Basic Classes in Conceptual Modeling: Theory and Practical Guidelines

Arturo Castellanos[1], Monica Chiarini Tremblay[2], Roman Lukyanenko[3], Binny Samuel[4]

[1]Baruch College (CUNY), USA, arturo.castellanos@baruch.cuny.edu
[2]William and Mary, USA, monica.tremblay@mason.wm.edu
[3]HEC Montréal, Canada, roman.lukyanenko@hec.ca
[4]University of Cincinnati, USA, samuelby@uc.edu

Abstract

Since the 1970s, many approaches to representing domains have been suggested. Each approach maintains the assumption that the information about the objects represented in the information system (IS) is specified and verified by domain experts and potential users. Yet, as more IS are developed to support a larger diversity of users such as customers, suppliers, and members of the general public (e.g., in the case of many multiuser online systems), analysts can no longer rely on a stable single group of people for the complete specification of domains; therefore, prior research has questioned the efficacy of conceptual modeling in these heterogeneous settings. This paper aims to address this problem by providing theoretical foundations rooted in psychology research supporting the existence and importance of special classes that are termed basic-level categories. Based on this research, we formulate principles for identifying basic classes in a domain. These classes can guide conceptual modeling, database design, and user interface development in a wide variety of traditional and emergent domains.

Keywords: Conceptual Modeling, Basic-Level Categories, Ontology, Psychology

1 Introduction

Although conceptual modeling has a long history (Borgida, 1985, p. 1; Brodie, 1984; Hirschheim & Heinz, 2010; Peckham & Maryanski, 1988), until recently, it has been predominantly conducted in internal organizational settings (e.g., to develop transaction processing systems). Greater control in organizational environments has made it relatively feasible to reach and engage domain experts and future users of a system to collect complete, consistent, stable, and agreed upon information requirements (Checkland & Holwell, 2006; Hirschheim, Klein, & Lyytinen, 1995).

However, dramatic changes in the information systems (IS) landscape have occurred over the last decade, including developments such as "big data" and big data technologies (e.g., NoSQL databases, Hadoop), social media and rapid content creation online by regular users, mobile and ubiquitous computing, business analytics, sensors, internet of things, and artificial intelligence. These changes have dramatically altered the ways in which IS are designed and used, necessitating innovative approaches to conceptual modeling to better support these developments (Jabbari Sabegh et al., 2017; Rai, 2016, 2017; Recker, 2015; Storey & Song, 2017).

In this paper, we focus on one major development: the explosive growth of information created by ordinary people (as opposed to organizational employees), known as user-generated content (UGC). As of 2015, more than two-thirds of smartphone users report using their devices to create digital content online (Gantz & Reinsel, 2012; Melumad, Inman, & Pham, 2019; A. Smith & Page, 2015). UGC takes on many forms, including forums, chats, product reviews, social networking (e.g., Facebook, Twitter, Instagram), as well as custom content creation platforms (e.g., YouTube, Flickr, WordPress, GalaxyZoo, Amazon's Mechanical Turk, Slack) (Brynjolfsson, Geva, & Reichman, 2015; Johnson, Safadi, & Faraj, 2015; Levina & Arriaga, 2014; Luo et al., 2013; Susarla, Oh, & Tan, 2012; Wattal et al, 2010).

The contextual environment for UGC differs from traditional corporate settings, which challenges some of the assumptions of conceptual modeling (Jabbari Sabegh et al., 2017). One difficulty is finding appropriate domain structures that are natural and familiar to all users in open and anonymous settings like UGC. Major conceptual modeling grammars like UML and ER rely on domain structures such as classes (also known as concepts, categories, kinds, or entity types).[1] Classes distill essential features of objects for storage and use in an IS (Borgida, 1985; Parsons & Wand, 1997). Once specified, classes constrain the user input that can be captured and used, directly impacting IS objects such as database tables, data collection fields, user interface options, and reports (Hirschheim et al., 1995; Teorey, Yang, & Fry, 1986). The identification of classes is traditionally one of the most important steps in IS development:

> *The first step in designing a database, a knowledge base, or an object-oriented system is to select [an] appropriate collection of ontological categories … the selection of categories determines everything that can be represented in a computer application or an entire family of applications. Incompleteness, distortions, or restrictions of the framework of categories must inevitably limit the flexibility and generality of every program and database that use those categories* (Sowa, 1995, p. 670).

Compared with traditional corporate environments, class selection may be even more critical in UGC settings. Online usage is often volitional—i.e., not mandatory—and thus users may abandon an online

system at any time for any reason without giving advance warning to project owners. In UGC settings, if the classes chosen in the model do not match those preferred by users, uncommitted users may produce low-quality data (e.g., because of low domain expertise or difficulty interpreting the classes) (He & Wiggins, 2015; Kosmala et al., 2016; Lewandowski & Specht, 2015; Lukyanenko, Parsons, & Wiersma, 2014b) or may become less engaged with the project (Lukyanenko, Parsons, & Wiersma, 2014a; Nov, Arazy, & Anderson, 2011; van Kleek et al., 2011), thus limiting the effectiveness of UGC.

In this paper, we propose a novel approach in conceptual modeling and IS development that, while generally applicable, is especially tailored to UGC. Following theoretical foundations in psychology on basic-level categories (BLCs), we suggest selecting basic classes (BCs) during conceptual modeling. In the context of information management and IS development, we define basic classes as labeled sets of attributes of objects or events that are most familiar to and shared among all system users regardless of the users' backgrounds, knowledge, and domain expertise.

BCs are classes for which user consensus on both the label and the attribute set is likely to be high regardless of the diversity of the user base. For example, most people are familiar with the class "bird" and have similar conceptions of common bird attributes (e.g., has wings, has beak, lays eggs, has feathers, most can fly). This contrasts with more specialized classes, such as "greater yellowlegs," which require specialized domain knowledge that may not be common in the general population. The idea of basic, preferred, or universal classes is not new, but despite arguments and evidence supporting the benefits of such classes (Lukyanenko et al., 2019; Lukyanenko et al., 2014b), no work in IS has deeply investigated the nature of these classes or provided guidelines for their identification and application. Our work adds a novel perspective to existing IS development research that has sought to support specialized tasks and thus has mainly focused on determining appropriate, typically specialized, classes. In our work, we explore unchartered territory for conceptual modeling and IS development—selecting generic classes maximally familiar to all users. As we discuss and show in our paper, such classes of maximal agreement can be beneficial in a wide range of applications, such as UGC, mobile apps, and even the design of traditional software.

---

[1] Following research practice, we use the terms (e.g., classes, attributes) based on the domain of discourse. Conceptual modeling research typically uses terms such as (1) classes, sets, or entity types, (2) objects, members, entities, or instances, (3) attributes or properties; whereas psychology

prefers terms such as (1) concepts, categories, or kinds, (2) objects, individuals, or members, and (3) features, characteristics, attributes, or properties, respectively (Murphy, 2004; Parsons & Wand, 1997; Sowa, 1995)

The rest of the paper is organized as follows: In the next section, we provide a motivating example to elucidate the challenges of UGC without BCs. We then survey state-of-the-art literature in conceptual modeling and psychology and use it as a foundation to formulate principles for identifying BCs in a domain. We develop practical guidelines for the identification and use of BCs and evaluate these guidelines.

2 Motivating Example and Problem Formulation

To better understand conceptual modeling challenges in UGC and potential applications of BCs, consider a case of citizen science—a type of UGC that harnesses contributions of ordinary people for scientific research (Burgess et al., 2017; Kosmala et al., 2016; Levy & Germonprez, 2017). Many citizen science projects have three characteristics that are common in UGCs: purpose-driven information collection, project openness, and lean user profiles. We briefly discuss each of these characteristics below. In addition, we provide a list of existing UGC projects in Appendix A and categorize them based on these three characteristics.

- **Purpose-driven information collection:** Although many uses of UGC come from mining existing sources such as Twitter, organizations are rapidly developing specialized UGC platforms to harness its power. Examples include BeingGirl.com (by Procter & Gamble), eBird.org (by Cornell University), and FEMA Disaster Reporter App (by the US Federal Emergency Management Agency). This type of organization-sponsored UGC promises to deliver targeted and less noisy data that are better aligned with organizational information needs (Brynjolfsson et al., 2015; Deng, Joshi, & Galliers, 2016; Lukyanenko et al., 2017).

- **Project openness:** Many UGC platforms are completely open and invite anybody to join and participate. This results in a need to model a system that is (ideally) capable of collecting data from users with vast differences in domain views and expertise and varying levels of motivation to contribute information. Because of the prevalence of mobile devices, such users are often on the go and may be prone to short attention spans. Citizen science projects are one example that broadly upholds the ideal of open participation (Hand, 2010; Philippoff & Baumgartner, 2016).

- **Lean user profile.** Many UGC environments are anonymous or semianonymous. Citizen

science projects often deliberately avoid maintaining persistent user profiles to comply with ethics protocols and to avoid placing participation barriers for people that do not want to provide personal information (Burgess et al., 2017; Louv & Fitzpatrick, 2012). Likewise, many mobile UGC projects choose not to collect extensive profile data, as doing so is viewed as a usability barrier (Hosseinmardi et al., 2014; Van Kleek et al., 2011; Wiggins & He, 2016). This suggests that IS developers must assume that an IS has little information about each user and that it will be difficult to mine data to predefine data structures appropriate for every user.

Consider a high-profile example of the citizen science project iSpot (www.ispotnature.org) run by The Open University in the UK (Clow & Makriyannis, 2011; Scanlon, Woods, & Clow, 2014; Silvertown, 2010). The objective of iSpot is to expand scientific knowledge by asking people to observe plants, animals, and other taxa across the globe and report these sightings on their custom online platform. Selecting classes for an open, purpose-driven application such as iSpot can be challenging. The open nature of participation means there are no established guidelines for ensuring that all potential users can be engaged for requirements elicitation and analysis. Since iSpot is designed to deliver data for scientific research, analysts may elicit a list of species from scientists together with higher-level classes to group the species. Representing these structures using conventional conceptual modeling grammars (e.g., UML) may produce a model (i.e., script) akin to the modeling fragment presented in Figure 1.

As seen from the script, the focal data collection class on iSpot is the species level of classification (e.g., *brown bear, sugarbag bee, spotted sandpiper*) as these are standard units of conservation and measurement in biology (Crall et al., 2011; Mayden, 2002). Based on scripts similar to that of Figure 1, developers can then create database tables and user interfaces.

Existing research has focused on the problem of ensuring that distributed, nonexpert online users are able to report information using specific classes needed by data scientists (e.g., biological species). For example, research has been investigating innovative means for training online volunteers or collecting data without forcing users to identify objects at the species level (Kosmala et al., 2016; Lukyanenko et al., 2017). Figure 2 shows a sample online quiz on iSpot that trains online volunteers to identify species of interest to the project.
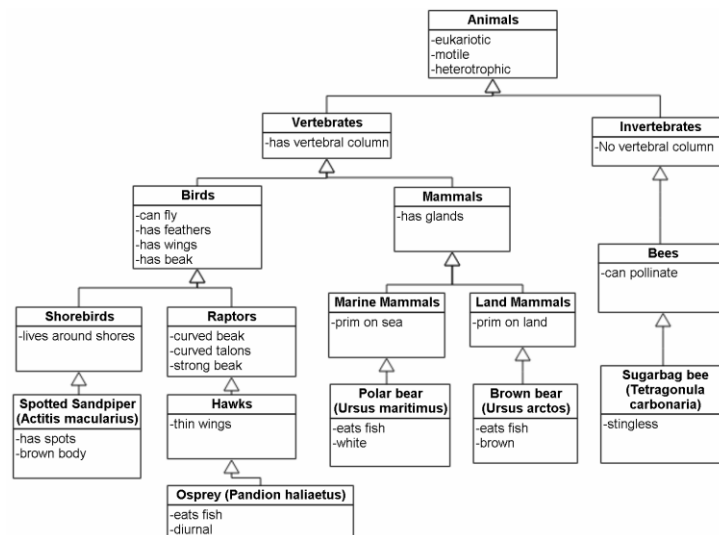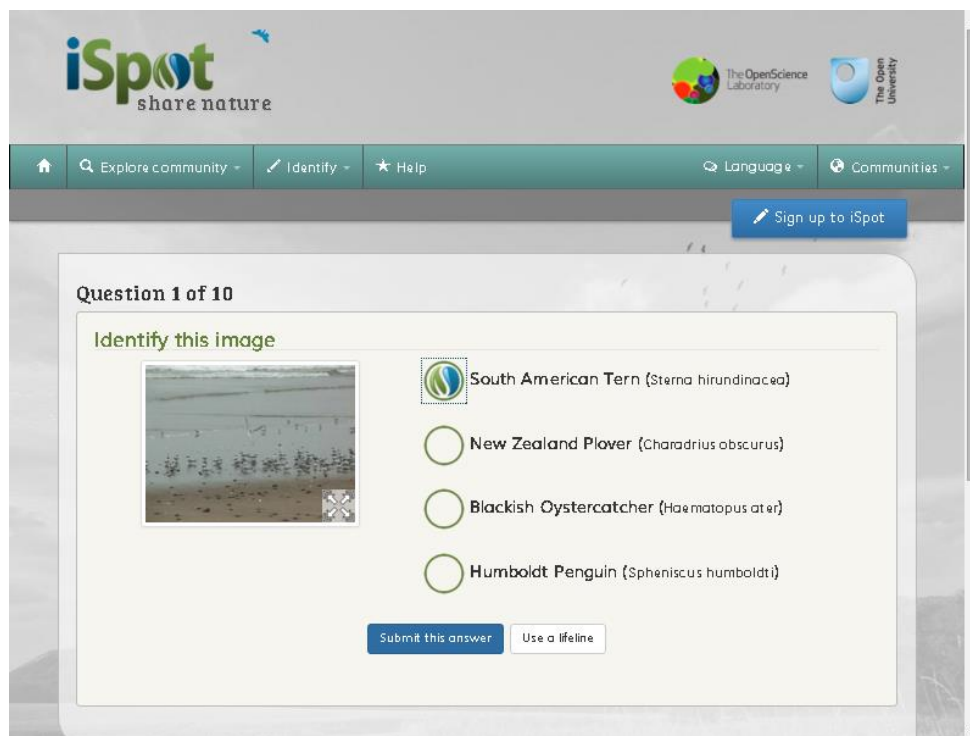
Figure 1. Fragment of a Candidate Script for iSpot

Figure 2. Online Quiz on iSpot that Trains Online Volunteers to Identify Species of Interest (https://www.ispotnature.org/quiz)

A major challenge that has received considerably less attention is organizing specialized categories (e.g., species in Figure 2) into higher-level classes. The list of specialized classes can be extremely long—researchers estimate that there are 8.7 million (±1.3 million SE) eukaryotic species globally that fall in the domain of iSpot (Mora et al., 2011). A natural way of handing specialized classes is by organizing data collection around more general classes so that navigational elements, menus, tutorials, and other interface choices can be presented and filtered by these more general classes.

The problem of effective organization of data collection and other design elements is becoming even more pressing for projects in which developers face severe space limitations and constraints—for example,

mobile or miniaturized environments (Melumad et al., 2019). Generally, this requires using classes at higher taxonomic levels (e.g., in Figure 1, these include raptor, bird, mammal, vertebrate, invertebrate, marine mammal, shore bird, land mammal, bee) but, in most domains, there can be hundreds of more general classes. Developers are provided with limited guidance on how to select the best generic classes.

Requirements elicitation and conceptual modeling activities are typically used to understand how application domains should be structured. Since not all potential users are involved in the development of the IS, these techniques may not work for UGC. It is possible that some nonexpert users may prefer (or be only familiar with) classes other than those presented in Figure 1. For instance, the fact that polar bears are bears and spend considerable amount of time on land may lead nonexperts to incorrectly (from the point of view of the project sponsors) conceptualize them as land mammals (Kaufman, 1999); likewise, some users may fail to classify spotted sandpipers as shorebirds. Furthermore, seeing ospreys near shores, nonexperts may consider them to be a type of shorebird—an incorrect classification from the point of view of scientific taxonomy and one incongruent with the script in Figure 1. Previous research has found that misalignments between the chosen classes and those preferred by users have an impact on data quality and effective system use (Burton-Jones & Volkoff, 2017; Lukyanenko et al., 2019). The success of a project might be threatened because some users may not be able to navigate structures of the project, contribute observations, or find desired information.

Prior research in conceptual modeling has shown that there are instances in which user agreement for classes may be very high (Lukyanenko et al., 2014b; McGinnes, 2011). Previous work has suggested that these conceptual modeling classes correspond to basic-level categories proposed in the reference discipline of psychology (Rosch et al. 1976). Recently, Lukyanenko et al. (2019) experimentally showed that when nonexpert observers were able to use basic-level categories to describe both familiar and unfamiliar objects, the resulting accuracy was nearly 100%. However, despite the strong empirical evidence of their effectiveness, no work in IS has deeply investigated the nature of these classes or, more importantly, how to identify and select these classes.

With this motivation, our research questions are: (1) *Which classes are the most appropriate to all potential users in a UGC project*? (2) *How can IS developers and researchers identify and select these classes?*

3   Existing Conceptual Modeling Approaches and UGC

We begin by considering prior research in IS related to our work. Existing studies in areas dealing with the diversity of users, difficulties in eliciting information requirements, and variable user needs lay the groundwork for our research. Research in conceptual modeling and related fields (e.g., human-computer interaction, computer-supported cooperative work, and social computing) has long recognized the value of selecting classes appropriate for the intended users' levels of domain familiarity and expertise (Sowa, 1995). However, it is commonly assumed that selecting these classes requires potential users to be directly involved in requirements elicitation and conceptual modeling (Dobing & Parsons, 2006; Erickson, Lyytinen, & Siau, 2005; Gemino & Wand, 2004), which is a prominent approach in the participatory design research tradition in IS (Björgvinsson, Ehn, & Hillgren, 2012; Bødker, 1996; Ehn, 1988; Kyng, 1995; Robertson & Simonsen, 2012).

Participatory design has been applied to development challenges in distributed, multiuser and open settings, including UGC (Gumm, 2006; Lukyanenko, 2016; Obendorf, Janneck, & Finck, 2009). Acknowledging the difficulties of developing IS in these settings, researchers have suggested leveraging innovative communication technologies and continuous system improvements and seeking frequent user feedback and the use of user surrogates (e.g., usability experts) as potential solutions (Anand & Mobasher, 2003; Le Dantec et al., 2015). Practical constraints in the real world have limited the efficacy of these solutions, as projects tend to effectively engage only a handful of prospective users (typically those most readily accessible to the development team) (Bratteteig & Wagner, 2014).

Researchers and developers have also conceptualized average users (also called "personas") whose attributes best represent the average attributes (e.g., personality, domain expertise, and model of the world) of the user population (Ehn, 1988; Iivari, 2011; Muller, Millen, & Strohecker, 2001). In practice, "politically representative users" are common (Muller et al., 2001, p. 102)—i.e., users who are delegates of established organizational units (e.g., trade unions, functional units, team leaders, managers) (Baskerville, de Marco, & Spagnoletti, 2013; Ehn, 1988; Kraft & Bansler, 1994). In general, no ideal solution has emerged and researchers have increasingly called for more work on adapting participatory design approaches to UGC settings (DiSalvo & DiSalvo, 2014; Lukyanenko, 2016).

Another emerging solution that has shown early promise is skipping conceptual modeling entirely and avoiding traditional domain representations (as shown in Figure 1). The resulting "lightweight" or "no conceptual modeling" approach then simply selects a flexible data model (e.g., a schemaless noSQL database) and presents users with an interface that allows them to suggest any attributes or classes they wish to report in a free-form manner (Jabbari Sabegh et al., 2017; Kaur & Rani, 2013). Yet, even when data collection is flexible, projects may wish to partition interfaces and navigational structures and provide tutorials organized by subjects, suggesting that even these types of projects may benefit from identifying effective higher-level classes.

Finally, a promising approach is to generate classes that match a user profile or varying user preferences "on demand." This work is particularly active in the human-computer interaction community (with the focus on interface design) but has also been pursued in conceptual modeling contexts (Ho, Davern, & Tam, 2008; Mobasher, Cooley, & Srivastava, 2000). Of particular relevance to our paper is TAXIS, a design language developed by Mylopoulos and colleagues (Mylopoulos & Wong, 1980; Nixon et al., 1987). TAXIS has the capability to detect user expertise, facilitating matching the most appropriate class to a user. These approaches are powerful and, in many situations, can be an adequate solution for constructing models (and corresponding user interfaces) that are appropriate to different users. However, for on-demand approaches such as TAXIS to be effective, the system needs to have a vast amount of prior data about a particular user (e.g., user domain views, expertise, abilities), which is challenging in cases where user participation is open and largely anonymous.

In general, dynamic on-demand solutions presuppose that the classes vary from one user to another. However, we also believe there is merit in keeping some conceptual structures unchanged and static. This may become useful when analyzing data for scientific purposes because drawing inferences from UGC projects may require that conditions under which observations were made by online crowds (including data collection interface choice) be as similar as possible. The static conceptual structure can also become valuable when building the static elements of a project—e.g., navigation menu, help items, process flow, etc.

Motivated by the limitations of existing approaches, we develop an alternative method that selects classes for which the interuser agreement is maximized before IS development. Prior conceptual modeling research suggests the existence of such classes following theoretical work in psychology on basic-level categories. Seeing promise in such categories, this research has called for practical guidelines to leverage

these categories as potential classes (Lukyanenko et al., 2014b; Lukyanenko & Samuel, 2017; McGinnes, 2011).

We heed the recommendation made by prior research and, in the next section, turn to psychology research in search of theoretical guidance. We then develop guidelines for the identification and application of basic-level categories as basic classes (BCs).

3.1 Theoretical Foundations for the Guidelines

The special status of basic-level categories is rooted in its position within a cognitive knowledge hierarchy. Before addressing relevant theories in psychology, we briefly consider the notion of levels or hierarchies of classes in conceptual modeling (Borgida, 1985; Purao & Storey, 1997; Smith & Smith, 1977). Conceptual models use the notion of hierarchy extensively with grammars such as the ER or UML, object-oriented database design, and supporting class hierarchy representations (e.g., via relationships of generalization/specialization, property inheritance) (Storey, 1993). Moreover, many scripts contain hierarchical structures (Dey, Storey, & Barron, 1999; Storey, 1993; Ullrich, Purao, & Storey, 2000; Wand, Storey, & Weber, 1999). Despite the centrality of hierarchies to scripts, the assumption was that there are no special classes since the classes are subject to someone's perceived reality, which reflects different user needs. In contrast, psychology research since the 1970s has begun to consider whether certain classes in a hierarchy have innately privileged standings.

According to psychology research, humans routinely form class hierarchies based on the need to maintain classes at different levels of abstraction since these levels perform fundamental functions of classification differently. Specifically, classes support *cognitive economy* and *inferential utility* (Lakoff, 1987; Roach et al., 1978; Smith & Medin, 1981)—two vital functions of organisms and one of the defining mechanisms of human cognition and behavior (Corter & Gluck, 1992; Roach et al., 1978). These functions compete for the same *limited* cognitive resources of human *memory, attention,* and *processing power.*

Cognitive economy is achieved by maximally abstracting from individual differences among objects and then grouping objects in classes of larger scope (Fodor, 1998; Murphy, 2004; Smith & Medin, 1981). In a biology domain, such classes could be *animals* and *plants*. By storing only a few classes, humans can easily memorize identifying characteristics of different class members (e.g., objects). Having only a few classes in the vocabulary maximizes the likelihood that two different people would have the same classes, which promotes communication efficiency and social interaction (Murphy, 2004). Cognitive economy

becomes increasingly vital because the environment continuously supplies organisms with massive amounts of unique sensory data. Thus, having fewer classes helps people cope with the changing diversity of the world. Strictly focusing on the benefit of cognitive economy therefore suggests that the best candidates for maximal agreement classes are those classes with the broadest scope—those at the top of the classification hierarchy.

Overemphasizing cognitive economy, however, comes at the expense of ignoring certain individual characteristics of objects that may be vital for the organism's function and survival via inductive inference (inferential utility). For example, suppose we are interested in a particular property of an object we encounter (e.g., we wish to discern if a mushroom is poisonous or edible). Classifying this object as a *fungus* (a high-level class) versus a *Clitocybe rivulosa* (a particular lower-level kind of poisonous mushroom) is associated with different probabilities of this object having the property of interest. The probability that a *Clitocybe rivulosa* is poisonous is substantially higher than the probability of any fungus being poisonous.[2] Thus, the ability to predict attributes of instances of a class, or the inferential power, increases as the scope of the class decreases. It follows that to maximize predictive power, humans should prefer classes with a narrower scope. While classes with a narrower scope are useful in many ways, memorizing, organizing, and communicating these categories require more cognitive and social effort.

Based on the tradeoff between cognitive economy and inferential utility, psychology research hypothesizes that humans favor (e.g., learn, communicate) those classes that maximally exploit both predictive power of classes and their cognitive economy. For example, Rosch et al. (1976) argued that in the world of "infinite number of discriminately different stimuli" and facing the tradeoff between cognitive economy and inferential power, humans favor classes that are most capable of supporting these competing objectives of classification. Based on converging evidence from anthropology and psychology (Berlin, Breedlove, & Raven, 1973; Raven, Berlin, & Breedlove, 1971; Rosch et al., 1976), Rosch et al. (1976) proposed that there is a set of "privileged" classes that they coin *basic-level categories,* which have become the subject of active research in psychology and cognitive sciences and have generated a considerable amount of

evidence, making the concept of basic-level categories one of the most established ideas in modern psychology (Lassaline, Wisniewski, & Medin, 1992; Murphy, 2004).

We review conclusions regarding basic-level categories generated by forty years of psychology research (Lassaline et al., 1992; Murphy, 2004). We organized these conclusions about basic-level categories into theoretical propositions that lay a foundation for their use in conceptual modeling (see Appendix B for references to specific papers supporting each proposition).

3.1.1 Theoretical Proposition 1: The Taxonomic Middle

As follows from the special function of basic-level categories of optimizing the tradeoff between cognitive economy and inferential utility, the basic level tends to be the taxonomic middle. Concepts that belong to this level tend to reside between the highest and lowest level in a conceptual hierarchy (e.g., *dog* is higher than *collie* and lower than *animal*). Basic-level categories tend to be common words such as *bird*, *tree*, *fish*, *cup*, *chair*, and *house* (Table 1 shows examples of basic-level categories identified by prior research) that occupy middle levels in the respective domain taxonomies.

3.1.2 Theoretical Proposition 2: Entry Category

Psychologists argue that a basic-level category is often an entry category—i.e., the first concept thought by a user when encountering a phenomenon (Jolicoeur et al., 1984). Murphy and Brownell (1985) called it the "necessary first step" of identification (p. 72). As entry categories, they tend to be retrieved from memory extremely quickly and accurately (Lukyanenko et al., 2014b; Zhou et al., 2010). In contrast, more precise and greater inference-bearing subcategories (e.g., dachshund) are contingent on expertise (e.g., dog experts may bypass the basic-level category and think of a specific breed but are still quite aware of the basic-level *dog*). An entry category may be different in situations when a phenomenon is an atypical representative of its basic class (e.g., subordinate *chicken* of the basic-class *bird*) (Murphy & Brownell, 1985).[3]

---

[2] This example also demonstrates why many UGC projects are interested in finer levels of classification (e.g., specific product categories, biological species). Knowing that a phenomenon is Clitocybe rivulosa affords greater inferences and action than knowing it is a fungus. This is the key reason why many biologically focused UGC projects would insist on collecting information at the species level, despite obvious difficulties this may create for amateur data

contributors (Kosmala et al., 2016; Lewandowski & Specht, 2015).

[3] This raises the question that there might be multiple basic-level categories (e.g., bird, duck; bird, chicken) within the same taxonomic tree. We contemplate this for future research opportunities.

Table 1. Some Examples of Basic-Level Categories from Psychology Studies

Basic-level category	Reference
Bird, dog	Tanaka & Taylor (1991)
Bear, rhino, pig, seal, bug, cat, turtle, crab, dog, fish, elephant, rabbit, horse, lizard, hippo, duck, snake, frog	Waxman & Klibanoff (2000)
Horse, rhino, lizard, pig, hippo, bug, duck, turtle, snake	Klibanoff & Waxman (2000)
Tree, fish, bird	Rosch et al. (1976)
Flower	Mervis et al. (1994)
Dog, duck, cat	Rhemtulla & Hall (2009)
Mouse, fish, butterfly, bird, rabbit, beetle, dolphin, horse, dog, tree, monkey, chicken	Op de Beeck & Wagemans (2001)
Apple, pear, orange, lime, coconut, pineapple, carrot, peas, corn, pepper, pumpkin, avocado, bird, dog	Jolicoeur et al. (1984)
Birds, dogs, fish, other common animals	Johnson & Mervis (1997)
Apple, melon, berry	Wales et al. (1983)
Horse, spider, chicken, fish, dog	Mandler & Bauer (1988)
Cat, dog, horse, bird, bat	Younger & Fearing (2000)
Bush, tree, flower	Murphy & Wisniewski (1989)
Cow, sheep	Zhou et al. (2010)
Cat, dog, horse, cow, apple, pear, daffodil, sunflower	Bowers & Jones (2008)
Dog, tree	Rorissa (2008)
Bird, flower, tree	Barr & Caplan (1987)

3.1.3 Theoretical Proposition 3: Frequently Used Words

Basic-level categories are words that occur most often in ordinary daily discourse, as communication is driven by the pragmatic need to exchange more information with the least effort. These findings originate in the work of Zipf (1935), who found that the length of a word is inversely related to its frequency of use (e.g., there is a small number of short words that are used frequently, while most long words are used less frequently). Shorter words tend to be the most frequently used words (Lassaline et al., 1992). Frequently used words indicate a balance of predictive power and cognitive economy and are thus uniquely suited for efficient communication.

3.1.4 Theoretical Proposition 4: Cohesion and Coupling

Compared to other levels, subcategories within basic-level categories are perceived as being the most similar to each other (Rhemtulla & Hall, 2009), i.e., having *cohesion*, while two neighboring basic-level categories have many psychologically relevant differences (Markman, 1991), i.e., exhibit *coupling*. In general, the basic level maximizes "both within-category similarity and between-category dissimilarity" (Mandler & Bauer, 1988, p. 247). Basic-level categories are generally the most differentiated (Murphy & Brownell, 1985). Thus, by knowing that a canary is a bird, we can confidently generalize to items with similar characteristics (e.g., other kind of birds) but not with items that are dissimilar (e.g., other kind of animals)

(Patterson, Nestor, & Rogers, 2007). Individual dogs are all represented using very similar patterns, whereas other kinds of animals (e.g., pigs, goats, birds) are represented using somewhat different patterns, and nonanimals are represented using dramatically different patterns.

3.1.5 Theoretical Proposition 5: Object Visualization

Basic-level categories are the most inclusive categories that allow for the construal of a visual gestalt (i.e., an organized whole that is perceived as greater than the sum of its parts), which is an image of a category schema compatible with most category members. For example, the outer shapes of most members of the category *dog* are so similar that it is possible to imagine a picture of a dog "as such." This is clearly impossible for superordinate categories (e.g., animal) because their members' outer shapes are too divergent (e.g., dog vs. bird). Considering psychological mechanisms of object visualization is especially important, as vision is perhaps the most important sensory organ for humans (O'Callaghan, 2017).

3.1.6 Theoretical Proposition 6: Simplest Words

Likely because of frequent use, words in the basic taxonomic level are generally morphologically simple (Craig, 1986). These words are known as primary lexemes (e.g., dog, home, food) (Brown, 1958; Rosch et al., 1976), whereas subordinate terms tend to be secondary lexemes that are formed from the basic level

term and a modifier (e.g., stray dog, family home, tasty food) (Berlin et al., 1973). While, in general, basic-level categories tend to be short, this proposition stresses the lexical complexity, thus explaining why some relatively long words such as lizard or elephant are also basic-level categories.

3.1.7 Theoretical Proposition 7: Original Words

Psychologists have further demonstrated that children learn basic-level categories first (Mervis et al., 1994). Mervis and Crisafi (1982) suggest that children's categorization ability is acquired in this order: basic, superordinate, and subordinate. This is partially based on the children's unsupervised way of discovering the world and partially driven by the influence of the adults. Adults have notions about the kind of language appropriate for use with children (e.g., long names are troublesome for children). Thus, adults do not necessarily provide a child with the name that is at the level of usual utility in the adult world (e.g., they might refer to an object as a *coin* rather than a *dime* since the monetary value of the coin is of little relevance to young children) (Brown, 1958).

3.1.8 Theoretical Proposition 8: General Predictive Utility

Inferences are one of the fundamental functions of categories, and the basic level disproportionally contributes to inferences. Rosch et al. (1976) hypothesized that because of their exceptional familiarity to humans and high frequency of usage, basic-level categories contain a large number of attributes that people think of when they think of a basic level (e.g., many attributes for birds—can fly, has feathers, lays eggs, builds nests, etc.—versus few additional attributes for shorebirds, for example). Corter and Gluck (1992) expanded Rosch et al.'s (1976) hypothesis by adding a base rate frequency of

categories (see Appendix C for more details). They reasoned that while a subordinate class, such as *chickadee* necessarily has more attributes (and inferences) than the basic *bird*, *chickadee* is used much less frequently than *bird*. This means that in the absence of much knowledge about an object, inferences to basic-level attributes are sound as a cognitive strategy. Thus, while basic inferences are cruder (e.g., lays eggs vs. lays blue eggs), they are more reliable in most daily situations.

To summarize, classification theory in psychology amasses considerable evidence for the existence of classes that maximize agreement among people with different backgrounds, education, and functional needs. So-called basic-level categories have been shown to carry a multitude of benefits resulting in a significant cognitive bias toward these categories. In the next section, we use and expand upon the theoretical foundations from basic-level categories to develop guidelines for identifying what we call *basic classes* (BCs) in conceptual modeling.

4 Guidelines for Identifying Basic Classes in Conceptual Modeling

A natural application of the theoretical propositions in psychology is to construct a set of design guidelines for conceptual modeling—as has been done in other design science studies (e.g., Evermann & Wand, 2005; Parsons & Wand, 2008; Soffer, Wand, & Kaner, 2015). As demonstrated above by the discussion of basic-level categories in psychology research, no claim will be definitively diagnostic for identifying BCs and there will be exceptions to the propositions. Some claims made about the basic-level categories lack operational precision (e.g., it may be unclear how to determine the middle among even levels in a taxonomy).

Table 2. Guidelines for Identifying Basic Classes of a Domain in Conceptual Modeling

Guideline name	Guideline description
G1: Middle Level	Identify classes in a domain in the middle of the conceptual hierarchy.
G2: Entry Category	Elicit entry classes from a sample of potential users for objects of interest.
G3: Frequently Used Words	Identify the most frequently used domain words used in a typical discourse.
G4: Cohesion and Coupling	Find a domain taxonomic level, for which sibling domain classes have maximal difference and their respective children have maximal similarity.
G5: Object Visualization	Find the highest class in the domain taxonomy for which class members can be easily visualized.
G6: Simplest Words	Among the classes in a domain, identify shortest and morphologically simple words.
G7: Original Words	Identify the first words or concepts in the domain learned by children or used by mothers to talk to children.
G8: General Predictive Utility	Identify classes in the domain with the greatest general predictive utility.

This difficulty in operationalizing theoretical ideas from reference disciplines into IS design principles is common (Arazy, Kumar, & Shapira, 2010; Hevner, 2007; Iivari, 2007). To overcome the lack of definitive guidance from psychology, we suggest using all eight propositions to intentionally introduce overlap and increase the likelihood of finding all relevant BCs in a domain. At the same time, we introduce additional precision in order to support a more consistent operationalization of the propositions (Chandra Kruse, Seidel, & Gregor, 2015; Chandra Kruse, Seidel, & Purao, 2016).

Paralleling the eight conclusions about basic-level categories in psychology research, we propose eight modeling guidelines (see Table 2) that an analyst (or agent) could follow to identify the basic classes (BCs). We illustrate the application of each guideline with at least one example to aid in their use (as done, for example, in Soffer et al., 2015).

4.1 Guideline 1: Middle Level

Knowledge about objects in the world can be organized hierarchically (de Beeck & Wagemans, 2001; Rosch et al., 1976). Indeed, the conceptual model in Figure 1 depicts classes that are organized in a hierarchy proceeding from more abstract (e.g., animals) to more specific (e.g., osprey). Psychology research predicts that the basic level should be in the middle of a taxonomy. Incorporating the notion of basic-level categories in the taxonomic middle leads to the following conceptual modeling guideline.

Guideline 1: Identify classes in a domain in the middle of the conceptual hierarchy.

To apply this guideline, analysts could arrange classes in a domain as a hierarchy (e.g., similar to the one in Figure 1) and select classes in the middle. Much human knowledge is already organized hierarchically, thus analysts could also leverage many existing repositories (e.g., research databases, wikis, books) to identify core concepts within a particular domain. This process could also be automated with an ontology as input to an algorithm that outputs classes in the taxonomic middle. For example, in the Catalogue of Life (www.catalogoflife.com), a comprehensive index of species containing information on names and relationships for over 1.6 million species, each object includes a taxonomic hierarchy whose range includes the most abstract (e.g., kingdom), the middle (e.g., class, order, family), and the most specific (e.g., genus, species, and subspecies).

Psychology research does not offer precise guidance on determining which classes should be selected when the hierarchy is deep (e.g., contains more than three levels). It is also unclear how to select the middle class when the number of levels is even. Hence, to ensure

consistent application of the guideline, we refine Guideline 1 by introducing the following heuristics.

Heuristic 1.1: Select the class in the middle of the hierarchy when the number of taxonomic levels is greater than or equal to three and odd.

Heuristic 1.2: Select the two classes in the middle of the hierarchy when the number of taxonomic levels is greater than two and even.

For example, we propose that if the number of classes in the vertical axis of the hierarchy is odd and greater or equal to three—e.g., animal, bird, and osprey ($n = 3$)—the basic class would be that of the taxonomic middle, in this example *bird*. Similarly, if the number of classes in the vertical axis of the taxonomy is even and greater than two—e.g., animal, vertebrate, bird, osprey ($n = 4$)—the middle two classes should be chosen, both *vertebrate* and *bird*. We realize the number of classes may become unwieldy following this guideline alone; however, this conservative practice prevents prematurely eliminating BCs as candidate classes, which will be further refined through subsequent guidelines.

4.2 Guideline 2: Entry Category

Basic-level categories often become the first concepts (entry category) thought of by a user when encountering a phenomenon. As discussed above, entry-level effects are contingent on a user's domain expertise and the typicality of the exemplar. However, even experts readily relate to the basic-level categories (in contrast to lower levels that require familiarity and expertise) (Tanaka & Taylor, 1991). In addition, expertise rarely spans an entire domain. For example, a person who owns a collie might be considered a "collie expert," but not an expert in other dog breeds (Tanaka & Taylor, 1991). Thus, entry-level effects offer a strong diagnostic for identifying BCs, leading to the next guideline:

Guideline 2: Elicit entry classes from a sample of potential users for the domain objects of interest.

A natural way to apply Guideline 2 is to elicit entry classes from potential users. In doing so, analysts should be aware that some responses might not be basic (because of the confounding effects of typicality and expertise). We suggest retaining all responses regardless of their perceived BC status and refinement among candidate BCs can occur after all the guidelines are considered together. The entry category also considers the effectiveness of the BC as way of organizing information (e.g., on a user interface). Guideline 2 provides a mechanism to elicit relevant classes from users, including nonexpert users. The entry categories can be easily crowdsourced online, especially with the use of crowdsourcing platforms, such as Amazon's Mechanical Turk or CrowdFlower

(Deng et al., 2016; Ipeirotis, Provost, & Wang, 2010) and further refined by the analyst. Figure 3 shows a prototype interface we built to elicit a nonexpert's (e.g., a citizen who lives in the geographical region of interest) classes of instances observed.

To determine a stopping point, we suggest applying stopping rules suggested in previous conceptual modeling research (Browne, Pitts, & Wetherbe, 2007; Browne & Ramesh, 2002). Stopping rules describe how individuals make a judgment of sufficiency when searching for information in order to move to the next stage in a problem-solving or decision-making process. Generally, a person will invoke a heuristic, or a stopping rule (Browne & Pitts, 2004b; Browne et al., 2007; Nickles, Curley, & Benson, 1995). Examples of some of the stopping rules that apply are outlined in Table 3 (based on Browne & Pitts, 2004a).

4.3 Guideline 3. Frequently Used Words

BCs are typically the most frequently used words in a language, making word-use frequency a convenient diagnostic feature of the basic level.

Guideline 3: Identify the most frequently used domain words used in a typical discourse**.**

A natural application of this guideline would involve mining existing sources of data and selecting the most frequently used words in a domain based on a threshold. For the citizen science example in Figure 1, for example, one could parse information from scientific publications, biology ontologies, or UGC (e.g., social media sources, such as Twitter) to identify common words that may suggest potential basic classes.

Figure 3. Prototype Interface for Eliciting Entry-Level Categories

Table 3. Stopping Rules in Support of Guideline 2

Stopping rule	Application
Difference threshold	Using the difference threshold stopping rule, developers assess the marginal value of the latest piece of information acquired (Nickles et al., 1995). Developers then stop eliciting entry-level classes when they determine they are no longer learning anything new.
Mental list	Developers have a mental list of items that must be satisfied before they stop collecting information. For example, a developer's mental list could include a minimal number of classes, types of classes, category of classes that must be covered. As each entry-level class is obtained, arguments are made for or against using each piece of information to fulfill requirements on his or her mental list. Once the developer reasons that all of the items contained on the list or set have been attained, the gathering of additional entry-level class ceases.
Representational stability	The developer elicits information until her or his mental model stops shifting and stabilizes, with the focus being the stability of the representation. When a new entry-level class is obtained, the developer either decides that this new class supports the use of this class to modify the representation or rejects the use of the new entry-level class. When the developer's mental representation of the problem is no longer being developed, he or she ceases collecting additional entry-level classes.

Psychology research does not offer guidance on the frequency threshold when identifying the most frequently used words. Following previous research investigating distribution frequencies, we suggest identifying an inflection point—a point that separates exponential and uniform distribution in a frequency plot (Friedman, 1967; Reynolds, Scott, & Nussbaum, 1980). To ensure consistency in application, we refine Guideline 3 by introducing the following heuristic.

Heuristic 3.1: Identify the most frequently used domain words used in a typical discourse by using the inflection point in the distribution frequencies as a frequency threshold.

To illustrate a possible implementation for Guideline 3, we used basic text mining techniques to parse an electronic document in the biology domain, which could be helpful in designing an application to collect UGC (Feldman & Dagan, 1995; Vequist & Licht, 2013; Weller, 1999)—specifically, visitors' sightings of wildlife in a national park. Toward this end, we mined a relevant source, the popular book *Wildlife Watching in America's National Parks* (Vequist & Licht, 2013), employing standard natural language processing techniques to identify the most frequently used words (Kao & Poteet, 2007). The general process involved transforming words into lower case, tokenizing (i.e., extruding basic linguistic units such as

words, punctuation, and numbers), filtering stop words (i.e., common words that do not add value to the analysis such as *the, as, for*), and counting the term occurrences in each of the documents. Figure 4 shows the outcome of this process—a list of the most common words ranked by their frequency of use. As expected, a handful of words (candidate BCs) appeared more often than others (e.g., 10% of distinct nonstop words account for 50% of all words in the corpus).

Then, we applied Heuristic 3.1 to identify which words were used most frequently by finding an inflection point. We used common statistical software to obtain the equation for the frequency distribution and solved the equation for the inflection point, resulting in a cut-off at 204 words. Our process allowed us to reduce the corpus of almost 10,000 words to 204 candidate BCs. Examining this list more closely reveals numerous categories that psychology research (surveyed in Table 1 above) previously identified as basic (e.g., *park, turtle, desert, north, bird, photo, beach, nest, night, bat, water, bear*). It is also entirely possible that some of the words among the 204 retained are BCs not previously addressed in psychology research. This demonstrates a robust potential of Guideline 3 operationalized through the Heuristic 3.1 to uncover relevant BCs for the domain, including the potential to discover new BCs.
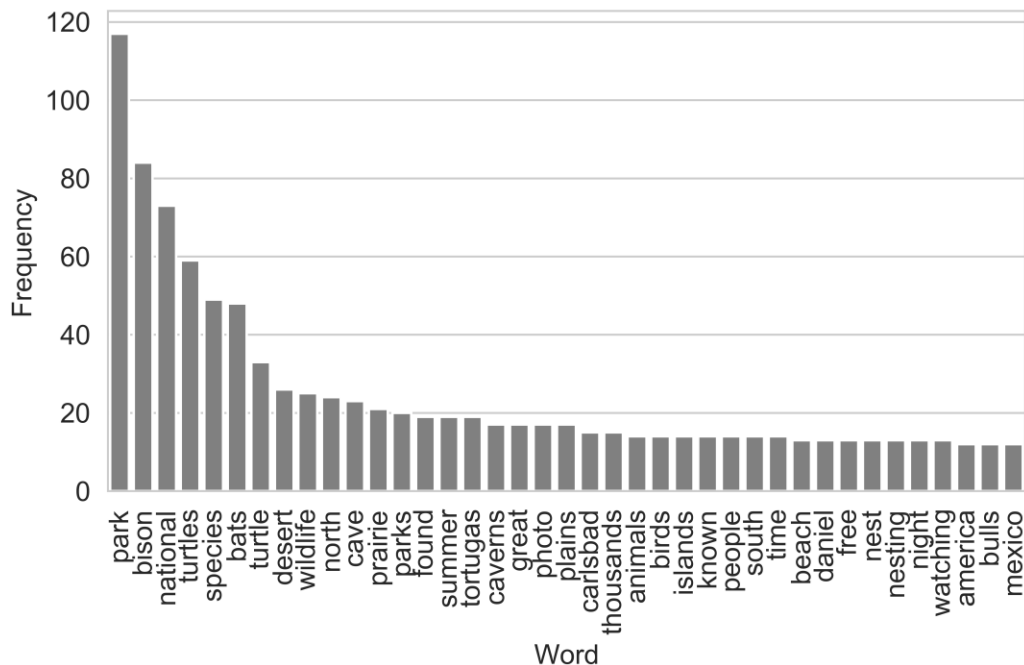

Figure 4. The Most Frequently Used Words in the *Wildlife Watching in America's National Parks* Corpus (Vequist & Licht, 2013).

4.4 Guideline 4: Cohesion and Coupling

Psychologists often suggest that basic-level categories carve the world at its natural joints (Gangestad & Snyder, 1985). Likewise, we expect BCs to contain members that are highly similar to one another, and highly dissimilar to members of other BCs.

Guideline 4: Find a domain taxonomic level, for which sibling domain classes have maximal difference and their respective children have maximal similarity.

To apply this guideline, analysts could interview prospective users of a system (e.g., sample of citizens) and ask them which classes in a domain are most dissimilar from one another. Alternatively, this guideline could be applied computationally if relevant data are available for mining and analysis. A common technique that could be leveraged here is cluster analysis, whereby a list of known attributes of objects of interest could be clustered using unsupervised learning techniques (e.g., *k*-means) to discover potential BCs. The resulting clusters could then be shown to domain experts (or prospective users) to validate and label the clusters.

4.5 Guideline 5: Object Visualization

Much of human experience is shaped by visual signals. Basic-level categories are particularly easy to visualize, leading to the following guideline.

Guideline 5: Find the highest class in the domain taxonomy for which class members can be easily visualized.

Following the citizen science taxonomy from Figure 1, an analyst could list the classes at the bottom of the hierarchy and ask users to identify a single visual object that represents that class. For instance, in Figure 1, the classes at the bottom of the hierarchy would be *spotted sandpiper, osprey, polar bear, brown bear, and sugarbag bee*. The task for the user would be to identify the highest class in the taxonomy for which class members could be visualized uniquely from other classes. For example, for a polar bear, the highest class the user may think of is *bear*. If the user chose a more abstract class (e.g., animal) it would be difficult to derive a shared visual image for all members of the class because *animal* also refers to birds, snakes, and bees, for example, which are very different from each other visually.

4.6 Guideline 6: Simplest Words

Basic-level categories are commonly short and morphologically simple words, offering a convenient diagnostic for BCs.

Guideline 6: Among the classes in a domain, identify the shortest and morphologically simple words.

To implement these guidelines, analysts could leverage text mining techniques to parse a domain-specific corpus and derive candidate BCs. The process is similar to the one followed to derive the most frequently used words (see Guideline 3) but we add an additional constraint to retain only tokens that are morphologically simple (i.e., single words such as *chair* rather than *bachelor's chair*) and short (e.g., low number of characters) before using an inflection point cut-off. As word length varies by language, analysts might consider the average length of words in a language (which can be obtained by computing an average on the words from a dictionary for the language of interest) and select words below the average. For example, in the English language, the average number of letters in words is between four and five (Welsh, 1988). Setting five letters as a threshold we would retain words—such as *bird, tree, fish, snake, home, dog, food, shop, new, old*—that psychology research identifies as basic-level categories based on the same data used in Guideline 3.

4.7 Guideline 7: Original Words

Reflecting the special psychological status of basic-level categories, these words are typically the first to be learned by children. When possible, this could be used as a diagnostic feature for BCs:

Guideline 7: Identify the first words or concepts in the domain learned by children or used by mothers to talk to children.

This guideline could be applied by interviewing children, parents with children, or mining existing sources. Following the citizen science taxonomy in Figure 1, analysts could parse the content from children's books relevant to their project and perform statistical analysis (e.g., term frequency-inverse document frequency, latent semantic analysis) to identify common words and or concepts and build a dictionary of these words used in children's books. An inflection point threshold (see Heuristic 3.1) could be applied to narrow the list to the most common words.

4.8 Guideline 8: General Predictive Utility

Often people reason about objects in the world not in terms of classes, but rather using attributes. This is a preferred strategy when reasoning and communicating about unknown objects or objects that may be difficult to definitively classify (e.g., due to an obscured image or when learning new objects) (Lukyanenko et al., 2017).

The world is not an unstructured total set of equiprobable co-occurring attributes. Rather, objects in the world are perceived as possessing high correlational structure (e.g., wings co-occur with feathers more than with fur). Words in basic-level

categories contain many learned associations (e.g., builds nests and lays eggs), which, combined with the higher use-frequency of words in basic-level categories, results in a unique advantage of having the most generally predictive attributes (Corter & Gluck, 1992). This property of basic-level categories may be used to identify BCs based on attributes.

Guideline 8: Identify classes in the domain with the greatest general predictive utility.

While this guideline may be implemented by asking stakeholders to reflect on the classes and attributes that are most predictive in general situations, the category utility (CU) proposed by Corter and Gluck provides an established quantitative procedure for identifying classes with the most general predictive utility (see Appendix C for more details). This is applicable in cases where domain information can be mined from existing data sources and thus may serve as a supplement or substitute for traditional requirements elicitation from individuals.

To demonstrate application of the CU function, consider the iSpot example above and the hierarchy animal—bird—osprey depicted in Figure 1 (for simplicity we ignore other classes). Assume the corresponding hypothetical feature probabilities (attributes) are those given in Table 4. Computing these probabilities for each class gives the CU measures shown in Table 5. Based on these calculations, *bird* has the greatest CU coefficient. According to Corter and Gluck, this result is explained by the relative balance between the use-frequency of the class *bird* and its predictive power relative to other classes.

4.9 General Considerations when Applying the Guidelines

Having discussed specific ways to apply each guideline, we consider a general strategy for implementing them in a project. Based on the lack of consensus in psychology research, no procedure can be definitive in identifying BCs. We therefore recommend attempting to *apply every guideline* and leveraging any overlap between guidelines *before determining the set of BCs*. Rather than viewing these guidelines as necessary and sufficient, we consider them as cumulative evidence supporting a hypothesis for a particular BC. For example, Guideline 3 (G3) provides a list of frequently used words for a particular domain (e.g., *animal, dog, cat, collie, snowshoe, siamese*). G1 represents a subset of classes that are in the middle of the hierarchy (e.g., *dog, cat*). Here, G1 is a subset of G3. Once these guidelines are followed, analysts should generate a list of candidate BCs. When guidelines create overlap in identified classes, analysts should select all classes generated by the guidelines, if the goal is to have a comprehensive list of BCs, or retain only the classes that are identified by most or all guidelines, if they wish to extract the most universal of the BCs.

To illustrate the application of the guidelines in a given context, we consider the scenario of building a hypothetical Smart City app that collects sightings of animals that are seen in urban settings. People are increasingly living in urban areas; thus, the development of the Smart Cities app could help urban planners, managers and decision makers collect a range of environmental and human-use data related to urban life (Maccani, Donnellan, & Helfert, 2014; Purao, Seng, & Wu, 2013; Ranchordás, 2018; Sinkonde et al., 2018).

Table 4. Feature Probabilities to Illustrate the Corter & Gluck Model

	Base-rate	P(k | animal)	P(k | bird)	P(k | osprey)
Motile	0.9	1	1	1
Can fly	0.4	0.5	0.95	1
Eats fish	0.006	0.007	0.01	0.9

Table 5. Category Probabilities and CU Measures to Illustrate Corter & Gluck Model

Class	Animal	Bird	Osprey
Probability of category, $P(c)$	0.9	0.33	0.005
CU measure*	0.25	0.31	0.01

Table 6. Sample Application of the Guidelines for a Hypothetical Smart City Project

Guideline name	Source of classes	Application of guideline	Potential resulting classes
G1: Middle Level	Structured analysis of a domain ontology	Using an existing ontology (e.g., Catalogue of Life) and looking for classes in the taxonomic middle by applying Heuristics 1.1 and 1.2 for odd and even numbers of classes, respectively.	Cat, dog, tree, fish, rabbit, hare, snake
G2: Entry Category	Interview with users	Design a prototype interface similar to the one in Figure 3 to elicit classes from potential users (both experts and nonexperts) on objects observed.	Cat, dog, bulldog, chicken, bunny, German shepherd, lab (Labrador retriever), fish, goldfish
G3: Frequently Used Words	Text-mining of relevant domain documentation	Retrieve a corpus of documents related to the application domain (e.g., Florida Wetlands). Parse the document and derive a frequency plot of the most frequently used words (e.g., words in the short tail of the distribution), find an inflection point to select potential basic classes.	Cat, dog, tree, German shepherd, snake, lab (Labrador retriever), fish
G4: Cohesion and Coupling	Structured analysis of a domain ontology	Calculate the total possible combinations and ask users (e.g., Amazon Mechanical Turk users) which classes are most dissimilar.	Tree, fish, snake
G5: Object Visualization	Interview with users	Ask potential users to identify the highest category in the taxonomy for which category members can be visualized uniquely from other classes.	Canine, feline, rodent, snake, tree, fish
G6: Simplest Words	Text-mining of relevant domain documentation	Follow-up analysis of the data used for G1 and identify morphologically simple and short words specific to the domain.	Cat, dog, tree, fish, rabbit, hare, snake
G7: Original Words	Text-mining of relevant children books	Use children books as the corpus look for words that are typically used by children.	Cat, kitty, dog, doggy, bunny, fish, fishy, tree, snake
G8: General Predictive Utility	Structured analysis of a domain ontology	Calculate and select classes with the highest CU coefficient.	Tree, snake, fish

Sensors such as air, noise, water monitoring devices, and traffic counters can be deployed to gather data in urban spaces. However, the Smart City app would also benefit from a different kind of sensor: the human sensor (Goodchild, 2007). Human sensors have advantages over other types of sensors in that they have the capacity to interpret real-world events and act upon them, thus making sense of unanticipated phenomena that would get coded as "errors" or "outliers" by most electronic sensors. As cities bring together people with different backgrounds, points of view, and perspectives, BCs could be used to create data collection interfaces, process flows, menus, navigational tools, and tutorials to make smart seeking apps accessible to as many people as possible.

Table 6 provides the details and outcomes of following our guidelines in the hypothetical Smart City context. As shown in Table 6, in this example, only *tree, snake,* and *fish* are selected by all eight guidelines, making them the maximally universal classes and excellent candidates for major navigation elements and other highly visible and used project features. In contrast, since other classes are selected by some but not all guidelines, this suggests that individuals can readily relate to the classes but that they may have certain limitations. To better understand why such classes could have limitations, one might consider why certain guidelines do not identify these classes. For example, *cat* and *dog* are present in five of the eight guidelines but are absent from G4, G5, and G8—i.e., the guidelines that deal with visual uniqueness—suggesting that cats resemble dogs much more than snakes resemble trees or fish. This does not necessarily disqualify these classes from being BCs but may indicate that more caution should be taken when selecting them as BCs. Indeed, dogs and cats, while although often displaying different behavior, share many morphological and relational features (e.g., have tails, fur, four legs and two ears, snouts, live with or close to humans). This means that, in some situations, discriminating between them may not be as easy as discriminating between the BCs selected by all guidelines (i.e., snake vs. tree). This may present an issue for certain projects. This information could also be used to interpret data generated through such a Smart City app (e.g., dogs may be mistaken for cats at a distance, thus observations of cats and dogs may not be as reliable during poor visibility conditions).

It is also possible to automate some or, in extreme cases, all of the guidelines. Indeed, with the ongoing digitization of human experience (reflected in big data phenomena), more and more of human knowledge is becoming accessible to computer-based analysis (Lazer, Brewer, Christakis, Fowler, & King, 2009; Maass, Parsons, Purao, Storey, & Woo, 2018). Guidelines G1, G3, G4, G6, G7, and G8 are especially conducive to automation, as they rely on an existing corpus of data. In contrast, Guidelines G2 and G5 are most naturally pursued by interviewing potential user, and thus may be more challenging to automate. Maas et al. (2018) suggest that domain expertise can help refine data obtained through data-driven approaches. Likewise, in our context, experts (e.g., domain experts, regular users, designers) could review the BCs produced by the guidelines (many of which may be obtained in a data-driven manner, such as through text mining), rank the classes, and select the best candidates based on the needs of the application.

In the end, following these guidelines should result in a list of effective BCs for the Smart City app. These classes could be used to develop menu items, label major sections of the project, organize and design data collection processes (e.g., users may be presented with the list of BCs and asked to select from it to report what they have observed), and even inform promotional material about the project. The universality of the BCs generated by following the guidelines should make the design features and processes informed by these classes more readily accessible to large audiences, expert and nonexpert alike, which would support wider participation and broader engagement with projects, facilitate more faithful communication of information, and, ultimately, contribute to the success of the processes and applications designed using the guidelines.

5 Evaluation of the Guidelines via Focus Groups

In this section, we evaluate the utility of our proposed guidelines and assess the usefulness of the guidelines in identifying appropriate domain structures familiar to users, regardless of the diversity of their backgrounds, knowledge, and domain expertise. We chose to evaluate our guidelines by engaging with analysts, developers, and other practitioners who would potentially benefit from using BCs in the design and implementation of an IS interface (Myers & Newman, 2007).

We chose to use a focus group methodology for our evaluation for a number of reasons. First, our research is the first to propose the notion of BCs for conceptual modeling in UGC contexts. Given the preliminary development of BCs, it was imperative to richly explore their utility to help us identify where there

might be a need for more development or clarification regarding BC use—something that might be missed through other forms of evaluation (Mazza & Berre, 2007; Prat, Comyn-Wattiau, & Akoka, 2015; Samuel, Khatri, & Ramesh, 2018; Tremblay, Hevner, & Berndt, 2010). Our focus groups allowed us to obtain feedback on our BCs using participants' natural ways of expressing themselves without restriction and offered us the opportunity to follow-up with them and ask questions to further clarify our understanding of any issues. As noted by van Aken, Chandrasekaran, and Halman (2016), focus groups "can be very informative and lead to better and more relevant management implications" since they facilitate direct interaction with participants.

Second, focus groups allow researchers to gain perspectives on a topic from a set of individuals interacting with a moderator and each other. In our case, this allowed us to glean perspectives of BCs from several individuals simultaneously, based on their understanding of the guidelines as well as any novel/nuanced perspectives that emerged during interactions with the moderator and each other. Interaction is a key strength of focus groups, as it provides the opportunity to receive feedback that might not surface with other evaluation strategies such as one-on-one interviews, surveys, or lab experiments (Krueger & Casey, 2000). Third, focus groups have been used in prior IS research to design and evaluate the utility and relevance of design artifacts (Ploesser, 2013; Prat et al., 2015; Samuel et al., 2018; Stahl, Tremblay, & LeRouge, 2011; Tremblay et al., 2010; Tremblay, Hevner, & Berndt, 2012). Thus, we determined that using focus groups was a viable evaluation strategy in the development of our guidelines.

5.1 Focus Group Design

We followed the approach outlined by Tremblay et al. (2010) in the design of our focus groups. In order to appropriately design the focus group and identify qualified participants, we defined the goals of the focus group as follows: (1) *Introduce participants to BLCs and BCs in the application of the guidelines*. (2) *Evaluate the utility of the BCs obtained from the eight guidelines and discuss if BCs would improve the design of user interfaces in a UGC application*

Per the advice of Tremblay et al. (2010), we ran multiple focus group sessions (hereafter we refer to the multiple sessions as focus groups) with different individuals of various backgrounds to mitigate potential bias in our findings. To ensure consistency in our focus groups, we created a moderator protocol with the planned procedures (see Appendix D) and detailed the tasks performed by participants (see Appendix E). Each focus group entailed a welcome, description of the procedures that would take place, an introduction

to BLCs and BCs, an initial impressions task to provide participants a chance to become more familiar with our BC guidelines, a design task in which participants could use BCs and our guidelines if they deemed them useful, and time to allow participants to offer any final feedback, thoughts, or comments on the guidelines. We utilized think-aloud techniques to maximize interaction with participants (Cotton & Gresty, 2006; Newell & Simon, 1972; Nielsen, Clemmensen, & Yssing, 2002; Stewart & Shamdasani, 2014).

In the following sections, we describe the focus group setting, the training exercise that introduced the guidelines to participants, the task in which participants used the guidelines to derive a user interface for a wildlife citizen science application, and the focus group results.

5.1.1 Focus Group Setting

The focus groups took place in a conference room at a large US urban university. The conference room was arranged in a U-shape to encourage collaboration between the participants (Krueger & Casey, 2000) and allow participants to easily see material used by the moderator (e.g., presentation slides to demonstrate the guidelines, whiteboard to document participant ideas, etc.). The conference room was also equipped with audio recording capabilities for later data analysis.

5.1.2 Participants

Table 7 describes the demographics of our participants across the focus groups. These participants were recruited because they had industry experience or formal training in systems analysis and design and conceptual modeling and were either alumni or current

graduate students in an intensive master's degree program in analytics at a large US university. There was no compensation beyond refreshments during the session, which lasted for 1.5 hours. As Table 7 indicates, our participants represented different age groups, females and males were equally represented, and, in general, they were experienced professionals that could comment on the usability of the BC guidelines in practice, as their roles comprised typical analysts and application developers at both senior and junior levels. Hence, we deemed this sample suitable for our focus group goals.

5.1.3 Focus Group Tasks

After introducing BLCs, BCs, and our guidelines, the first task was an initial impressions task that allowed participants to become more familiar with BCs and our guidelines by discussing how they impact user interfaces and underlying class structures. We identified two existing web applications to aid the discussion: WebMD and Mayo Clinic's online symptom-checker (see Appendix E). Despite the same goal of these two existing web applications (e.g., helping a user narrow down potential diagnoses based on their symptoms), there are key differentiators between the user interface designs that allowed us to discuss BLCs, BCs, and our guidelines. For example, a key differentiator between the two web applications is the application's entry category (i.e., Guideline 2). While the Mayo Clinic symptom checker organizes the information based on whether the individual is an adult or a child the WebMD symptom checker organizes information based on the body part where the symptom is located.

Table 7. Demographics of Focus Group Participants

IT experience	Age	Sex	Job title
Focus Group 1			
Less than 5 years	35-44	F	Business intelligence consultant
10-14 years	25-34	M	Clinical business technical consultant
15-19 years	45-54	M	Associate fellow
10-14 years	35-44	F	IT manager
10-14 years	35-44	F	IT project manager, information education, IRB
10-14 years	35-44	F	Senior quality assurance analyst
Focus Group 2			
Less than 5 years	25-34	F	Senior software developer
Less than 5 years	25-34	M	Sales associate/systems analyst
Less than 5 years	25-34	M	Test coordinator
Zero	25-34	F	Graduate student in analytics
Focus Group 3			
15-19 years	45-54	F	Business intelligence manager
Zero	20-24	M	Graduate student in analytics
Less than 5 years	20-24	M	Systems engineer
20-29 years	45-54	M	Systems analyst

We asked participants to draw parallels from each of the proposed BC guidelines to useful/not-useful design aspects of the symptom-checker applications (Question 1 in Appendix E). We then asked participants to discuss whether the BC guidelines could have been useful in the process of deriving classes for the conceptual model script supporting the design of either of these applications (Question 2 in Appendix E). This last discussion also served as a warm-up activity to help participants become more comfortable applying the guidelines to the design task.

5.2 Focus Group Design Task

The initial impressions task served to introduce participants to BCs and our BC guidelines and expose them to connecting them to conceptual modeling and interface design. Thus, at this point, the participants were ready to evaluate the usefulness of BCs and our BC guidelines for designing a new UGC application (our research context). The moderator distributed a printed version of the guidelines to each participant and explained that the goal was to design a conceptual structure for a UGC application that would allow citizen scientists to report wildlife encounters in the Everglades National Park in Florida. To get a sense of what the users of such an app might try to classify, we showed participants a slideshow with images of common wildlife in the Everglades (see Figure E3 in Appendix E). We then asked participants to apply the guidelines when designing this application (see Question 3 in Appendix E). The moderator captured ideas on a whiteboard and asked participants to discuss the utility of the guidelines and articulate how the guidelines might provide guidance in the approach they would use to design such an application—in particular, which guideline(s) they would apply. Finally, the moderator discussed the guidelines' utility and the implications of using them (e.g., quality of data, familiarity) (see Question 4 in Appendix E).

5.3 Data Analysis Approach

Each focus group audio recording was professionally transcribed for subsequent analysis. We conducted the analysis using Dedoose version 7.6.17 (www.dedoose.com), a popular qualitative research software (Silver & Lewins, 2014). Several approaches are available for analyzing qualitative data, including grounded theory (Corbin & Strauss, 1990) and interpretive phenomenological analysis (Smith, 1996). For our study, we used a template analysis (King, 1998, 2004), which has fewer defined procedures, compared to more formal alternatives, and is adaptable to our requirements. We created an initial template using our BC guidelines as the higher-order codes. These higher-order codes indexed sections of text as relating to a theme or issue in the data, which the

researcher had identified as important to his or her interpretation (King, 2004). We developed a coding scheme based on the guidelines to identify discussions, reactions, comments, or criticisms for each guideline.

The coding was completed in two rounds. Initially, one of the authors and a graduate student (MS student in business analytics with experience in systems design) independently coded the transcripts and used the guidelines as labels for the excerpts. The two coders systematically worked through a portion (30%) of the focus groups' transcripts in order to identify sections of the transcripts that were relevant to our aim of evaluating the utility of the eight design guidelines. Initially, any given excerpt could have multiple codes attached to it. The two coders discussed the areas of initial disagreement to reconcile differences in coding interpretation. The rest of the transcripts were then coded based on the agreement between the two coders (Tremblay et al., 2010). A pooled Cohen's kappa (Cohen, 1960) interrater agreement of 0.64 was achieved in the first round, which reflects good agreement between the two coders (Miles & Huberman, 1994). After discussing the areas of initial disagreement and completing the coding of all the transcripts, an interrater agreement of 0.88 was achieved, which reflects an excellent agreement between both coders (Miles & Huberman, 1994).

In the next section, we discuss the results generated from both tasks. We provide a summary table (Appendix Table F1) that evidences the utility of each guideline and addresses potential challenges that analysts may face when applying such guidelines, using insights from participants as support.

5.4 Focus Groups Results

Overall, our analysis of the data offered evidence for the utility of the guidelines. Table 8 illustrates the coding support for the guidelines across the three focus group sessions. Table 8 shows that Guidelines G1 (middle level) and G2 (entry category) were the most used across the different focus groups. One plausible explanation for this is that these guidelines are intuitive and require less information processing (e.g., calculate a frequency, compare to other guidelines, assess their cognitive utility, or realize whether the classes are morphologically short). Participants had difficulties with G5 (object visualization), likely because they were thinking of visual rather than prototypical images. Our initial impressions task pinpointed the utility of G1, middle level. A participant from Focus Group 1 (FG1) suggested that using middle-level categories would allow both nonexpert and expert users to contribute. A participant in FG2 reasoned that neglecting middle-level classes could lead to poor design choices due to discrepancies in data entry from nonexpert users.

Table 8. Code Application in Each of the Focus Groups

Code	Focus Group 1	Focus Group 2	Focus Group 3	Total
G1: Middle Level	7	9	2	18
G2: Entry Category	6	6	10	22
G3: Frequently Used words	1	2	2	5
G4: Cohesion and Coupling	4	5	2	11
G5: Object Visualization	3	2	1	6
G6: Simplest Words	--	5	6	11
G7: Original Words	3	4	4	11
G8: General Predictive Utility	7	2	3	12
Totals	31	35	30	96

Participants understood the importance of the generality of BCs when organizing information for a broader audience using different entry categories (G2). Participants stated that entry points are fundamental to help differentiate and help reduce redundant information. One interesting finding is that entry-level categories appear to be contextual to the user, leading to a variety of valid conceptualizations of entry-level BCs. The consensus among participants was that, ultimately, it is the role of the analyst to define which entry categories are better aligned to their goals.

Participants identified that words used frequently within a context (G3) can help organize information in an efficient manner and that basic-level categories are generally the most differentiated (G4), providing UGC app users with classes capable of helping them traverse a knowledge base effectively. Although participants noted the utility of our fifth guideline (G5: Object Visualization) to identify meaningful BCs, some of the participants expressed confusion regarding the application of this guideline. The term "visualization" made participants think of visual cues (from a UI/UX view) rather than whether the BC triggered a mental image—i.e., of a prototypical object such as a dog or a bird. Thus, we note the importance of stressing that G5 applies to the classes in the domain rather than to interface objects.

Participants considered the predictive utility (G8) guideline to be intuitive because of the products and services they use on a regular basis (e.g., Amazon's recommendation system, Netflix's movie recommendation engine). There was a general consensus that G8 can help organize information based on the likelihood of an event (e.g., the likelihood of a bird flying is higher than the likelihood of any animal flying). The WebMD symptom checker uses a bar meter that determines conditions on the basis of symptoms selected by the user. In this interface, conditions are listed according to likelihood. The focus group participants argued that the likelihood of events (as reflected by the bar meter) improved the user experience by providing relevant recommendations. However, a participant in FG1 suggested that the range of plausible diagnoses provided by the Mayo Clinic symptom checker was too extreme. Nevertheless, users maintained that for the citizen science app, likelihood could serve as a way of inferring objects based on the object's characteristics (features). For example, if a citizen scientist stated that they saw a white bird with a long neck, long legs, and a yellow beak in the Florida Everglades, a biologist would most likely infer that a great white heron was seen. A participant in FG3 felt that interfaces should allow users to enter features (attributes) about the object in order to gather information about the object.

A recurrent theme in our FGs was the value of considering all the guidelines together. For instance, although the word *mushroom* is not simpler (morphologically) (G6) than *fungus*, it is more likely to be learned by children first (G7). Adults are mindful about the kind of language that is appropriate for use with children (e.g., long names are troublesome for children). In general, participants agreed that BCs tend to be at a level that is easily relatable to users. A participant in FG3 gave an example that went beyond our task and highlighted how different organizations can leverage the idea of entry categories to organize information effectively. Frequently used words derived from interactions with existing users can help organize information that will be consumed by future users.

In summary, the focus groups demonstrated the utility of the BCs guidelines and the value of the BCs in developing and using applications. We are encouraged by the reception of the guidelines by our participants. We also noted the emergence of rich ideas and concepts from the focus group methodology. For example, we learned about the importance of both user-context and application-context (environment) information. The focus group discussions also provided strong support for our contention that an overlap exists between guidelines. Moreover, participants argued that certain guidelines (e.g., G1, G2) were easier to adopt and were thus referred to more often (see Table 8 and Table F1), suggesting that the totality of evidence should be taken into account when selecting the most appropriate BCs for a project.

6   Implications for Research and Practice

Traditionally, conceptual modeling research has relied extensively on users for the identification and selection of classes in a domain. Analysts are advised to represent views of users no matter how deficient they appear (Gemino & Wand, 2004). However, in an increasingly expanding range of applications, this practice has become problematic. For example, when modeling systems to capture UGC, analysts may no longer rely on the ability to reach all relevant users. Even if each user is reached, these users may not be subject matter experts and their requirements may not be as accurate and reliable as in traditional settings. In online settings, user views may be extremely diverse and even change over time—further complicating the ability to achieve consensus and generate a common unified view of the domain. In each case, traditional approaches to conceptual modeling may be limited. This paper contributes to the theory and practice of conceptual modeling and development of emerging IS by proposing a novel approach to conceptual modeling in UGC applications based on the notion of basic-level categories, a widely researched topic in psychology.

Having identified basic-level categories as a valuable idea for conceptual modeling, this paper proposed guidelines for identifying BCs in a domain. These guidelines are derived from well-established propositions in psychology research that have been corroborated in numerous empirical studies. These guidelines provide concrete practical procedures that analysts could follow when performing conceptual modeling.

As there can be substantial procedural ambiguity when applying theoretical design guidelines in practice (Chandra Kruse et al., 2015; Gregor & Jones, 2007; Chandra Kruse et al., 2016; Lukyanenko & Parsons, 2013), we took additional steps to further support practice (Iivari, 2007). First, we provided operational definitions, and when necessary, application heuristics to ensure that the application of the theoretical claims in psychology weas precise and consistent. Second, we provided examples to illustrate the application of each guideline and discussed potential pitfalls in implementation by referencing the relevant work in psychology. Third, we evaluated the utility of these guidelines via focus groups and found that, to different extents, these guidelines are beneficial when eliciting classes from potential users and different knowledge bases. Finally, recognizing that the guidelines we proposed in this paper can be automated, enabling the discovery of BCs in big data sets, we offered suggestions for building automatic routines (e.g., based on the CU formula in Appendix C). Finally, we evaluated the utility of the proposed guidelines in a series of focus groups with perspective analysts and

developers. The focus group evidence shows that practitioners appreciated the value of the proposed guidelines and found the notion of the basic level useful in identifying classes. Taken together, we believe that the proposed guidelines and application strategies constitute an important novel addition to the conceptual and practical toolbox in IS development.

Grounded in established research in psychology, we believe that the guidelines for identifying and applying BCs constitute a powerful tool for design and action. Our primary motivation in this paper was the need to support UGC. We suggest that BCs may safely be relied upon as starting points of data collection, as they can help narrow design possibilities (e.g., filter lists of more specialized classes from which online users can select to report on observed or experienced phenomena). These types of classes may be also used, for example, in the development of major sections of a project, for organizing menus, or to create training and tutorial elements. However, we do not believe that the potential uses of BCs end there. We strongly encourage future research to leverage the concept of BCs and the guidelines for choosing them in a variety of other applications. To motivate this work, we briefly suggest some of the possibilities for future research and discuss extensions of the notions proposed in this paper.

First, BCs open a novel opportunity to increase rigor in IS studies that use classes or categories. For example, experimental work in conceptual modeling commonly involves giving analysts and users a conceptual modeling script that represents a domain (Bodart et al., 2001; Burton-Jones & Meso, 2008; Burton-Jones, Wand, & Weber, 2009; Gemino & Wand, 2003; Parsons & Cole, 2005). While such scripts can be constructed using meaningless words (Parsons, 2011), the scripts often contain meaningful concepts at various levels of familiarity to the analysts (e.g., (Khatri et al., 2006)). Some of these concepts could be deemed BCs. The presence of BCs in such scripts can potentially confound experimental findings due to their cognitive privilege and people might be attracted to those levels in answering questions. Likewise, BCs can inadvertently appear in experimental work on human-computer interaction (e.g., as choices, section headers, or data collection or navigational elements). Thus far, we are not aware of any work that considers the potential confounding effects based on the presence of BCs in research.

Second, we believe our work opens exciting opportunities for the development of new theoretical concepts in conceptual modeling and knowledge management. Conceptual modeling research generally does not distinguish classes within a taxonomy (e.g., it assumes that all classes elicited from users may be equally relevant); however, not all classification levels are equally salient for different users. We suggest that

some classes in a domain have particularly interesting properties. An intriguing theoretical consequence of the BC concept is the idea of an *information gradient*. The salience of basic-level categories for individuals suggests that classes in a domain can be arranged in the order of their category utility, salience, and familiarity, rather than taxonomically. For example, using the category utility criteria (Appendix C), classes in Figure 1 can be arranged in descending order of category utility, which would result in the sequence of *bird*, *animal*, *osprey*. We call such arrangement of classes an *information gradient* (in contrast to the traditional generalization and specialization hierarchy that is based on property inheritance). The gradient concept can be used as an alternative to hierarchical representations of knowledge that are based on category utility, category salience, or other functions derived from research on BCs.

As taxonomies underlie much of modern science and technology, we believe that the concept of information gradient has the potential to make a broad contribution. Information gradients become a novel form of knowledge organization. They can be used to compare common knowledge with expert hierarchies, identify inconsistencies between intuitive and expert knowledge, and suggest potential conflicts. Information gradients can provide valuable input for information technology design (e.g., by suggesting which concepts among many are more and less salient for people, potentially affecting information collection, search, and retrieval). Gradients may naturally differ in their shapes (e.g., some may have multiple minima and maxima or sharp vertical distances between nodes), leading to different outcomes for how people use information and relate to the world. We hope that future research will build on the intriguing possibilities implied by the special status of BCs and expand the notion of the information gradient.

Third, a particularly important potential application of BCs is in the design of mobile and wearable devices. The challenge when creating mobile, wearable, or miniaturized interfaces is the scarcity of visual spaces and novel ergonomic restrictions (e.g., smaller screens, particular convenient input/output facilities, and lower processing capabilities) (Adipat, Zhang, & Zhou, 2011; Chae et al., 2002; Chittaro, 2006). BCs can help develop and manage systems with constrained visual spaces by providing natural content and flow content partitions that are easy for average users to understand and relate to.

Fourth, as projects are beginning to leverage natural language processing and artificial intelligence in guiding user input and analyzing user data (Gantz & Reinsel, 2012; Kao & Poteet, 2007), the knowledge of BCs may be leveraged in the design of artificial algorithms. For example, a conversational artificial agent can be modeled with the knowledge of BCs and this can be leveraged in supporting seamless communication with nonexpert users. Another promising application of BCs is in enhancing the transparency and understandability of complex machine learning models (e.g., neural networks) (Adadi & Berrada, 2018). It is feasible to posit that the intricate paths within a neural network could be abstracted to a set of BCs, which would subsequently offer a generic, high-level overview of the kinds of objects a neural network acts upon that would be accessible to nonexpert users.

Fifth, we hope that future research begins to investigate the best usage of BCs in conjunction with other classes. Despite the many benefits, it is important to underscore that relying on BCs *alone* for collecting or processing information may be insufficient for many projects. An IS designed using only BCs will collect information that, in most cases, is too general for any specialized use of the data. Returning to the context of iSpot, for example, the data consumers of the project—i.e., scientists and environmental agencies—would not likely find contributions expressed merely as BCs useful for their typical needs. For example, knowing that there are 50 birds and 10 trees observed does not carry significant utility for most projects (because of inferential utility, as discussed above). Instead, for most applications, it would be important to collect additional information at higher levels of specificity or precision (e.g., specific species of birds, health symptoms, geographical features, product categories). We believe BCs can be most useful to organize data collection into sections or subsections. For example, a project could provide a list of BCs as the first step, which would narrow the options down to only birds or only trees. Having achieved this narrowing, projects could then apply the other principles for collecting information (e.g., ask additional questions, allow users to type additional attributes of the observed bird or tree, or ask users to select from a predefined list of bird species, provided contributors have sufficient expertise to perform this task—see Lukyanenko et al., 2014b; Wiggins & He, 2016). Being equipped with the new tool of BCs, we call upon researchers and practitioners to find creative ways to leverage this tool in conjunction with other design solutions.

Finally, we note the degree of subjectivity in the application of the guidelines. In this paper, we adopted the notion of basic-level categories from psychology research and took numerous steps to enable practitioners to operationalize this important concept in IS (e.g., we turned theoretical propositions into actionable guidelines, added heuristics, provided multiple examples on how to apply the guidelines, and utilized focus groups to evaluate the ability of practitioners to work with the guidelines). Despite

these efforts, local adaptation and contextual choices may be required for practitioners to implement these guidelines in their specific projects. For example, our guidelines do not inform developers where to obtain data sources (e.g., taxonomical hierarchies, ontologies, or text corpus), or interviewees. Researchers have argued that it is important to avoid overprescribing design decisions in order to promote creativity, freedom of expression, and increase the applicability of design science research to a variety of future situations (Chandra Kruse et al., 2016). At the same time, as demonstrated in several studies, especially in the design science research community in IS, local choices on how to follow and implement design guidelines can measurably affect project outcomes (Tiefenbeck, 2017). We therefore urge practitioners to consult other relevant design research and best practices to inform the most effective application of our ideas in their projects. We also encourage future researchers to continue developing the notion and identification of basic classes and evaluating their boundary conditions and application in real projects (Seidel et al., 2018).

Acknowledgments

References

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI). *IEEE Access, 6*, 52138-52160.

Adipat, B., Zhang, D., & Zhou, L. (2011). The effects of tree-view based presentation adaptation on mobile web browsing. *MIS Quarterly, 35*(1), 99-121.

Anand, S., & Mobasher, B. (2003). Intelligent techniques for web personalization. *Proceedings of the 2003 international conference on Intelligent Techniques for Web Personalization.*

Arazy, O., Kumar, N., & Shapira, B. (2010). A theory-driven design framework for social recommender systems. *Journal of the Association for Information Systems, 11*(9), 455-490.

Archambault, A., Gosselin, F., & Schyns, P. G. (2000). A natural bias for the basic level. *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society.*

Barr, R., & Caplan, L. (1987). Category representations and their implications for category structure. *Memory & Cognition, 15*(5), 397-418.

Baskerville, R., De Marco, M., & Spagnoletti, P. (2013). *Designing Organizational Systems*. Springer.

Berlin, B., Breedlove, D., & Raven, P. (1973). General principles of classification and nomenclature in folk biology. *American Anthropologist, 75*(1), 214-242.

Björgvinsson, E., Ehn, P., & Hillgren, P.-A. (2012). Agonistic participatory design: working with marginalised social movements. *8*(2-3), 127-144.

Bodart, F., Patel, A., Sim, M., & Weber, R. (2001). Should optional properties be used in conceptual modelling? A theory and three empirical tests. *Information Systems Research, 12*(4), 384-405.

Bødker, S. (1996). Creating conditions for participation: conflicts and resources in systems development. *Human-Computer Interaction, 11*(3), 215-236.

Borgida, A. (1985). Features of languages for the development of information systems at the conceptual level. *IEEE Software, 2*(1), 63-72.

Boster, J. S. (1986). Exchange of varieties and information between Aguaruna manioc cultivators. *American Anthropologist, 88*(2), 428-436.

Bowers, J. S., & Jones, K. W. (2008). Detecting objects is easier than categorizing them. *Quarterly Journal of Experimental Psychology, 61*(4), 552-557.

Bratteteig, T., & Wagner, I. (2014). *Disentangling participation: Power and decision-making in participatory design*. Springer.

Brodie, M. L. (1984). On the development of data models. In Brodie M. L., Mylopoulos J., Schmidt J. W. (Eds.), *On conceptual modelling* (pp. 19-47). Springer.

Brown, R. (1958). How shall a thing be called? *Psychological Review, 65*(1), 14-21.

Browne, G. J., & Pitts, M. G. (2004a). Stopping rule use during information search in design problems. *Organizational Behavior Human Decision Processes, 95*(2), 208-224.

Browne, G. J., & Pitts, M. G. (2004b). Stopping rule use during information search in design problems. *Organizational Behavior and Human Decision Processes, 95*(2), 208-224.

Browne, G. J., Pitts, M. G., & Wetherbe, J. C. (2007). Cognitive stopping rules for terminating information search in online tasks. *MIS Quarterly*, *31*(1), 89-104.

Browne, G. J., & Ramesh, V. (2002). Improving information requirements determination: A cognitive perspective. *Information & Management, 39*(8), 625-645.

Brynjolfsson, E., Geva, T., & Reichman, S. (2015). Crowd-squared: Amplifying the predictive power of search trend data. *MIS Quarterly, 40*(4), 941-961.

Burgess, H., DeBey, L., Froehlich, H., Schmidt, N., Theobald, E., Ettinger, A., HilleRisLambers, J., Tewksbury, J., Parrish, J. (2017). The science of citizen science: Exploring barriers to use as a primary research tool. *Biological Conservation, 208*, 113-120.

Burton-Jones, A., & Meso, P. N. (2008). The effects of decomposition quality and multiple forms of information on novices' understanding of a domain from a conceptual model. *Journal of the Association for Information Systems, 9*(12), 748-802.

Burton-Jones, A., Wand, Y., & Weber, R. (2009). Guidelines for empirical evaluations of conceptual modeling grammars. *Journal of the*

Association for Information Systems, 10(6), 495-532.

Cantor, N., Smith, E. E., French, R. D., & Mezzich, J. (1980). Psychiatric diagnosis as prototype categorization. *Journal of Abnormal Psychology, 89*(2), 181-193.

Chae, M., Kim, J., Kim, H., & Ryu, H. (2002). Information quality for mobile internet services: A theoretical model with empirical validation. *Electronic Markets, 12*(1), 38-46.

Chandra Kruse, L., Seidel, S., & Gregor, S. (2015). *Prescriptive knowledge in IS research: Conceptualizing design principles in terms of materiality, action, and boundary conditions.* Paper presented at the 48th Hawaii International Conference on System Sciences.

Chandra Kruse, L., Seidel, S., & Purao, S. (2016). *Making use of design principles.* Paper presented at the International Conference on Design Science Research in Information Systems.

Checkland, P., & Holwell, S. (2006). *Information, systems and information systems: Making sense of the field.* Wiley.

Chittaro, L. (2006). Visualizing information on mobile devices. *Computer, 39*(3), 40-45.

Clark, E. V. (1979). Building a vocabulary: Words for objects, actions and relations. In P. Fletcher & M. Garman (Eds.), *Language Acquisition*, *Studies in first language development* (pp. 149-160). Cambridge University Press.

Clow, D., & Makriyannis, E. (2011). iSpot analysed: Participatory learning and reputation. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge.*

Corbin, J., & Strauss, A. (1990). Grounded theory research: Procedures, canons and evaluative criteria. *Zeitschrift für Soziologie, 19*(6), 418-427.

Corter, J., & Gluck, M. (1992). Explaining basic categories: Feature predictability and information. *Psychological Bulletin, 111*(2), 291-303.

Cotton, D., & Gresty, K. (2006). Reflecting on the think-aloud method for evaluating e-learning. *British Journal of Educational Technology, 37*(1), 45-54.

Craig, C. G. (1986). *Noun classes and categorization.* John Benjamins.

Crall, A. W., Newman, G. J., Stohlgren, T. J., Holfelder, K. A., Graham, J., & Waller, D. M. (2011). Assessing citizen science data quality: An invasive species case study. *Conservation Letters, 4*(6), 433-442.

de Beeck, H. O., & Wagemans, J. (2001). Visual object categorisation at distinct levels of abstraction: A new stimulus set. *Perception, 30*(11), 1337-1361.

Deng, X., Joshi, K., & Galliers, R. D. (2016). The duality of empowerment and marginalization in microtask crowdsourcing: Giving voice to the less powerful through value sensitive design. *MIS Quarterly, 40*(2), 279-302.

Dey, D., Storey, V. C., & Barron, T. M. (1999). Improving database design through the analysis of relationships. *ACM Transactions on Database Systems, 24*(4), 453-486.

DiSalvo, B., & DiSalvo, C. (2014). *Designing for democracy in education: Participatory design and the learning sciences.* International Society of the Learning Sciences.

Dobing, B., & Parsons, J. (2006). How UML is used. *Communications of the ACM, 49*(5), 109-113.

Downing, K., Ning, F., & Shin, K. (2011). Impact of problem-based learning on student experience and metacognitive development. *Multicultural Education Technology Journal, 5*(1), 55-69.

Ehn, P. (1988). *Work-oriented design of computer artifacts.* Arbetslivscentrum,

Erickson, J., Lyytinen, K., & Siau, K. (2005). Agile modeling, agile software development, and extreme programming: the state of research. *Journal of Database Management, 16*(4), 88-100.

Evermann, J., & Wand, Y. (2005). Ontology based object-oriented domain modelling: fundamental concepts. *Requirements Engineering, 10*(2), 146-160.

Feldman, R., & Dagan, I. (1995). Knowledge Discovery in Textual Databases (KDT). *Proceedings of the First International Conference on Knowledge Discovery and Data Mining.*

Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong.* Oxford University Press.

Friedman, G. M. (1967). Dynamic processes and statistical parameters compared for size frequency distribution of beach and river sands. *Journal of Sedimentary Research, 37*(2), 327-354.

Gangestad, S., & Snyder, M. (1985). "To carve nature at its joints": On the existence of discrete

classes in personality. *Psychological review, 92*(3), 317.

Gantz, J., & Reinsel, D. (2012). The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. *IDC iView: IDC Analyze the future, 2007*, 1-16, https://www.emc.com/leadership/digital-universe/2012iview/index.htm.

Gemino, A., & Wand, Y. (2003). Evaluating modeling techniques based on models of learning. *Communications of the ACM, 46*(10), 79-84.

Gemino, A., & Wand, Y. (2004). A framework for empirical evaluation of conceptual modeling techniques. *Requirements Engineering, 9*(4), 248-260.

Gennari, J. H., Langley, P., & Fisher, D. (1989). Models of incremental concept formation. *Artificial Intelligence, 40*(1-3), 11-61.

Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal, 69*(4), 211-221.

Gregor, S., & Jones, D. (2007). The anatomy of a design theory. *Journal of the Association for Information Systems, 8*(5), 312-335.

Gumm, D. C. (2006). Distributed participatory design: An inherent paradoxon? *Proceedings of IRIS29*.

Hand, E. (2010). People power. *Nature, 466*(7307), 685-687.

He, Y., & Wiggins, A. (2015). Community-as-a-service: data validation in citizen science. *Proceedings of the Fourth International Workshop on Methods for Establishing Trust of (Open) Data.*

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems, 19*(2), Article 4.

Hirschheim, R., & Heinz, K. (2012). A glorious and not-so-short history of the information systems field. *Journal of the Association of Information Systems, 13*(4), 188-235.

Hirschheim, R., Klein, H. K., & Lyytinen, K. (1995). *Information systems development and data modeling: Conceptual and philosophical foundations*. Cambridge University Press.

Ho, S. Y., Davern, M. J., & Tam, K. Y. (2008). Personalization and choice behavior: The role of personality traits. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems, 39*(4), 31-47.

Hosseinmardi, H., Ghasemianlangroodi, A., Han, R., Lv, Q., & Mishra, S. (2014). *Towards understanding cyberbullying behavior in a semi-anonymous social network.* Paper presented at the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining.

Iivari, J. (2007). A paradigmatic analysis of information systems as a design science. *Scandinavian Journal of Information Systems, 19*(2), Article 5.

Iivari, N. (2011). Participatory design in OSS development: Interpretive case studies in company and community OSS development contexts. *Behaviour Information Technology, 30*(3), 309-323.

Ipeirotis, P. G., Provost, F., & Wang, J. (2010). Quality management on Amazon Mechanical Turk. *Proceedings of the ACM SIGKDD Workshop on Human Computation.*

Jabbari Sabegh, M. A., Lukyanenko, R., Recker, J. C., Samuel, B., & Castellanos, A. (2017). Conceptual modeling research in information systems: What we now know and what we still do not know. *Proceedings of the AIS Symposium on Research in Systems Analysis and Design (SIGSAND).*

Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General, 126*(3), 248-277.

Johnson, S. L., Safadi, H., & Faraj, S. (2015). The emergence of online community leadership. *Information Systems Research, 26*(1), 165-187.

Jolicoeur, P., Gluck, M. A., & Kosslyn, S. M. (1984). Pictures and names: making the connection. *Cognitive Psychology, 16*(2), 243-275.

Jones, G. V. (1983). Identifying basic categories. *Psychological Bulletin, 94*(3), 423-428.

Kao, A., & Poteet, S. R. (2007). *Natural language processing and text mining..*

Kaufman, K. (1999). *Advanced birding*. Houghton Mifflin Harcourt.

Kaur, K., & Rani, R. (2013). *Modeling and querying data in NoSQL databases.* Paper presented at the 2013 IEEE International Conference on Big Data.

Khatri, V., Vessey, I., Ramesh, V., Clay, P., & Park, S.-J. (2006). Understanding conceptual schemas: Exploring the role of application and IS domain knowledge. *Information Systems Research, 17*(1), 81-99.

King, N. (1998). *Template analysis*. In G. Symon & C. Cassell (Eds.), *Qualitative methods and*

analysis in organizational research: A practical guide (p. 118-134). SAGE.

King, N. (2004). Using templates in the thematic analysis of text. *Essential Guide to Qualitative Methods in Organizational Research, 2*, 256-270.

Klibanoff, R. S., & Waxman, S. R. (2000). Basic level object categories support the acquisition of novel adjectives: evidence from preschool-aged children. *Child Development, 71*(3), 649-659.

Kosmala, M., Wiggins, A., Swanson, A., & Simmons, B. (2016). Assessing data quality in citizen science. *Frontiers in Ecology, 14*(10), 551-560.

Kraft, P., & Bansler, J. P. (1994). The collective resource approach: the Scandinavian experience. *Scandinavian Journal of Information Systems, 6*(1), Article 4.

Krueger, R. A., & Casey, M. A. (2000). *Focus groups: A practical guide for applied research*. SAGE.

Kyng, M. (1995). Making representations work. *Communications of the ACM, 38*(9), 46-55.

Lakoff, G. (1987). *Women, fire, and dangerous things.* University of Chicago Press.

Lakoff, G., & Johnson, M. (2008). *Metaphors we live by.* University of Chicago Press.

Lassaline, M. E., Wisniewski, E. J., & Medin, D. L. (1992). Basic levels in artificial and natural categories: Are all basic levels created equal? In B. Burns (Ed.), *Advances in psychology* (pp. 327-378). North-Holland.

Lazer, D., Brewer, D., Christakis, N., Fowler, J., & King, G. (2009). Life in the network: the coming age of computational social. *Science, 323*(5915), 721-723.

Le Dantec, C. A., Asad, M., Misra, A., & Watkins, K. E. (2015). Planning with crowdsourced data: Rhetoric and representation in transportation planning. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing.*

Levina, N., & Arriaga, M. (2014). Distinction and status production on user-generated content platforms: Using Bourdieu's theory of cultural production to understand social dynamics in online fields. *Information Systems Research, 25*(3), 468-488.

Levy, M., & Germonprez, M. (2017). The potential for citizen science in information systems research. *Communications of the Association for Information Systems, 40*(2), 22-39.

Lewandowski, E., & Specht, H. (2015). Influence of volunteer and project characteristics on data quality of biological surveys. *Conservation Biology, 29*(3), 713-723.

Louv, R., & Fitzpatrick, J. W. (2012). *Citizen science: Public participation in environmental research.* Cornell University Press.

Lukyanenko, R. (2016). Information quality research challenge: Information quality in the age of ubiquitous digital intermediation. *Journal of Data and Information Quality, 7*(1-2), 1-3.

Lukyanenko, R., & Parsons, J. (2013). Lukyanenko R., Parsons J. (2013) Reconciling theories with design choices in design science research. In J. vom Brocke, R. Hekkala, S. Ram, & M. Rossi (Eds.), *Design science at the intersection of physical and virtual design: DESRIST 2013.* Springer.

Lukyanenko, R., Parsons, J., & Wiersma, Y. (2014a). The impact of conceptual modeling on dataset completeness: A field experiment. *Proceedings of the International Conference on Information Systems.*

Lukyanenko, R., Parsons, J., Wiersma, Y., & Maddah, M. (2019). Expecting the unexpected: Effects of data collection design choices on the quality of crowdsourced user-generated content. *MIS Quarterly, 43*(2), 623-648.

Lukyanenko, R., Parsons, J., & Wiersma, Y. F. (2014b). The IQ of the crowd: Understanding and improving information quality in structured user-generated content. *Information Systems Research, 25*(4), 669-689.

Lukyanenko, R., & Samuel, B. M. (2017). Are All Classes Created Equal? Increasing Precision of Conceptual Modeling Grammars. *ACM Transactions on Management Information Systems, 8*(4), 1-15.

Lukyanenko, R., Wiersma, Y., Huber, B., Parsons, J., Wachinger, G., & Meldt, R. (2017). Representing crowd knowledge: Guidelines for conceptual modeling of user-generated content. *Journal of the Association for Information Systems, 18*(4), 297.

Luo, X., Zhang, J. J., Gu, B., & Phang, C. (2013). Expert blogs and consumer perceptions of competing brands. *MIS Quarterly, 41*(2), 371-395.

Maass, W., Parsons, J., Purao, S., Storey, V. C., & Woo, C. (2018). Data-Driven Meets Theory-Driven Research in the Era of Big Data: Opportunities and Challenges for Information

Systems Research. *Journal of the Association for Information Systems, 19*(12), 1253-1273.

Maccani, G., Donnellan, B., & Helfert, M. (2014). *Action design research in practice: the case of smart cities.* Paper presented at the International Conference on Design Science Research in Information Systems.

Macé, M. J.-M., Joubert, O. R., Nespoulous, J.-L., & Fabre-Thorpe, M. (2009). The time-course of visual categorizations: You spot the animal faster than the bird. *PloS One, 4*(6), e5927.

Mandler, J. M., & Bauer, P. J. (1988). The cradle of categorization: Is the basic level basic? *Cognitive Development, 3*(3), 247-264.

Markman, E. M. (1991). *Categorization and naming in children: Problems of induction*. MIT Press.

Mayden, R. L. (2002). On biological species, species concepts and individuation in the natural world. *Fish and Fisheries, 3*(3), 171-196.

Mazza, R., & Berre, A. (2007). *Focus group methodology for evaluating information visualization techniques and tools.* Paper presented at the 2007 11th International Conference Information Visualization.

McGinnes, S. (2011). Conceptual modelling for web information systems: What semantics can be shared? In O. de Troyer, C. Bauzer Medeiros, R. Billen, P. Hallot, A. Simitsis, H. van Mingroot (Eds.), *Advances in Conceptual Modeling. Recent Developments and New Directions: ER 2011.* Springer .

Medin, D. L. (1983). Structural principles in categorization. In T. J. Tige, B. E. Sharp (Eds.), *Perception, cognition, development: Interactional analyses* (pp. 203-230). Erlbaum.

Medin, D. L., Lynch, E. B., Coley, J. D., & Atran, S. (1997). Categorization and reasoning among tree experts: Do all roads lead to Rome? *Cognitive Psychology, 32*(1), 49-96.

Melumad, S., Inman, J. J., & Pham, M. T. (2019). Selectively Emotional: How Smartphone Use Changes User-Generated Content. *Journal of Marketing Research*, *56*(2), 259-275.

Mervis, C., & Crisafi, M. (1982). Order of acquisition of subordinate-, basic-, and superordinate-level categories. *Child Development*, *53*(1), 258-266.

Mervis, C., Golinkoff, R., & Bertrand, J. (1994). Two-year-olds readily learn multiple labels for the same basic-level category. *Child Development, 65*(4), 1163-1177.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: an expanded sourcebook*. SAGE.

Mobasher, B., Cooley, R., & Srivastava, J. (2000). Automatic personalization based on web usage mining. *Communications of the ACM, 43*(8), 142-151.

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G., & Worm, B. (2011). How many species are there on Earth and in the ocean? *PLoS Biology, 9*(8), e1001127.

Muller, M., Millen, D. R., & Strohecker, C. (2001). *What makes a representative user representative? a participatory poster.* Presented at the CHI'01 Extended Abstracts on Human Factors in Computing Systems.

Murphy, G. (2004). *The big book of concepts*. MIT Press.

Murphy, G. L. (1982). Cue validity and levels of categorization. *Psychological Bulletin, 91*(1), 174-177.

Murphy, G. L., & Brownell, H. H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(1), 70-84.

Murphy, G. L., & Wisniewski, E. J. (1989). Categorizing objects in isolation and in scenes: What a superordinate is good for. *Journal of Experimental Psychology-Learning Memory and Cognition, 15*(4), 572-586.

Myers, M. D., & Newman, M. (2007). The qualitative interview in IS research: Examining the craft. *Information and Organization, 17*(1), 2-26.

Mylopoulos, J., & Wong, H. K. (1980). Some features of the Taxis data model. *Proceedings of the Sixth International Conference on Very Large Data Bases*.

Nakamura, G. V., Medin, D. L., & Taraban, R. (1993). *Categorization by humans and machines*. Academic Press.

Neisser, U. (1987). From direct perception to conceptual structure. In U. Neisser (Ed.), *Emory symposia in cognition, 1. Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 11-24). Cambridge University Press.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Prentice Hall

Nickles, K. R., Curley, S. P., & Benson, P. G. (1995). *Judgment-based and reasonging-based stopping rules in decision making under uncertainty*. University of Minnesota Press.

Nielsen, J., Clemmensen, T., & Yssing, C. (2002). Getting access to what goes on in people's

heads?: Reflections on the think-aloud technique. *Proceedings of the second Nordic conference on Human-Computer Interaction.*

Nixon, B., Chung, L., Mylopoulos, J., Lauzon, D., Borgida, A., & Stanley, M. (1987). Implementation of a compiler for a semantic data model: Experiences with taxis. *ACM SIGMOD Record, 16*(3), 118-131.

Nov, O., Arazy, O., & Anderson, D. (2011). Dusting for science: Motivation and participation of digital citizen science volunteers. *Proceedings of the 2011 iConference.*

O'Callaghan, C. (2017). *Beyond vision: Philosophical essays*: Oxford University Press.

Obendorf, H., Janneck, M., & Finck, M. (2009). Inter-contextual distributed participatory design. *Scandinavian Journal of Information Systems, 21*(1), Article 2.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology, 17*(4), 273-281.

Ozcan, E., van Egmond, R., & Jacobs, J. (2014). Product sounds: Basic concepts and categories. *International Journal of Design, 8*(3), 97-111.

Parsons, J. (2011). An experimental study of the effects of representing property precedence on the comprehension of conceptual schemas. *Journal of the Association for Information Systems, 12*(6), 441-462.

Parsons, J., & Cole, L. (2005). What do the pictures mean? Guidelines for experimental evaluation of representation fidelity in diagrammatical conceptual modeling techniques. *Data & Knowledge Engineering, 55*(3), 327-342.

Parsons, J., & Wand, Y. (1997). Choosing classes in conceptual modeling. *Communications of the ACM, 40*(6), 63-69.

Parsons, J., & Wand, Y. (2008). Using cognitive principles to guide classification in information systems modeling. *MIS Quarterly*, *32*(4), 839-868.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8*(12), 976-987.

Peckham, J., & Maryanski, F. (1988). Semantic data models. *Computing Surveys, 20*(3), 153-189.

Philippoff, J., & Baumgartner, E. (2016). Addressing common student technical errors in field data collection: An analysis of a citizen-science

monitoring project. *Journal of Microbiology and Biology Education, 17*(1), 51-55.

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences, 108*(9), 3526-3529.

Ploesser, K. (2013). *A design theory for context-aware information systems* (Doctoral dissertation, Queensland University of Technology). http://eprints.qut.edu.au/60865/1/Karsten_Ploesser_Thesis.pdf

Prat, N., Comyn-Wattiau, I., & Akoka, J. (2015). A taxonomy of evaluation methods for information systems artifacts. *Journal of Management Information Systems, 32*(3), 229-267.

Purao, S., Seng, T. C., & Wu, A. (2013). *Modeling citizen-centric services in smart cities.* Paper presented at the International Conference on Conceptual Modeling.

Purao, S., & Storey, V. C. (1997). *Intelligent support for retrieval and synthesis of patterns for object-oriented design.* Paper presented at the International Conference on Conceptual Modeling.

Rai, A. (2016). Editor's comments: Synergies between big data and theory. *MIS Quarterly, 40*(2), iii-ix.

Rai, A. (2017). Editor's comments: Diversity of design science research. *MIS Quarterly, 41*(1), iii-xviii.

Ranchordás, S. (2018). Citizens as consumers in the data economy: The case of smart cities. *Journal of European Consumer Market Law, 7*(4), 154-161.

Raven, P. H., Berlin, B., & Breedlove, D. E. (1971). The origins of taxonomy. *Science, 174*(4015), 1210-1213.

Recker, J. (2015). *Research on conceptual modelling: less known knowns and more unknown unknowns, please.* Paper presented at the Proceedings of the 11th Asia-Pacific Conference on Conceptual Modelling.

Reynolds, R. T., Scott, J. M., & Nussbaum, R. A. (1980). A variable circular-plot method for estimating bird numbers. *The Condor, 82*(3), 309-313.

Rhemtulla, M., & Hall, D. (2009). Basic-level kinds and object persistence. *Memory & Cognition, 37*(3), 292-301.

Roach, E., Lloyd, B. B., Wiles, J., & Rosch, E. (1978). Principles of categorization. In L. Roach & B.

Lloyd (Eds.), *Foundations of Cognitive Psychology* (pp. 189-206). Erlbaum.

Robertson, T., & Simonsen, J. (2012). Challenges and opportunities in contemporary participatory design. *Design Issues, 28*(3), 3-9.

Rorissa, A. (2008). User-generated descriptions of individual images versus labels of groups of images: A comparison using basic level theory. *Information Processing & Management, 44*(5), 1741-1753.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyesbraem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*(3), 382-439.

Samuel, B. M., Khatri, V., & Ramesh, V. (2018). Exploring the effects of extensional versus intentional representations on domain understanding. *MIS Quarterly, 42*(4), 1187-1209.

Scanlon, E., Woods, W., & Clow, D. (2014). Informal participation in science in the UK: Identification, location and mobility with iSpot. *Educational Technology & Society, 17*(2), 58-71.

Seidel, S., Chandra Kruse, L., Székely, N., Gau, M., & Stieger, D. (2018). Design principles for sensemaking support systems in environmental sustainability transformations. *European Journal of Information Systems, 27*(2), 221-247.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27*(4), 623-656.

Silver, C., & Lewins, A. (2014). *Using software in qualitative research: A step-by-step guide*: SAGE.

Silvertown, J. (2010). Taxonomy: Include social networking. *Nature, 467*(7317), 788-788.

Sinkonde, D., Mselle, L., Shidende, N., Comai, S., & Matteucci, M. (2018). Developing an intelligent PostGIS database to support accessibility tools for urban pedestrians. *Urban Science, 2*(3), 52, 1-13.

Smith, A., & Page, D. (2015). US smartphone use in 2015. *Pew Research Center.* https://www.pewresearch.org/internet/2015/04/01/us-smartphone-use-in-2015/

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Harvard University Press.

Smith, J. A. (1996). Beyond the divide between cognition and discourse: Using interpretative phenomenological analysis in health psychology. *Psychology and Health, 11*(2), 261-271.

Smith, J. M., & Smith, D. C. (1977). Database abstractions: aggregation and generalization. *ACM Transactions on Database Systems, 2*(2), 105-133.

Soffer, P., Wand, Y., & Kaner, M. (2015). Conceptualizing routing decisions in business processes: Theoretical analysis and empirical testing. *Journal of the Association for Information Systems, 16*(5), 345-393.

Sowa, J. F. (1995). Top-level ontological categories. *International Journal of Human-Computer Studies, 43*(5-6), 669-685.

Stahl, B. C., Tremblay, M. C., & LeRouge, C. M. (2011). Focus groups and critical social IS research: How the choice of method can promote emancipation of respondents and researchers. *European Journal of Information Systems, 20*(4), 378-394.

Stewart, D. W., & Shamdasani, P. N. (2014). *Focus groups: Theory and practice*. SAGE.

Storey, V. C. (1993). Understanding semantic relationships. *The VLDB Jorunal, 2*(4), 455-488.

Storey, V. C., & Song, I.-Y. (2017). Big data technologies and management: what conceptual modeling can do. *108*, 50-67.

Susarla, A., Oh, J.-H., & Tan, Y. (2012). Social networks and the diffusion of user-generated content: Evidence from YouTube. *Information Systems Research, 23*(1), 23-41.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology, 23*(3), 457-482.

Teorey, T. J., Yang, D., & Fry, J. P. (1986). A logical design methodology for relational databases using the extended entity-relationship model. *ACM Computing Surveys, 18*(2), 197-222.

Tiefenbeck, V. (2017). Bring behaviour into the digital transformation. *Nature Energy, 2*(6), 1-3.

Tremblay, M. C., Hevner, A. R., & Berndt, D. J. (2010). Focus groups for artifact refinement and evaluation in design research. *Communications of the Association for Information Systems, 26*, 599-618.

Tremblay, M. C., Hevner, A. R., & Berndt, D. J. (2012). Design of an information volatility measure for health care decision making. *Decision Support Systems, 52*(2), 331-341.

Tversky, B., & Hemenway, K. (1983). Categories of environmental scenes. *Cognitive Psychology, 15*(1), 121-149.

Ullrich, H., Purao, S., & Storey, V. C. (2000). *An ontology for classifying the semantics of relationships in database design.* Paper presented at the International Conference on Application of Natural Language to Information Systems.

Ulrich, K. (1995). The role of product architecture in the manufacturing firm. *Research Policy, 24*(3), 419-440.

van Aken, J., Chandrasekaran, A., & Halman, J. (2016). Conducting and publishing design science research: Inaugural essay of the design science department of the *Journal of Operations Management. Journal of Operations Management, 47*, 1-8.

Van Kleek, M. G., Styke, W., Schraefel, M., & Karger, D. (2011). Finders/keepers: A longitudinal study of people managing information scraps in a micro-note tool. *Proceedings of the International Conference on Human Factors in Computing Systems*,

Vequist, G. W., & Licht, D. S. (2013). *Wildlife watching in America's national parks: A seasonal guide*: Texas A&M University Press.

Wales, R., Colman, M., & Pattison, P. (1983). How a thing is called: A study of mothers' and children's naming. *Journal of Experimental Child Psychology, 36*(1), 1-17.

Wand, Y., Storey, V. C., & Weber, R. (1999). An ontological analysis of the relationship construct in conceptual modeling. *ACM Transactions on Database Systems, 24*(4), 494-528.

Wattal, S., Schuff, D., Mandviwalla, M., & Williams, C. B. (2010). Web 2.0 and politics: The 2008 US presidential election and an e-politics research agenda. *MIS Quarterly*, *34*(4), 669-688.

Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology, 36*(5), 571-581.

Weller, M. W. (1999). *Wetland birds: Habitat resources and conservation implications*. Cambridge University Press.

Welsh, D. (1988). *Codes and cryptography*. Oxford University Press.

Wiggins, A., & He, Y. (2016). Community-based data validation practices in citizen science. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing.*

Younger, B. A., & Fearing, D. D. (2000). A global-to-basic trend in early categorization: evidence from a dual-category habituation task. *Infancy*, *1*(1), 47-58.

Zhou, H., Liu, J., Jing, W., Qin, Y., Lu, S., Yao, Y., & Zhong, N. (2010). Basic level advantage and its switching during information retrieval: An fMRI study. In Y. Yao, R. Sun, T. Poggio, J. Liu, N. Zhong, & J. Huang (Eds.), *Brain Informatics* (pp. 427-436). Springer.

Zipf, G. K. (1935). *The psycho-biology of language: An introduction to dynamic philology*. Houghton Mifflin.

Appendix A: Sample Projects with Characteristics of Interest to Our Paper

In Table A1 we present a series of UGC projects classified by industry and our characteristics of interest:

- Project-purpose-driven information collection: the specific kind of information a project is designed to collect, generally to meet organizational information needs)

- Project openness: projects are open when participation is not restricted to some subgroup in a population or domain experts (i.e., when anyone interested can register and participate)

- Lean user profile: when profile information is insufficient for reliable assessment of an individual user's level of domain expertise, skills, and motivation.

Table A1. UGC Projects Classified by Scope, Openness, and Lean User Profile

Organization/industry	Project purpose and scope	Project openness	Lean user profile
SalesForce Success community *Business*	Share and vote for ideas through an online forum ("IdeaExchange") to improve the product.	Available only to individuals with Salesforce credentials.	The user profile includes picture, description, industry, and products used. It also allows the user to link other social profiles. Providing social profile data is optional.
My Starbucks Idea *Business*	Help increase the company's focus on customers and their needs. Structured (vote on existing ideas) and unstructured (submit new ideas). Users submit their ideas in 500 characters or less together with their contact information. Users must choose a category for the idea (e.g., store, coffee, milks offered).	Open to anyone. Tailored to existing customers knowledgeable about Starbucks products and services with ideas for improving service.	Only name and email are required to participate. These data are required to submit an idea.
Amazon *e-Commerce*	Crowd-sourced reviews about products sold. Semistructured (predefined categories) and unstructured. Predefined dimensions plus text field.	Anyone can create an account. Amazon account users can register for other services provided by the company.	An Amazon account is necessary to submit a written or video review. For other services, more data is required (including method of payment).
Yelp *Business*	Crowd-sourced reviews about local businesses. Semistructured (predefined categories) and unstructured. Predefined dimensions plus text field.	No account needed to view reviews. Anyone can create an account to write a review.	Registration is required to post a review for a local business. Name, email, and zip code are required to register, birthday is optional. Signup via Facebook account is also permitted.
Trip Advisor *Travel*	Reviews of travel-related content, including forums. Semistructured (predefined categories) and unstructured.	No account needed to view reviews. Anyone can create an account to write a review.	Registration is required to post a review. Name, email, and zip code are required to register, birthday is optional. Signup via Facebook account is also permitted.
Asteroid Zoo *Astronomy / citizen science*	Classify unknown asteroids. Semistructured—predefined characteristics to identify in an image. Within 24 hours of launch, the site was receiving almost 70,000 classifications per hour.	Available to anyone (citizen scientists).	No account needed to start classifying galaxies. Creation of a profile requires username and email and is optional.
Bee Spotter *Entomology / citizen science*	A citizen science project where users register, take pictures of bees, and try to classify their observations. Color pattern, female vs. male, bee anatomy. Structured—image similarity.	No account needed to view bee spottings. Anyone can create an account to submit bee spottings.	Name, email, and username are required to register.

Riskmap.us Flood reporting / crowdsourcing	Used for Hurricane Irma via user-generated reports of flooding. Open source initiative to map urban flooding and provide real-time information to emergency responders and citizens posted by citizens. Structured and unstructured. Geolocation + depth + image + free text.	Open to anyone. Account needed to report flooding but anyone can view zones at risk.	To input flood reports registration is required via Facebook, Twitter, or Telegram.
iNaturalist *Wildlife / citizen science*	Records user encounters with other organisms and connects users with experts who can identify the organisms observed. Structured (select from list) and unstructured (report observations).	Available to anyone (citizen scientists).	User must create an account to participate. Name, username, and email are required.
Fix My Street *Government*	Sinkhole and pothole mapping by citizens. Unstructured (forum allowing users to discuss findings).	Available to anyone (citizen scientists).	User must create an account to participate. Name, username, and email are required.

Appendix B: Theoretical Support for Each of the Guidelines

Table B1. Theoretical Support for Each of the Guidelines

Theoretical underpinnings for G1:	• People consistently use middle-level concepts in speech (Brown, 1958). • A hierarchy develops in both directions from the middle level of abstraction (Brown, 1958). Objects at the subordinate (lower than basic) levels need higher perceptual processing compared to those at the basic level (Jolicoeur et al., 1984). *The basic level falls somewhere in the middle* of taxonomic hierarchies, regardless of how many levels of inclusiveness they contain (Ulrich, 1995). • Objects are typically identified at a particular level of abstraction that is *neither the most general nor the most specific* possible (Jolicoeur et al., 1984) but an intermediate one called basic level (Rosch et al., 1976). • "The middle level is the first level where one finds rich prototypes … the features at this level are distinctive, as opposed to those of specific categories" (Cantor et al., 1980). • The basic level falls somewhere in the *middle of taxonomic hierarchies*, regardless of how many levels of inclusiveness they contain (Neisser, 1987). • "the most natural, preferred level at which to conceptually carve up the world. The basic level can be seen as a compromise between the accuracy of classification at a maximally general level and the predictive power of a maximally specific level (Murphy, 2004)." • Middle-level categories are learned most quickly or could be named most quickly after they were learned (Corter & Gluck, 1992).
Derived guideline:	**G1 – Middle Level: Identify classes in a domain in the middle of the conceptual hierarchy.** If the following hierarchy: animal—bird—osprey, the basic-level category would be that of the taxonomic middle, in this example, bird.
Theoretical underpinnings for G2:	• It has been suggested that basic-level categories often become an *entry-level category*—the first concept thought of by a user when encountering a phenomenon (Jolicoeur et al., 1984). Murphy and Brownell (1985) called it the "necessary first step" of identification (p. 72). These classes tend to be retrieved extremely fast, accurately, and efficiently. • Jolicoeur et al. (1984) and Murphy and Brownell (1985) introduced the concept of entry-level category to explain the shorter reaction times found at the subordinate level for some atypical members of basic-level categories (e.g., a penguin is categorized faster as a penguin than as a bird—since its appearance is distant from the prototypical bird) (Macé, Joubert, Nespoulous, & Fabre-Thorpe, 2009). • For typical members of basic-level categories, the entry point is usually at the basic level. Expertise is likely to shift the entry category toward subordinate levels (Rosch et al., 1976). • Entry categories are usually at the basic level but not always. To access categories below the entry point, additional information is required (Archambault, Gosselin, & Schyns, 2000). • The particular entry point for a given object covaries with its typicality, which affects whether or not the object will be identified at the basic level (Jolicoeur et al., 1984). • The entry point in the formation of a hierarchical categorization system may be at the "unique beginner level" or at the next level down (Berlin et al., 1973). • One of the most important features of basic-level concepts consists in the fact that they provide us with *much information with little cognitive effort* (Murphy, 2004; Roach et al., 1978). • Experts should be able to categorize objects at the subordinate level as quickly as objects at the basic level because their basic- and subordinate-level *categories are equally differentiated* (Tanaka & Taylor, 1991).
Derived guideline:	**G2 – Entry Category: Elicit entry categories from a sample of potential users for the domain objects of interest.** Example: a visual stimulus such as a robin first activates the bird category, providing rapid access to the name "bird" and other typical bird properties (e.g., has wings and can fly) (Patterson et al., 2007). A bird expert could verify an object as a robin or as a bird with equal speed. In the novice domain, verification times are fastest at the basic level (Tanaka & Taylor, 1991).

Theoretical underpinnings for G3:	• The more frequently a word is used, the smaller its average length tends to be and the fewer synonyms it has (Zipf, 1935). • Individual differences of classification can be a function of idiosyncratic life *experiences and/or culture* and, thus, the importance of eliciting entry categories from potential users (e.g., tree vs. oak) can be context dependent. • Words in basic-level categories tend to be used more frequently in English than words in superordinate or subordinate categories (Corter & Gluck, 1992). • Boster (1986) found that Aguarana women, who are typically engaged in in cultivating manioc, tended to refer to manioc plants with highly specific (species-level) names. Other members who interacted less with manioc named these plants at the basic level (Brown, 1958; Wales et al., 1983).
Derived guideline:	**G3 – Frequently Used Words: Identify the most frequently used domain words used in a typical discourse.** For example, people can more quickly categorize a boxing glove as a boxing glove than as a glove, even though the latter is the basic-level category (Murphy & Brownell, 1985). People across cultures tend to use the same level of concepts in naming animals and plants (B. Berlin, Breedlove, Raven, & Hammel, 2013)
Theoretical underpinnings for G4:	• The ratio of *within-category to between-category similarity* is highest for the middle level (Tversky & Hemenway, 1983). • One way to characterize categories at a privileged level is in terms of *similarity relationships,* or patterns of *common and distinctive properties* or features that define the within and the between-category similarity. A privileged level is one at which within-category similarity is high relative to between-category similarity (Medin et al., 1997). • Basic-level categories maximize within-category similarity relative to between-category similarity (Murphy & Brownell, 1985). Within-category similarity is maximal for categories that are more specific, and between-category similarity is minimal for the most general categories (Medin, 1983). • A privileged category is one in which category members are very *similar to each other and not very similar to members of other categories* (Murphy & Brownell, 1985).
Derived guideline:	**G4 – Cohesion and Coupling: Find a domain taxonomic level, for which sibling categories have maximal difference and their respective children have maximal similarity** In biology, such classes could be *animals* and *plants*. By storing only a few classes, humans can easily memorize the identifying characteristics of different classes.
Theoretical underpinnings for G5:	• A concept is a *mental representation of an object* or a class of similar objects (Lakoff & Johnson, 2008; Murphy, 2004). Categories can occur as a result of sensory perception and the cognitive, conceptual, and emotional processing of objects (Ozcan, van Egmond, & Jacobs, 2014). • Basic level is the most abstract level at which people are able to form an integrated perceptual representation of a category. Basic-level concepts are activated more quickly than subordinate concepts because they are perceptually distinctive (Rosch et al., 1976). • The basic level is a *level of abstraction of visual concepts* that maximizes between-category distinctiveness and within-category informativeness. Basic-level categories in which objects share a characteristic shape have the highest level of abstraction (Rosch et al., 1976). • There are exceptions to the finding that people classify images more quickly at the basic level than at the subordinate level (Jolicoeur et al., 1984). For example, a picture of a penguin is classified more quickly as a penguin than as a bird. • Expertise causes categories at subordinate levels to function as basic. As expertise is acquired, *overall shape also can be used to identify objects* at the subordinate level (Johnson & Mervis, 1997).
Derived guideline:	**G5 – Object Visualization: Find the highest category in the taxonomy for which category members can be easily visualized.** The outer shapes of most members of the category *dog* are so similar that it is possible to imagine a picture of a dog "as such." This is clearly impossible for superordinate categories because their members' outer shapes are too divergent. When shown a picture of a sparrow, most people think of it as a bird, not a sparrow (subordinate) or animal (superordinate). An apple is matched with the name "apple" faster than with "delicious apple" or with "fruit" A visual stimulus such as a shorebird first activates the bird node, providing rapid access to the name bird and other typical bird properties (e.g., has wings and can fly) (Patterson et al., 2007)

Theoretical underpinnings for G6:	• Zipf's law predicts that words belonging to the basic taxonomic level, because of their frequent use, will be labeled with *shorter, morphologically simpler terms* than words belonging to superordinate and subordinate levels (Craig, 1986). • The shorter names for anything will usually be the most frequently used names for that thing (Brown, 1958). • Words belonging to basic-level categories tend to be shorter and more frequently used in English than names of superordinate or subordinate categories (Corter & Gluck, 1992). • Infrequently used object names take longer to name than frequently used object names (Oldfield & Wingfield, 1965). Frequently used words tend to be short: "The magnitude of words tends, on the whole, to stand in an inverse (not necessarily proportionate) relationship to the number of occurrences (Zipf, 1949).
Derived guideline:	**G6 – Simplest Words: Among the classes in a domain, identify the shortest and morphologically simplest words.** The monosyllable *dog* is used with much higher frequency than *boxer*. It sometimes happens, however, that the frequency-brevity principle makes the wrong prediction. A pineapple is a fruit, yet the former word is more frequently used to refer to it.
Theoretical underpinnings for G7:	• The sequence in which words are acquired is not determined by the cognitive preferences of children so much as by the naming practices of adults" (Brown, 1958, p. 20). Mothers use more general and frequently used terms for their children (Wales et al., 1983). The names used to refer to *categories at this level tend to be brief*. Considerable agreement exists across time, languages, and children in the *first words children acquire* (Clark, 1979). • The basic level is the most frequently used in speech, and the *first learned by children* (Downing, Ning, & Shin, 2011). Mervis and Crisafi (1982) suggest that children's categorization ability is acquired in the order basic, superordinate, and subordinate. • Categories that are in the middle of the taxonomic hierarchy are learned first; then, children work up the hierarchy generalizing and down the hierarchy specializing (Lakoff, 1987). • When naming the same object for a child and an adult, adults will sometimes provide the child with a different name than the name they use with the adult (Anglin, 1977).
Derived guideline:	**G7 – Original Words: Identify the first words or concepts learned by children or used by mothers to talk to children.** A child might refer to a coin as a coin rather than a dime since children do not necessarily focus on the monetary value of the coin) (Brown, 1958). An adult would refer to the abdomen as *tummy* or *belly* to make it simpler for the child.
Theoretical underpinnings for G8:	• The best categories are those that maximize feature predictability and optimize information transfer (Corter & Gluck, 1992). • Natural language use is highly nonstationary as word probabilities change depending on their context (Piantadosi, Tily, & Gibson, 2011). • Mervis and Rosch (1981) found that basic-level categories are those that carry the most information about attributes. • One critically important function of categories is supporting inductive inferences; categories extend knowledge via inferences (Anderson, 1985).
Derived guideline:	**G8 – General Predictive Utility: Identify classes with the greatest general predictive utility.**

Appendix C: Formal Models of Basic-level categories

Psychology research has produced a number of formal models of basic-level categories, suggesting the potential for automation of the basic-level selection process.

An early model by Rosch et al. (1976) advocated *cue validity*, a sum of the conditional probabilities that an object belongs to a target class (e.g., fish) given that it possesses a set of attributes (e.g., can swim, has scales). Rosch et al. (1976) argued that since basic-level categories hold the greatest number of attributes, cue validity of such classes would be maximal. Murphy (1982) refuted this argument by pointing out that the cue validity model lacked constraints (e.g., limited cognitive capacity constraint) and was unbounded. To balance cue validity, another measure, *category validity* was proposed (Gregory L. Murphy, 1982). It reversed the conditional probability of cue validity and measured the probability of an object having features of interest (e.g., can fly, has wings) given that it is assigned a particular category (e.g., bat).

Combining cue and category validity models appeared to offer a mathematical balance to compensate for the lack of binding constraints. The problem, however, is that it is unclear how to combine category and cue validity in such a way that their individual contributions genuinely reflect the importance of these functions to humans. Several heuristic approaches and algorithms, mainly in artificial intelligence, cognitive science, and economics have been proposed. For instance, Jones (1983) developed a *collocation model* in which cue and category validity are multiplied to produce a concave function with a unique maximum. The collocation measure was argued to be maximal for basic-level categories (Jones, 1983). While the collocation model resolved the unboundedness issue of cue and category validity, it lacked a theoretical rationale for combining the two measures in a particular way (Corter & Gluck, 1992).

Building on the above theories, a model of classification optimality and category utility was proposed by Corter and Gluck (1985, 1992). This model is designed to directly operationalize the trade-off between cognitive economy and inferential utility in a way that adheres to the widely held propositions about human cognition in psychology research. This model has been applied in artificial intelligence and used as part of more complex algorithms (Gennari, Langley, & Fisher, 1989; Nakamura, Medin, & Taraban, 1993); it assumes a class hierarchy (e.g., animal—bird—osprey, as presented in Figure 1 above). Corter and Gluck (1985, 1992) argue that the usefulness of a class is rooted in the ability to predict unobservable attributes (inferential utility) and optimize information processing and transfer (cognitive economy). Corter and Gluck (1992) posit that classes with the highest CU will also be most universal among all humans, since knowing and storing them provides the greatest value. They can therefore also be considered basic. The category utility function is calculated as follows:

$$max\ CU = f(c, F) = P(c) \sum_{k=1}^{m} [P(f_k|c)^2 - P(f_k)^2] \qquad (1)$$

In this formula, some class c is defined by a set of objects o. Each object is characterized by a finite feature (attribute) set $F = \{f_1, f_2, ..., f_m\}$. Consider that with no knowledge about a class membership, f_1 (or a set F) can be predicted using its base-rate probability $P(f_1)$. This probability, in turn, reflects the occurrence of that feature in reality. Such random guesses, will be, on average, correct $P(f_1)$ times, leading to the final probability of correct guessing in the absence of a class as the product of the two probabilities, or $P(f_1)^2$. Extending the same rationale to the probability of guessing a feature under the assumption of a class membership the correct guess will be $P(f_1|c_1)^2$. Thus, the difference between $P(f_1)^2$ and $P(f_1|c_1)^2$ denotes the additional benefit gained from the class membership. This difference, however, needs to be weighted by the probability of a class c_1 occurring, since the guess is made under the condition of c_1 identification.

Category utility ranges between 0 (when predicted frequencies are equal to the base rate) and 1 (if the base-rate frequencies are low while conditional probabilities are high). An interesting property of CU is its relationship to the communication theory by Shannon and Weaver (Shannon, 1948). CU can be considered as the expected reduction of uncertainty due to communication of category information through some cue. The uncertainty is maximal when no category is present and is reduced as the category becomes more "informative"; but this is balanced by the use-frequency of the category. The category utility offers opportunities for computational approaches to conceptual modeling and the automatic discovery of basic-level categories in situations where the required parameters are known or can be estimated for a domain of interest.

Appendix D: Moderator Focus Group Protocol

Each focus group session adheres to the rolling interview presented below to ensure consistency across sessions (Stewart & Shamdasani, 2014). A think-aloud technique was used to collect data (Cotton & Gresty, 2006) while emphasizing participant reflection and discussion during the tasks. The moderator spent time listening to the discussions while allowing the participants to interact with each other.

- Welcome (5 minutes)
    - Greet participants as they arrive
    - Give participants consent form to review and sign prior to participation
    - Ask participants to complete demographics questionnaire

- Describe focus group procedures (5-10 minutes):
    - Describe the objectives of the study
    - Describe the goal of the focus group

- Introduce basic-level categories (BLCs) and basic classes (BCs) (15-20 minutes):
    - Provide a description of BLCs and their use as a motivation for BCs in systems analysis and design
    - Present BC guidelines with BC examples generated from the guidelines

- Task: initial impressions of BCs (10-15) minutes)
    - Introduce two different symptom checker applications (Mayo Clinic and WebMD)
    - For each BC guideline, ask if/how the guideline applies to the applications
    - Discuss whether guidelines are useful in the process of deriving classes for the conceptual model supporting the applications

- Task: designing a wildlife sighting app (30-45 minutes)
    - Provide a sheet with our BC guidelines to participants
    - Describe task (see Appendix E)
    - Discuss the potential classes a mobile app used for citizen scientists (experts and nonexperts) should use in capturing sightings of wildlife
    - Ask participants to discuss how our proposed guidelines could help in modeling such a design (e.g., capture relevant information)

- Closing (10-15 minutes)
    - Are the guidelines useful? Is there any guideline that stands out (or needs improvement) when deriving useful classes?
    - Do you see yourself using these guidelines in the future?
    - Is there anything we missed?

Appendix E: Focus Group Task Details

Initial impression of BCs task: design decisions when developing a symptom checker application

Take a look at the following interfaces shown below. Although the goal of both of the following user interfaces is similar (e.g., find the cause of a set of symptoms), the experience—how the information is organized, is different in both applications.

User Interface 1: Figure E1 presents the symptom checker developed by Mayo Clinic (https://www.mayoclinic.org/symptom-checker/select-symptom/itt-20009075). The symptom checker consists of three steps: (1) Choose a symptom: This step is further divided into two categories (adult symptoms and child symptoms). Since some of these symptoms exist both for adults and children (e.g., abdominal pain), there are some symptoms that are repeated in both lists. (2) Select related factors: Once the user selects a symptom, the second step provides one or more factors that apply to the selected symptom. (3) View possible causes: The symptom checker provides a list of diseases and conditions that match at least one of the factors selected by the user.
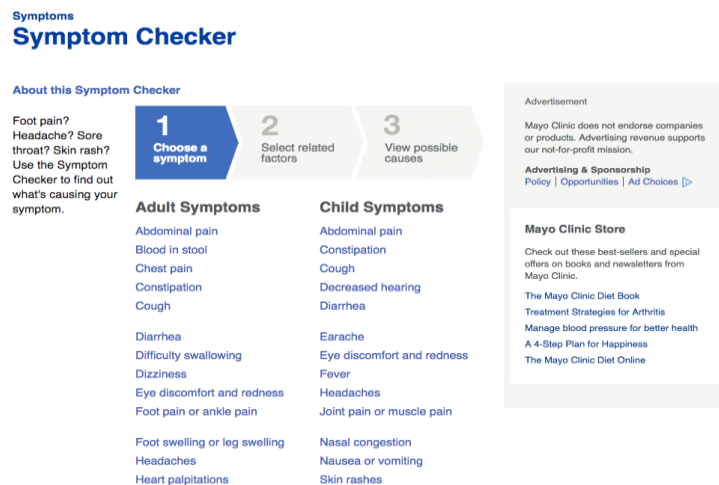

Figure E1: Symptom Checker Splash Screen – Choose a Symptom (Mayo Clinic)

User Interface 2: Figure E2 presents the symptom checker developed by WebMD (https://symptoms.webmd.com/default.htm#introView). The symptom checker has a three-step process similar to the one designed by Mayo Clinic; however, prior to the first step, users are required to provide both gender and age and optionally provide their zip code and email (see Figure E2a below). Based on the gender selected by the user, the first step includes a visual cue of a male or female body. The user selects the part of the body where the symptom originates; the options are gender specific.
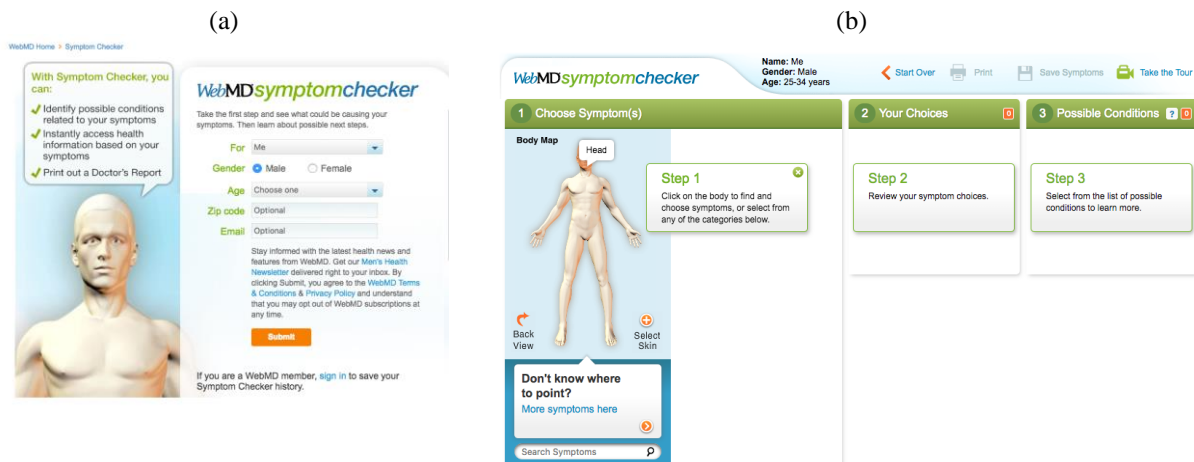
(a)   (b)

Figure E2. Symptom Checker Splash Screen – Choose a Symptom (WebMD)

Questions for discussion:

1. Consider each of the guidelines for identifying basic classes. Do you think the guideline could be applicable to the interfaces above? If yes, how?
2. Do you think the guidelines would be useful in the process of deriving classes for the conceptual model supporting the interfaces?

Focus Group Design Task: designing a wildlife sighting app

The task is to design an app that can be used by people of diverse backgrounds. The goal of the app is to capture information about the wildlife of some region seen by the members of the general public (see examples in Figure E3). Good design avoids creating a user interface that lacks effective organization of sections and data collection processes, which may hinder participation and thus threaten the success of the project.

Figure E3. Common Wildlife Objects Used for Designing a Citizen Science Application (Displayed but Not Distributed to Participants During Focus Group Session)

Questions for discussion:

3. Consider each of the guidelines for identifying basic classes. [For each of the guidelines,] what basic classes can we derive that could be useful for both experts and nonexperts using the app?
4. Which of the guidelines you think is the most useful in deriving classes for the conceptual model supporting our citizen science app?

Subjects participating in these tasks have different domain expertise (e.g., based on the individual's background and experience). We wanted to capture candidate classes from all these individuals and a subgroup of these classes would be useful to both experts and nonexperts. We then asked participants to comment on the usefulness of the guidelines.

Appendix F: Focus Group Results Details

We provide details of the interactions between participants in the two focus groups. Overall, participants valued the utility of using these guidelines in identifying basic-level categories in a given domain. Our experience was that some guidelines were readily understood and subjects were able to apply them naturally to a given context, whereas other guidelines needed further clarification and their application led to some counterintuitive evidence. In Table F1 we summarize elements that provide evidence and counterevidence of the utility of these guidelines.

Table F1. Code Application in Each of the Focus Groups

Guideline name	Evidence [of utility]	Counterevidence [of utility]
G1: Middle Level	Can be readily identified by a nonexpert Allows both experts and nonexperts to contribute	Neglecting middle-level classes could lead to misalignment between the mental model of the user and that of the system.
G2: Entry Category	Entry categories are contextual to the user. Thus, applying G2 brings a diverse set of candidate BCs. It is the role of the analyst to define the goal of the application and choose entry categories that are better aligned to this goal.	From a user perspective, there may be a discrepancy between what a good entry category should be.
G3: Frequently Used Words	Within a domain, BCs organize information efficiently	Depending on the domain and source data, the BC candidates can be large.
G4: Cohesion and Coupling	Helps to identify BCs that are most differentiated from one another	The most differentiated categories may serve as the subject of the chosen BC.
G5: Object Visualization	Mental images of a group of objects can help identify BCs.	May trigger visuals that are less useful compared to BCs
G6: Simplest Words	Selects morphologically simpler candidate BCs	Simpler words may exist in the long tail of a domain
G7: Original Words	Can help further refine candidate BCs (e.g., selecting between two candidate BCs)	Identify BCs that are relatable to individuals, regardless of their backgrounds.
G8: General Predictive Utility	Helps organize information based on likelihood	May be difficult to assess the likelihood of all BCs

Our initial impressions task pinpointed the utility of Guideline 1 (G1), Middle Level. The WebMD application allowed users to select a middle-level body part to arrive at a diagnosis. A middle-level body part lies in the middle of the conceptual hierarchy and can be readily identified by a nonexpert to provide more information about the source of a symptom than a superordinate class such as "entire body." Similarly, a nonexpert user can select a part of the body that is better recognized and known to guide their use of the interface, as opposed to a subordinate one that is overly detailed and possibly unknown e.g., "spleen." The Mayo Clinic symptom checker interface did not provide a basic-class queue, and instead asked participants to select symptoms from a list, which varied in terms of the level of specificity (e.g., blood in stool, lower back pain). A participant of Focus Group 1 (FG1) highlights this by stating:

> *[In] The application with the images, [referring to the WebMD application—See Appendix D] it's a lot easier for like say nonexpert users, like somebody you know I'm just having a symptom, I know where it hurts, I'm going to point it out because when you're interacting with the application you're not talking to somebody that can understand your terminology so **it's easier to pinpoint the places where you're having the symptom [part of the body—a basic class]** where with the other one if you're not familiar with the correct terminology it might lead you to the wrong diagnosis and you don't have a professional there telling you, this I what I'm feeling but it's not exactly here—you pinpoint areas.*

We also saw evidence of utility of G1 in the design task. A participant from FG1 indicated that when you use middle-level categories, both nonexpert and expert users could contribute: *"I know I took a picture of a bird. **I don't know what type of bird it is.** So, it really gets you in the right place for people who are bird experts to then go and contribute what they know is the name of the bird."*

In FG2, participants had an interesting insight regarding G1. Neglecting middle-level classes could lead to poor design, which in turn results in discrepancies in data-entry from non-expert users. For example, in FG2 two participants struggled to differentiate the type of fish (the BC) that was shown to them because they focused below the middle level:

FG2 Participant A said: "*That's a salmon?"* Participant B: *"I think that's sea bass.*"

We saw additional evidence of the importance of the generality of BCs when organizing information for a broader user base—particularly for entry categories (G2). For example, in the training task, participants in FG1 disagreed with the effectiveness of each of the interfaces based on their entry points as first concepts thought of by a user [referring to WebMD's body part entry point and Mayo Clinic's child vs. adult entry point]. A participant in FG1 stated: *"When I see adult/child [referring to Mayo Clinic's interface] to me it doesn't bring much of a difference because we're human beings. **Male or female would be more distinctive**. But I can have a stomach-ache and be an adult and a child could also have a stomach-ache so to me **it wasn't very useful to have the child there** [as an entry-level category]."*

Another participant in FG2 also pointed out that the adult vs. child classification was not optimal given that there are many redundant diagnoses between adults and children (e.g., hay fever is independent of whether the patient is a child or an adult): *"The categories on the list, are similar for the child versus the adult, which is a bit **confusing** because then you have **duplicate items**."*

Notably, a key feature of BCs is their ability to differentiate between objects (i.e., G4); clearly, participants felt the adult vs. child distinction in this setting was not achieving a good differentiation. Yet, there was not full agreement on which of the two websites had the best entry-level category. Entry-level categories appear to be *contextual to the user*. For example, one of the participants had a child and considered the entry category child vs. adult to be a valid one: *"I like the child vs. adult. **Now that I have a kid, I feel that the diagnosis might be different.** I don't know, the kid might be teething versus an adult wouldn't be teething."*

As we further investigated the role of user context for G2, *we saw a variation of opinion on what the correct entry category could be.* For the wildlife application design task, a participant in FG1 stated that an entry category could be the size of the object—i.e., one can classify objects as being small, medium, or large (e.g., small vs. large breeds of dogs)—in reference to how adjectives can serve as descriptors of the entry categories (G2): *"Some people think of dogs as what you want to go get—**people that want a dog as a pet,** they tend to say, **I want a big dog, or I only want a little dog**."*

For both the symptom checker and citizen science application, there were a variety of valid conceptualizations of entry-level BCs based on the participants' personal views. In both cases, we did not provide participants with much detail concerning the goal of the app to encourage creativity. Ultimately, *it will be the role of the analyst to define what the goal of the application is and what entry categories are better aligned to this goal.*

While communicating the guidelines, participants were able to identify ways in which an app can derive words that are used frequently within a context (G3). In the wildlife app design task, a participant in FG3 stated that it was important to know what words are commonly used within the context of interest: *"'Is it a plant?' Because you're in the Everglades. You're trying to think of what could be there, what could be present in that environment."*

This reinforces our notion that BCs are also context-specific to the domain. As the participant above commented, if we were to obtain the ontology of species in the Everglades and plot the observation frequency, there might be a subset of objects that can be identified by citizen scientists and validated by expert users (e.g., biologists). Within a specific context, we seek categories that can organize information in an efficient manner. This supports G4, which states that *basic-level categories are generally the most differentiated*. Providing UGC app users with classes that are highly cohesive and loosely coupled can help a user traverse a knowledge base effectively. The participants of our Focus Groups agreed. For example, one of the participants in FG2 argued in favor of creating categories that are most differentiated from one another: *"Try to group things together that are similar underneath the higher-level category … a fungus is quite a bit different from a flower but a mushroom might have attributes that are similar."*

Although some participants noted the utility of our fifth guideline (G5: Object Visualization) to identify meaningful BCs, there was confusion from some of the participants when applying this guideline. The term "visualization" triggered participants to think of *visual cues* (from a UI/UX view) *rather than* whether the BC triggered *a mental image* (e.g., creating a mental image of a prototypical object such as dog or bird). This misconception could have been triggered by the fact that one of the training tasks had an image of the human body (with its body parts) whereas the other interface did not (see Appendix E). Notwithstanding, some participants understood the value of object visualization. For example, a participant from FG1 stated:

> *For me when I said plant/animal that was like the easiest visualization to distinguish that was like very clear cut. Now if it would all have been animals in the picture, in the slides, and some were birds and some were fish and some were dogs, then the land/air/water would've been more appropriate because it would've all been animals and they would've been different types of animals, so that's kind of the visualization part ... "I see plants, I see animals, I see different kinds of animals, I see different kinds of plants."*

Predictive utility (G8) can further help organize information based on likelihood of an event (e.g., the likelihood of a bird flying is higher than the likelihood of any animal flying). For example, predictive modeling can be used to find patterns and likelihood within ontologies or historical data to provide better associations. The WebMD symptom checker has a bar meter that calculates conditions based on symptoms selected by the user. In this interface, conditions are listed based on likelihoods. The focus group participants argued that *likelihood of events* (as reflected by the bar meter) improved the user experience of users by providing *relevant recommendations*. For instance, a participant in FG1 claimed that the range of plausible diagnoses in the Mayo Clinic symptom checker was too extreme: *"[Once the symptoms are selected] you could just have a mild headache or you could have a brain tumor. The idea of **likelihood** kind of solves your issue in a way as far as putting symptoms of different parts of the body then there are certain likelihoods."*

In the citizen science app, users argued that likelihood could serve as a way to infer objects based on the object's characteristics (features). For example, if a citizen scientist states they saw a white bird with a long neck, long legs, and a yellow beak in the Florida Everglades, a biologist would most likely infer that they saw a great white heron. A participant in FG3 asked whether characteristics (features or attributes) about the object could be entered in the interface: *"Can we **add a feature**? [characteristic] Because that particular bird has a black half. So, any features like you have a dotted face. That'd be more specific."*

Similarly, a participant in FG1 hinted to the idea of *adding feature or attribute context* and made the following statement to represent the same idea of likelihood: *"I found this feather. What is it? It's a feather. Oh, it's got to be a bird. Oh, it's a blue feather. It's like it's likely to be a blue jay or something."*

A recurrent theme in our FGs was the possible **overlap** between BCs derived from different guidelines. For instance, the word *mushroom* is not necessarily simpler (morphologically) than *fungus*, but it is more likely to be learned by children first (G7): *"You would learn about a mushroom before you learned about fungus and [it would] thus be used more frequently [G3]."*

Another participant argued that G6 could help break the fungus/mushroom dichotomy: *"Fungus is not a simple word probably ... you would learn about a mushroom before you learned about fungus."*

Focus group participants found utility in the idea of entry in identifying BCs. In fact, a participant of FG1 indicated she found G1 and G2 to be the most useful:

> *I think that the entry and middle category are probably the most important ones that are interacting with the application because it will target to what you're looking for. And because I want to go to pull something from the Everglades but I don't really care about plants, I'm interested in animals, **I just go directly to the animal section** instead of having to search for all these things at the top level so that you classify—I mean **the entry level is really crucial for an application, entry and middle**. I think it will take you to where the user needs to be to input [or obtain] any information.*

Our participants found G7 useful. A participant in FG3 stated that: *"We're doing a class project right now, and when we start making our categories, I pretend I'm talking to my five-year-old son. If I'm explaining this to Frank [the participant's son], how would I do it to a level that he gets it?"*

Adults are mindful about the kind of language appropriate for use with children (e.g., long names are troublesome for children). Thus, adults do not necessarily provide a child with names that may be typical in the "adult world." It is about *"making things more **relatable** to people, like how to explain things better as well,"* a participant in FG2 emphasized. An individual in FG2 stated: *"Yeah, so just **making things more relatable to people**, like how to explain things better as well. ... In User Interface 2, you can see term abdominal, you could use belly or stomach instead,"* which would thus make it relatable to people.

Another participant in FG2 suggested that both G1 and G4 were related and that identifying the middle-level categories for the wildlife application that are distinct enough from each other can help derive the separation of different groups. Identifying objects in terms of their middle level can also help break these dichotomies: *"The middle level would be the bird, fish, flower, fungus because that has the most diversity I guess, most difference."*

In general, participants agreed that basic-level categories tend to be at a level that is easily relatable to users. A participant in FG2 went beyond our use cases to note that this idea reminds him of some e-commerce sites such as

BestBuy.com: *"It reminds me of like BestBuy.com where you log in and you start seeing some simple categories, computers, TVs, those types of things and then you start drilling down into more specific subcategories, the whole concept of segmentation."* One of the participants in FG2 argued that the idea of classification at a basic (middle) level applies to everyday tasks: *"I'm always thinking, how do you classify things? and it's —even for my own work sometimes I think, how can I organize it so that I can access that information faster."*

A participant in FG3 gave an example that went beyond our task and highlighted how different organizations can leverage on the idea of *entry categories to organize information effectively*. The participant recently had been looking to buy a car and he had a varied experience with different online marketplaces. Different websites had different entry categories:

> *I want a SUV. Everybody gets that. I want a sedan. I want a compact. There are other categories, but you can start there. Then as you start drilling down, at least all the car sites that we have looked at—that happens to be our subject—all of them give you the **ability to filter**—the good ones give you the ability to filter on everything or at least sort on anything that you chose. So, you pick sedan. You pick sedan, you pick your year range, 2015 to 2017, and then from there all of the options along the side give you the **ability to limit what has come up**.*

Frequently used words derived from the interaction with users can also help organize how information is organized and consumed by future users. For example, a participant argued that Google Maps could leverage frequently used item sets to organize information: *"They have so many categories and like they know what are the most-looked-for categories as well when you're on the road, for example. Where is the closest gas station? Where is the closest coffee shop? Things like that that cater towards different things, they don't tell you where's the closest mocha cappuccino."*

As closing remarks in the focus group, users found these guidelines useful and by applying them they were able to derive a diverse set of potential classes. A subset of these classes (i.e., basic classes) was considered superior, being useful for both novice and expert users. Since the users are primed to identify classes by following a set of guidelines, there is a possibility that we did not capture every candidate class. A participant in FG1 stated that *"some guidelines are more relevant than others."* Another participant related the intuitiveness of a guideline to understandability: *"I think that the entry and middle-level category are probably the most important ones … it will take you to where the user needs to be to input or obtain any information."* A participant in FG2 provided an example of utility in a different context—navigating websites: *"This reminds me of BestBuy.com where I start seeing simple categories, computers, TVs, and then you start drilling down into more specific subcategories."* Another participant added *"when you are presented a new thing you have to categorize that thing into a thing that you already know, which category are you going to put that new thing."*

About the Authors

Arturo Castellanos is an assistant professor in the department of information systems and statistics at the Zicklin School of Business, Baruch College (CUNY). His research interests are in the areas of system analysis and design, blockchain, and business analytics. Dr. Castellanos received his MS and PhD degrees in business (information systems) from Florida International University and holds a BS and MS in telecommunications engineering from the University of Navarre (Spain). He currently serves as the vice president for the AIS Special Interest Group in Systems Analysis and Design (SIGSAND).

Monica Chiarini Tremblay is an associate professor of business analytics at the Raymond A. Mason School of Business, William and Mary. Dr. Tremblay's research is focused on business analytics and technology adoption and sustainability in the healthcare context. Some of her publications appear in *Journal of American Medical Informatics, Decision Sciences, Decision Support Systems, European Journal of Information Systems, ACM Journal of Data and Information Quality, Communications of the AIS*. She was the principal investigator on several federal and state grants in the area of Health IT and analytics and was a study session member for the HITR section of Agency Healthcare Research Quality (AHRQ)

Roman Lukyanenko is an associate professor in the Department of Information Technologies at HEC Montreal, Canada. Roman obtained his PhD from Memorial University of Newfoundland under the supervision of Dr. Jeffrey Parsons. Roman's research interests include conceptual modeling, information quality, crowdsourcing, machine learning, design science research, and research methods (research validities, instantiation validity, and artifact sampling). Roman's work has been published in *Nature, MIS Quarterly, Information Systems Research, Journal of the Association for Information Systems, European Journal of Information Systems*, among other outlets. Roman is the current president of the AIS Special Interest Group on Systems Analysis and Design (SIGSAND).

Binny M. Samuel is an assistant professor in the Operations, Business Analytics, and Information Systems department of the Lindner College of Business at the University of Cincinnati, Cincinnati, OH, USA. He earned his PhD from the Kelley School of Business at Indiana University. He also holds a bachelor's of science degree in business and an MBA with concentrations in accounting and information systems. Prior to his doctoral education, he worked in IT roles at Ford Motor Company and at Indiana University. Before joining the University of Cincinnati, he worked at the Ivey Business School at the University of Western Ontario. His work has been published in *ACM Transactions on Management Information Systems, AIS Transactions on Replication Research, Communications of the ACM, European Journal of Information Systems, IEEE Transactions on Software Engineering, Information & Management, Journal of Information Technology*, and *MIS Quarterly*.