data mining*, (Vol. 21), AAAI press Menlo Park.

Fiedler, K. D., Grover, V., and Teng, J. T. C. 1996. "An Empirically Derived Taxonomy of Information Technology Structure and Its Relationship to Organizational Structure," *Journal of Management Information Systems*, (13:1), p. 9.

Hjørland, B. 2013. "Facet analysis: The logical approach to knowledge organization," *Information Processing & Management*, (49:2), pp. 545–557 (doi: 10.1016/j.ipm.2012.10.001).

Jacob, E. K. 2004. "Classification and categorization: a difference that makes a difference," *Library Trends*, (52:3).

Kant, I. 1992. *Lectures on logic*, The Cambridge edition of the works of Immanuel Kant, Cambridge ; New York: Cambridge University Press.

Lakoff, G. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind.*, Chicago and London: University of Chicago Press.

Larsen, K. R. 2003. "A taxonomy of antecedents of information systems success: variable analysis studies," *Journal of Management Information Systems*, (20:2), pp. 169–246.

LC. 2016. "Library of Congress Classification System," *Library of Congress Classification* (available at http://www.loc.gov/aba/cataloging/classification/; retrieved November 20, 2016).

Levitin, D. J. (Ed.). 2002. *Foundations of cognitive psychology: core readings*, Cambridge, Mass: MIT Press.

Levitin, D. J. 2014. *The organized mind: thinking straight in the age of information overload*, New York: Penguin Group.

Li, X., Chen, H., Zhang, Z., Li, J., and Nunamaker, J. F. 2009. "Managing Knowledge in Light of Its Evolution Process: An Empirical Study on Citation Network-Based Patent Classification," *Journal of Management Information Systems*, (26:1), pp. 129–154 (doi: 10.2753/MIS0742-1222260106).

Linnaeus, C. 1758. *Systema naturae per regna tria naturae :secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis (in Latin) (10th ed.).*, Stockholm: Laurentius Salvius.

Linnaeus, C. 1758. *Systema naturae. 10th ed.*, (Vol. v.1), Laurentii Salvii, Stockholm.

Marradi, A. 1990. "Classification, typology, taxonomy," *Quality & Quantity*, (24:2), pp. 129–157.

Mayr, E. 1999. *Systematics and the Origin of Species from the Viewpoint of a Zoologist*, Harvard University Press.

McKelvey, B. 1982. *Organizational Systematics: Taxonomy, Evolution, Classification*, University of California Press.

McKelvey, B. 1987. "Organizational Systematics: Taxonomic Lessons from Biology," *Management Science*, (24:13), pp. 1428–1440.

McKinney Jr, E. H., and Yoos, C. J. 2010. "Information about information: A taxonomy of views," *MIS Quarterly*, pp. 329–344.

McKnight, D. H., Choudhury, V., and Kacmar, C. 2002. "Developing and Validating Trust Measures for e-Commerce: An Integrative Typology," *Information Systems Research*, (13:3), p. 334.

Moore, J. H. 1979. "A framework for MIS software development projects," *Mis Quarterly*, pp. 29–38.

Nickerson, R. C., Varshney, U., and Muntermann, J. 2013. "A method for taxonomy development and its application in information systems," *European Journal of Information Systems*, (22:3), pp. 336–359.

Nickerson, R., Muntermann, J., Varshney, U., and Isaac, H. 2009. "Taxonomy development in information systems: Developing a taxonomy of mobile applications," in *European Conference in Information Systems*.

OCLC. 2016. "Online Computer Library Center (OCLC)," (available at https://www.oclc.org/; retrieved November 20, 2016).

Parrochia, D., and Neuville, P. 2013. *Towards a General Theory of Classifications*, Basel: Springer Basel.

Parsons, J., and Wand, Y. 2008. "Using cognitive principles to guide classification in information systems modeling," *MIS quarterly*, pp. 839–868.

Pavlinov, I. Y. 2013. "Taxonomic theory for nonclassical systematics," *Biology Bulletin Reviews*, (3:1), pp. 17–26 (doi: 10.1134/S2079086413010088).

Pinder, C. C., and Moore, L. F. 1979. "The Resurrection of Taxonomy to Aid the Development of Middle Range Theories of Organizational Behavior," *Administrative Science Quarterly*, (24:1), p. 99 (doi: 10.2307/2989878).

Plato. 427AD. "Dialogues of Plato," *Internet Sacred Texts Archive*, 347 BC (available at http://www.sacred-texts.com/cla/plato/index.htm; retrieved October 10, 2016).

Porphyry. 2006. *Porphyry Introduction*, Clarendon Press.

Posey, C., Roberts, T., Lowry, P. B., Bennett, B., and Courtney, J. 2013. "Insiders' protection of organizational information assets: Development of a systematics-based taxonomy and theory of diversity for protection-motivated behaviors," *MIS Quarterly*, (37:4), pp. 1189–1210.

Prat, N., Comyn-Wattiau, I., and Akoka, J. 2015. "A Taxonomy of Evaluation Methods for Information Systems Artifacts," *Journal of Management Information Systems*, (32:3), pp. 229–267 (doi: 10.1080/07421222.2015.1099390).

Ranganathan, S. R. 1946. *Suggestions for the Organization of Libraries in India*, Oxford University Press.

Rich, P. 1992. "The organizational taxonomy: Definition and design," *Academy of Management Review*, (17:4), pp. 758–781 (doi: 10.5465/AMR.1992.4279068).

RSC. 2004. "Dmitri Mendeleev in The History of the Periodic Table," *The History of the Periodic Table* (available at http://www.rsc.org/education/teachers/resources/periodictable/pre16/develop/mendeleev.htm).

Sabherwal, R., and King, W. R. 1995. "An Empirical Taxonomy of the Decision-Making Processes Concerning Strategic Applications in Information Systems," *Journal of Management Information Systems*, (11:4), p. 177.

Simons, A. J. H. 2002. "The theory of classification part 1: perspectives on type compatibility," *Journal of Object Technology*, (1:1), pp. 55–61.

Simons, A. J. H. 2006. "Theory of Classification," *University of Sheffield: Anthony J H Simons* (available at http://staffwww.dcs.shef.ac.uk/people/A.Simons/classify/).

Son, J.-Y., and Kim, S. S. 2008. "Internet users' information privacy-protective responses: a taxonomy and a nomological model," *MIS Quarterly*, pp. 503–529.

Sowa, J. F. 2006. "Categorization in Cognitive Computer Science," in *Handbook of Categorization in Cognitive Science*, Elsevier Science, p. 141–163.

Spangler, W. E., May, J. H., and Vargas, L. G. 1999. "Choosing Data-Mining Methods for Multiple Classification: Representational and Performance Measurement Implications for Decision Support," *Journal of Management Information Systems*, (16:1), p. 37.

Specia, L., and Motta, E. 2007. "Integrating Folksonomies with the Semantic Web," in *The Semantic Web: Research and Applications*, E. Franconi, M. Kifer, and W. May (eds.), (Vol. 4519), Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 624–639 (doi: 10.1007/978-3-540-72667-8_44).

Tan, J. K., and Benbasat, I. 1990. "Processing of graphical information: A decomposition taxonomy to match data extraction tasks and graphical representations," *Information Systems Research*, pp. 416–439.

Tsang, E. W., and Williams, J. N. 2012. "Generalization and induction: Misconceptions, clarifications, and a classification of induction," *MIS Quarterly*, (36:3), pp. 729–748.

UDC. 2017. "Universal Decimal Classicication," *UDC*, UDC Consortium, (available at http://www.udcc.org/; retrieved January 19, 2017).

Vlas, R. E., and Robinson, W. N. 2012. "Two Rule-Based Natural Language Strategies for Requirements Discovery and Classification in Open Source Software Development Projects," *Journal of Management Information Systems*, (28:4), pp. 11–38 (doi: 10.2753/MIS0742-1222280402).

W3C. 2017. "World Wide Web Consortium," (available at https://www.w3.org/; retrieved October 12, 2016).

Zhou, L., Burgoon, J. K., Twitchell, D. P., Qin, T., and Nunamaker Jr, J. F. 2004. "A comparison of classification methods for predicting deception in computer-mediated communication," *Journal of Management Information Systems*, (20:4), pp. 139–166.

Zurada, J. 2012. "Does Removing/Replacing Missing Values Improve The Models' Classification Performances?," *International Journal of Management & Information Systems (Online)*, (16:3), p. 215.