On Ordinal Data Science and its role in Socially Acceptable ICT Design

Gerd Stumme

Comparing and ordering things is a basal ability of mankind for organizing its physical and social environment. While many hierarchical relationships can be derived from numerical measures like length or voltage, many others cannot appropriately be captured this way. We argue that the newly emerging field of data science up to now lacks engagement in developing analysis methods for such ordinal data. By the example of an already existing approach in this domain, Formal Concept Analysis, we will discuss its capabilities as a knowledge representation and argue – based on its philosophical foundations – why it is an important building block for socially acceptable IT design.

I. Ordinal Data

Order is a predominant paradigm for perceiving and organizing our physical and social environment, to infer meaning and explanation from observation, and to search and rectify decisions. For instance, we admire the highest mountain on earth, observe pecking order among animals, schedule events in time, and organize our collaborations in hierarchies. The notion of order is deeply embedded in our language, as every adjective gives rise to a comparative (e.g., better, more expensive, more beautiful).

In many cases, entities can be ordered through real-valued valuation functions like size or price. This process of quantification has been boosted by different factors, including *i*) the development of scientific measuring instruments since the scientific revolution, *ii*) the claim that the social sciences should use the same numerical methods which had been successful in natural sciences, and *iii*) nowadays by the instant availability of an enormous range of datasets to almost all aspects of science and everyday life. As real numbers constitute an ordered field, i.e., a field equipped with a linear order, the analysis of such data benefits from the existence of the algebraic operators of fields (the existence of 0 and 1, addition, subtraction, multiplication, division) together with total comparability (i.e., every pair of its elements is comparable) – a combination that allows for various measures of tendency (such as mean, variance, and skewness) and transformations. If more than one real-valued dimension is present, this yields to a real vector space (as Cartesian product of the field of reals), which results in a multitude of additional descriptive measures and metric properties, such as volumes, angles, correlation, covariance. This is the standard setting for the majority of data analysis and machine learning models, and many algorithms (e.g., *k*-means clustering, logistic regression, neural networks, support vector machines, to name just a few) have been developed for these tasks.

However, organizing hierarchical relationships by means of numerical values is not always adequate, as this kind of organization presupposes two important conditions: *i*) every pair of entities has to be comparable, and *ii*) the sizes of differences between numerical values are meaningful and thus comparable. In many situations, however, this is not the case: (*i*) does not hold, for instance, in concept hierarchies ('mankind' is neither a subconcept nor a superconcept of 'ocean') nor in organizations (a member of parliament is neither above nor below a secretary of state); and (*ii*) does not hold, for instance, in school grades (In Germany, is the difference between 1 (very good) and 2 (good) equal to the difference between 4 (sufficient) and 5 (insufficient/fail)?) nor in organizations (In the European Commission, is an advisor closer to a deputy director general than a head of group to a director?).

To address such differences, S. S. Stevens distinguishes four levels of measurement: nominal, ordinal, interval, and ratio.[1] For data on the ratio level (e.g., height), all above-mentioned operations are allowed (division, for instance, provides ratios). Data on the interval level (e.g., temperature measured in Celsius or Fahrenheit) do not have a meaningful zero as point of reference and thus do not allow for ratios, while the comparison of differences is still meaningful. Ordinal data (e.g., the parent relation) only allow for comparisons, and nominal data (e.g., eye color) only for determining equality.[2]

---

1 *Stevens*, On the Theory of Scales of Measurement, Science 103.2684 (1946), 677, https://science.sciencemag.org/content/103/2684/677.full.pdf.

2 Stevens' levels of measurement have been (and still are) heavily disputed. A particularly controversial question that is discussed since Stevens' paper for over 70 years is whether computing the mean of ordinal data is an allowed operation or not, see, e.g., *Lord*, On the Statistical Treatment of Football Numbers, American Psychologist 8.12 (1953), 750 and *Zand Scholten/Borsboom*, A reanalysis of Lord's statistical treatment of football numbers, Journal of Mathematical Psychology 53.2 (2009),

Although there is a large range of preliminary work, Ordinal Data Science is only just emerging as distinct research field. It focusses on data science methods for ordinal data, as they lack the large variety of methods that have been developed for other data types. To this end, we define *ordinal data* as sets of entities ('data points') together with one or more order relations, i.e., binary relations that are reflexive, transitive and anti-reflexive (or variations thereof, such as quasiorders or weak orders).[3] Ordinal data belong thus to the large family of relational data which have received high interest of the computer science community in the last years, due to developments in related fields such as sociology ("relational turn")[4], genetics[5], or epidemiology[6], and socio-technical developments such as the rise of online social networks or knowledge graphs. This means that, for the analysis of ordinal data, one can benefit from all kinds of measures and methods for relational data, for instance, centrality measures and clustering algorithms for (social) network data or inductive logic programming or statistical relational learning from the field of relational data mining.[7] The specific structure of ordinal data, however, allows additionally to tap on the rich

69 as illustrative examples. In practice, this is frequently done (e.g., for aggregating the jury votes in gymnastics or figure skating), while in other cases this is considered bad practice (e.g., for aggregating reviewer judgments in scientific peer reviewing, where often consensus is sought in a discussion phase of the programme committee). Stevens' levels have been subject to several proposals of extensions (e.g., *Chrisman*, Rethinking Levels of Measurement for Cartography, Cartography and Geographic Information Systems 25.4 (1998), 231; *Mosteller*, Data Analysis and Regression: A Second Course in Statistics. Reading, Mass 1977), see *Tal*, Measurement in Science, in: Zalta (Ed.), The Stanford Encyclopedia of Philosophy. Fall 2017 Edition, Stanford 2017 for a systematic survey. All those extensions, however, consider ordinal data as a separate category.

3 The entities may – and in most cases will – have additional attributes on other levels of measurement. The development of hybrid analysis methods is of particular interest here.

4 *Mische*, Relational Sociology, Culture, and Agency, in: Scott/Carrington (Eds.), The SAGE Handbook of Social Network Analysis 2011, London 2011, 80.

5 *Goto/Bono/Ogata/Fujibuchi/Nishioka/Sato/Kanehisa*, Organizing and computing metabolic pathway data in terms of binary relations, Pac Symp Biocomput 1997, 175.

6 *Ciavarella/Fumanelli/Merler/Cattuto/Ajelli*, School closure policies at municipality level for mitigating influenza spread: a model-based evaluation, BMC Infectious Diseases 16.1 (2016), 576.

7 *Dzeroski/Lavrač* (Eds.), Relational Data Mining, Berlin/Heidelberg 2001.

– but up to date mostly unexploited for data science – toolset of mathematical order theory[8] and lattice theory[9].

For one-dimensional ordinal attributes (e.g., agreements in psychological surveys, or judgments in scientific peer reviewing), frequently 5- or 7-point Likert scales are used. For specific applications such as seismic activity or the hardness of minerals, there exist standard ordinal scales.[10] For these, most of the tendency measures for numerical values are not applicable, due to the incomparability of differences and/or the lack of a zero point. Only few tendencies, such as medoids and percentiles, remain. In the case of higher-dimensional ordinal data (e.g., family trees or dependencies in job scheduling tasks), also these tendency measures become meaningless. This is also the case when several ordinal attributes are combined – which is again done by the Cartesian product, yielding a higher-dimensional ordered set. Order theory provides a more appropriate toolset for describing and analyzing such data (e.g., order filters, order intervals, suprema and infima, Pareto optima etc.). Furthermore, specific technical and social processes have been established for dealing with ordinal structures, e.g., scheduling routines for aircraft take-offs, first-in, first-out queuing at bus stops, deriving the succession order as depth-first linear extension of the royal family tree, or discussing only the borderline cases in scientific programme committees. These processes, however, are rather task-specific – there exist only few generic data analysis and machine learning tasks that are particularly tailored for ordinal data. This indicates the need for the development of specific methods for ordinal data.

II. Formal Concept Analysis: A Mathematization of Concepts

A frequently found type of ordinal data arises from boolean/binary attributes, i.e., attributes such as 'has two legs' or 'is serving airports in Africa' which, for a given object, are either true or false: Every single attribute provides a split of the set of entities in two disjoint groups – those entities having the attribute, and those having not. Every such attribute provides a quasiorder (i.e., a reflexive, transitive relation), as those entities of the first group are 'more specific' (in some cases one may also say 'better

8 *Caspard/Leclerc/Monjardet*, Finite Ordered Sets: Concepts, Results and Uses, Cambridge 2012.

9 *Davey/Priestley*, Introduction to Lattices and Order, 2nd ed., Cambridge 2002.

10 https://de.wikipedia.org/wiki/Liste_ordinaler_Skalen.

equipped') than those of the second group. While for one attribute this quasiorder distinguishes two cases only, the number of potential combinations grows exponentially with the number of attributes under consideration: By combining several such attributes by logical conjunction, which is equivalent to computing the Cartesian product as mentioned in Section 1, this yields a rich mathematical structure known as *concept lattice*[11].[12] Concept lattices have been extensively studied in Formal Concept Analysis.[13]

Formal Concept Analysis (FCA) was introduced as a mathematical theory modeling the concept of 'concepts' in terms of lattice theory. The next few paragraphs are meant to be a 'mathematical appetizer' – to give the reader a taste of Formal Concept Analysis. They provide a brief illustration of the core notions.

The basic data structure in FCA is a *(formal) context*:

Definition 2.1. *A* (formal) context *is a triple* $\mathbb{K} := (G,M,I)$, *where G is a set whose elements are called* objects, *M is a set whose elements are called* attributes, *and I is a binary relation between G and M (i.e.,* $I \subseteq G \times M$*).* $(g, m) \in I$ *is read "the object g* has *the attribute m".*

	Latin America	Europe	Canada	Asia Pacific	Middle East	Africa	Mexico	Caribbean	United States
Air Canada	×	×	×	×	×		×	×	×
Air New Zealand		×		×					×
All Nippon Airways		×		×					×
Ansett Australia				×					
The Austrian Airlines Group		×	×	×	×	×			×
British Midland		×							
Lufthansa	×	×	×	×	×	×	×		×
Mexicana	×		×				×	×	×
Scandinavian Airlines	×	×		×		×			×
Singapore Airlines		×	×	×	×	×			×
Thai Airways International	×	×		×				×	×
United Airlines	×	×	×	×			×	×	×
VARIG	×	×		×		×	×		×

Figure 1. A formal context about the destinations of the Star Alliance members.

11 *Wille*, Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts, in: Rival (Ed.), Ordered Sets, Dordrecht/Boston 1982, 445.

12 Other logical operators (in particular Propositional Logic and Description Logics) have also been studied, but conjunction suffices in most applications.

13 *Ganter/Wille*, Formal Concept Analysis: Mathematical Foundations, Heidelberg 1999.

Figure 1 shows a formal context where the object set G comprises all airlines of the Star Alliance group and the attribute set M lists their destinations. The binary relation I is given by the cross table and describes which destinations are served by which Star Alliance member.

Definition 2.2. *For $A \subseteq G$, let*

$$A' := \{m \in M \mid \forall g \in A\colon (g,m) \in I\}$$

and, for $B \subseteq M$, let

$$B' := \{g \in G \mid \forall m \in B\colon (g,m) \in I\}.$$

A (formal) concept *of a formal context (G, M, I) is a pair (A, B) with $A \subseteq G$, $B \subseteq M$, $A' = B$ and $B' = A$. The sets A and B are called the* extent *and the* intent *of the formal concept (A, B), respectively. The subconcept–superconcept relation is formalized by*

$$(A_1, B_1) \leq (A_2, B_2) :\Leftrightarrow A_1 \subseteq A_2 \ (\Leftrightarrow B_1 \supseteq B_2).$$

The set of all formal concepts of a context $\mathbb{K}$ together with the order relation $\leq$ is always a complete lattice,[14] *called the* concept lattice *of $\mathbb{K}$ and denoted by $\underline{\mathfrak{B}}(\mathbb{K})$.*

Figure 2 shows the concept lattice of the context in Figure 1 by a *line diagram*. Line diagrams follow the conventions for the visualization of hierarchical concept systems as established in the German standard DIN 2331.[15] In a line diagram, each node represents a formal concept. A concept $\mathfrak{c}_1$ is a subconcept of a concept $\mathfrak{c}_2$ if and only if there is a path of descending edges from the node representing $\mathfrak{c}_2$ to the node representing $\mathfrak{c}_1$. The name of an object g is always attached to the node representing the smallest concept with g in its extent; dually, the name of an attribute m is always attached to the node representing the largest concept with m in its intent. We can read the context relation from the diagram because an object g has an attribute m if and only if the concept labeled by g is a subconcept of the one labeled by m. The extent of a concept consists of all objects whose labels are attached to subconcepts, and, dually, the intent consists of all attributes attached to superconcepts. For example, the concept labeled by 'Middle East' has {Singapore Airlines, The Austrian Airlines Group,

---

14 I.e., for each subset of concepts, there is always a unique greatest common subconcept and a unique least common superconcept.

15 *Deutsches Institut für Normung*, DIN 2331: Begriffssysteme und ihre Darstellung, 1980.

Lufthansa, Air Canada} as extent, and {Middle East, Canada, United States, Europe, Asia Pacific} as intent.

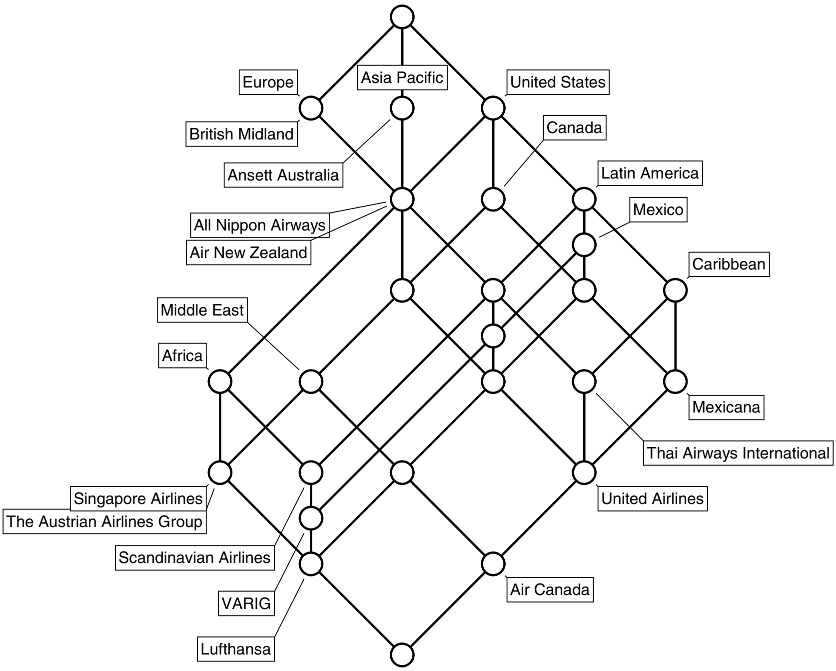

Figure 2. The concept lattice of the context in Figure 1.

In the top of the diagram, we find the destinations which are served by most of the members: Europe, Asia Pacific, and the United States. For instance, besides British Midland and Ansett Australia, all airlines are serving the United States. Those two airlines are located at the top of the diagram, as they serve the fewest destinations – they operate only in Europe and Asia Pacific, respectively.

The further we go down in the concept lattice, the more globally operating are the airlines. The most destinations are served by the airlines at the bottom of the diagram: Lufthansa (serving all destinations besides the Caribbean) and Air Canada (serving all destinations besides Africa). Also, the further we go down in the lattice, the fewer served are the destinations. For instance, Africa, the Middle East, and the Caribbean are served by relatively few Star Alliance members.

Dependencies between the attributes can be described by implications. For $X, Y \subseteq M$, we say that the *implication* $X \rightarrow Y$ *holds* in the context, if each object having all attributes in X also has all attributes in Y. For instance, the implication {Europe, United States} → {Asia Pacific} holds in the Star Alliance context. It can be read directly in the line diagram: the largest concept having both 'Europe' and 'United States' in its intent (i.e., the concept labeled by 'All Nippon Airways' and 'Air New Zealand') also has 'Asia Pacific' in its intent. Similarly, one can detect that the destinations 'Africa' and 'Canada' together imply the destination 'Middle East' (and also 'Europe', 'Asia Pacific', and 'United States').

III. Knowledge Representation with Formal Concept Analysis

The convergence of FCA – whose roots are in mathematics – with computer science demands for a discussion about their relationship. Several aspects of this relationship have already been studied.[16] In this paper we take up the discussion.[17] R. Davis, H. Shrobe, and P. Szolovits studied the question "What is a knowledge representation?".[18] They provided five principles a knowledge representation should follow. We will use these principles to "characterize and make explicit the 'spirit' of [Formal Concept Analysis], the important set of ideas and inspirations that lie behind [...] the concrete machinery used to implement the representation".[19] According to

16 *Hereth/Stumme/Wille/Wille*, Conceptual Knowledge Discovery – a Human-Centered Approach, Journal of Applied Artificial Intelligence (AAI) 17.3 (2003), 281; *Mineau/Stumme/Wille*, Conceptual Structures Represented by Conceptual Graphs and Formal Concept Analysis, in: Tepfenhart/Cyre (Eds.), Conceptual Structures: Standards and Practices, Berlin/Heidelberg/New York 1999, 423; *Stumme/Wille/Wille*, Conceptual Knowledge Discovery in Databases Using Formal Concept Analysis Methods, in: Zytkow/ Quafofou (Eds.), Principles of Data Mining and Knowledge Discovery, Berlin/Heidelberg/New York 1998, 450; Wille/Zickwolff (Eds.), Begriffliche Wissensverarbeitung – Grundfragen und Aufgaben, Mannheim 1994; *Wille*, Why can concept lattices support knowledge discovery in databases?, J. Exp. Theor. Artif. Intell. 14.2-3 (2004), 81; *Zickwolff*, Begriffliche Wissenssysteme in der Künstlichen Intelligenz. FB4–Preprint 1506. TH Darmstadt, 1992.

17 This section has first been published in *Stumme*, Off to New Shores – Conceptual Knowledge Discovery and Processing, Intl. J. Human-Computer Studies (IJHCS) 59.3 (2003), 287.

18 *Davis/Shrobe/Szolovits*, What is a Knowledge Representation, AI Magazine 14.1 (1993), 17.

19 Ibid.

the authors, a knowledge representation is (i) a medium of human expression, (ii) a set of ontological commitments, (iii) a surrogate, (iv) a fragmentary theory of intelligent reasoning, and (v) a medium for pragmatically efficient computation.[20] The authors claim that these principles offer a framework for making explicit the 'spirit' of a representation, and the way it emphasizes one or more of them characterizes the fundamental 'mindset' of the representation. Each knowledge representation formalism is in some way a trade-off between these principles. We will use these five criteria for discussing the role of FCA as knowledge representation method.

It will turn out that the first three principles (especially the first one) have been the driving forces for the development of FCA, while interest on the last two principles – although not completely absent at the beginning (see, for instance, knowledge acquisition with attribute exploration, implicational theories, and efficient computation of concept lattices)[21] – increased during the change of orientation of FCA towards computer science.

1. FCA as a medium of human expression

"Knowledge representations are [...] the medium of expression and communication in which we tell the machine (and perhaps one another) about the world. [...] Knowledge representation is thus a medium of expression and communication for the use by *us*".[22] In other words: "A representation is the language in which we communicate, hence we must be able to speak it without heroic effort".

This observation has always been predominant for the development of the theory of FCA and applications thereof, as the strong emphasis on its philosophical roots shows. When introducing FCA in 1982, R. Wille's intention was to restructure lattice theory: "*Restructuring lattice theory* is understood as an attempt to unfold lattice-theoretical concepts, results, and methods in a continuous relationship with their surroundings [...]. One basic aim is to promote better communication between lattice theorists and potential users of lattice theory".[23] The program of restructuring lat-

20 Davis et al. discuss these principles in the order 3–2–4–5–1. Here we reorder them to follow more closely the historical development of FCA.

21 *Ganter*, Algorithmen zur Formalen Begriffsanalyse, in: Ganter/Wille/Wolff (Eds.), Beiträge zur Begriffsanalyse, Mannheim 1987, 241.

22 *Davis/Shrobe/Szolovits*, AI Magazine 14.1 (1993) (Fn. 18), 17.

23 *Wille*, in: Rival (Ed.) (Fn. 11), 447.

tice theory followed a programmatic discussion about the role of sciences in our society by H. von Hentig.[24] Hentig requests that the sciences "uncover their non-intended aims, declare their intended aims, select and adjust their means according to those aims, discuss openly and understandably their justifications, expectations, and possible consequences, and therefore disseminate their means of research and results in common language".[25] As application, Wille referred to the scientific origin of lattices, namely a model for hierarchies of concepts, which were introduced by Ernst Schröder in the late 19th century and played an important role in attempts to formalize logic.[26] Wille discusses in his visionary article "how parts of arithmetic, structure and representation theory of lattices may be developed out of problems and questions which occur within the analysis of contexts and their hierarchies of concepts".[27]

A second philosophical foundation of FCA is the pragmatic philosophy of Ch. S. Peirce[28] and the Theory of Communicative Action of J. Habermas[29].[30] Peirce considers knowledge as always incomplete, formed and continuously assured by human discourse. J. Habermas took up these ideas in his Theory of Communicative Action where he emphasizes on the importance of the intersubjective community of communication. He observes that humans operate in argumentative dispute on the normative basis of practical-ethical rules. Even in scientific statements (i.e., in assertions), one tries to convince the listener and expects agreement or counterarguments. Hence even in these apparently objective domains the ethical norms of equality and acceptance are present.[31] Following this line of argumentation, the task for theories formalizing aspects of knowledge is thus

---

24 *v. Hentig*, Magier oder Magister? Über die Einheit der Wissenschaft im Verständigungsprozess, Frankfurt 1974.

25 *v. Hentig* (Fn. 24), 136 f., translated by the author.

26 *Schröder*, Algebra der Logik I, II, III. 1890, 1891, 1895, Bristol 2001.

27 *Wille*, in: Rival (Ed.) (Fn. 11), 448.

28 *Peirce*, Collected Papers. Ed. by Hartshorne/Weiss/Burks, Cambridge 1931–1935.

29 *Habermas*, Theorie des kommunikativen Handelns, Frankfurt 1981.

30 Cf. *Wille*, Conceptual Landscapes of Knowledge: A Pragmatic Paradigm for Knowledge Processing, in: Gaul/Locarek-Junge (Eds.), Classification in the Information Age, Berlin/Heidelberg 1999, 344; *Wille*, Plädoyer für eine philosophische Grundlegung der Begrifflichen Wissensverarbeitung, in: Wille/Zickwolff (Eds.), Begriffliche Wissensverarbeitung – Grundfragen und Aufgaben, Mannheim 1994, 11.

31 Cf. *Horster*, Habermas, Jürgen, in: Lutz (Ed.), Metzler Philosophen Lexikon. Von den Vorsokratikern bis zu den Neuen Philosophen, Stuttgart/Weimar 1995, 335, 338.

to provide means for rational communication. The observation that this understanding conflicts with the widely accepted view of mathematics as a means for mechanistic problem solving was certainly one of the main reasons for the change of orientation of FCA towards computer science, where human(–computer) interaction is considered as a research topic on its own (although large parts of computer science also follow a rather mechanistic view).

2. *The ontological commitment of FCA*

Knowledge Representation "is a *set of ontological commitments*, i.e., an answer to the following question: In what terms should I think about the world? [...] In selecting any representation, we are [...] making a set of decisions about how and what to see in the world. [...] We (and our reasoning machines) need guidance in deciding what in the world to attend to and what to ignore".[32]

Formal Concept Analysis formalizes the concepts "concept", "concept extension", "concept intension", and "conceptual hierarchy". We discuss this ontological commitment of FCA along two lines: a definition of concept given in a philosophical lexicon, and the international standard ISO 704.

> **Concept.** A concept is the most basic unit of thought, in contrast to judgment and conclusion, which are forms of thought composed of concepts. While a judgment makes an assertion about an issue, a concept is a notional, i.e., abstract–mental, representation of its 'whatness'; it captures an object based on 'what' it is, without already making an assertion about it. [...] For each concept one distinguishes its *intension* and *extension*. The intension of a concept comprises all attributes thought with it, the extension comprises all objects for which the concept can be predicated. In general, the richer the intension of a concept is, the lesser is its extension, and vice versa.[33]

This lexicon entry reflects a predominant understanding of concepts as being the most basic units of thought, based on which more complex entities of thought – i.e., judgments and conclusions – can be built. This under-

32 *Davis/Shrobe/Szolovits*, AI Magazine 14.1 (1993) (Fn. 18), 17.
33 *Brugger*, Philosophisches Wörterbuch, Freiburg 1976, 39 f., translated by the author.

standing has grown during centuries from Greek philosophy to late Scholastic and has been stated in modern terms in the 17th century in the Logic of Port Royal.[34] It is nowadays established in the standard ISO 704.[35] The definition of formal concepts in FCA follows closely this understanding. It explicitly formalizes extension and intension of a concept, their mutual relationships, and the fact that increasing intent implies decreasing extent and vice versa. Thus, the formalization of concepts by FCA follows a long philosophical tradition.

The standard ISO 704 distinguishes three levels: object level, concept level, and representation level (see Figure 3). There is no immediate relationship between objects and names. This relationship is rather provided by concepts. On the concept level, the objects under discussion constitute the extension of the concept, while their shared properties constitute the intension of the concept. On the representation level, a concept is specified by a definition and is referred to by a name.[36]

While other knowledge representation formalisms like Description Logics or Conceptual Graphs mainly focus on the representation level, the focus of FCA is on the concept level. In fact, the definition of formal concepts follows closely the description of that level in ISO 704:[37] formal concepts consist of extension and intension (only), while concept names and definitions are not within the (core) notions of FCA. Thus, FCA should not be considered as competing with the other mechanisms, but rather as a complement. This view has been followed, for instance, when FCA was combined with Description Logics[38] or with Conceptual Graphs[39].

---

34 *Arnauld/Nicole*, La logique ou l'art de penser – contenant, outre les règles communes, plusieurs observations nouvelles, propres à former le jugement, 1668.

35 *International Organization for Standardization*, ISO 704. Terminology Work – Principles and Methods, 2000.

36 After a discussion of the three levels, ISO 704 provides an overview over naming and definition principles and provides quality criteria for them.

37 *International Organization for Standardization* (Fn. 35).

38 *Baader/Ganter/Sertkaya/Sattler*, Completing Description Logic Knowledge Bases Using Formal Concept Analysis, in: Veloso (Ed.), Proc. IJCAI 2007, Menlo Park 2007, 230; *Prediger/Stumme*, Theory-Driven Logical Scaling, in: Franconi/Kifer (Eds.), Proc. KRDB'99. CEUR Workshop Proc. 21. (Also in: Lambrix/Borgida/Lenzerini/Möller/Patel-Schneider (Eds.), Proc. DL'99. CEUR Workshop Proc. 22), 1999, http://ceur-ws.org/Vol-21; *Prediger*, Logical Scaling in Formal Concept Analysis, in: Lukose/Delugach/Keeler/Searle/Sowa (Eds.), Conceptual Structures: Fulfilling Peirce's Dream, Berlin/Heidelberg/New York 1997; *Stumme*, The Concept Classification of a Terminology Extended by Conjunction and Disjunction,

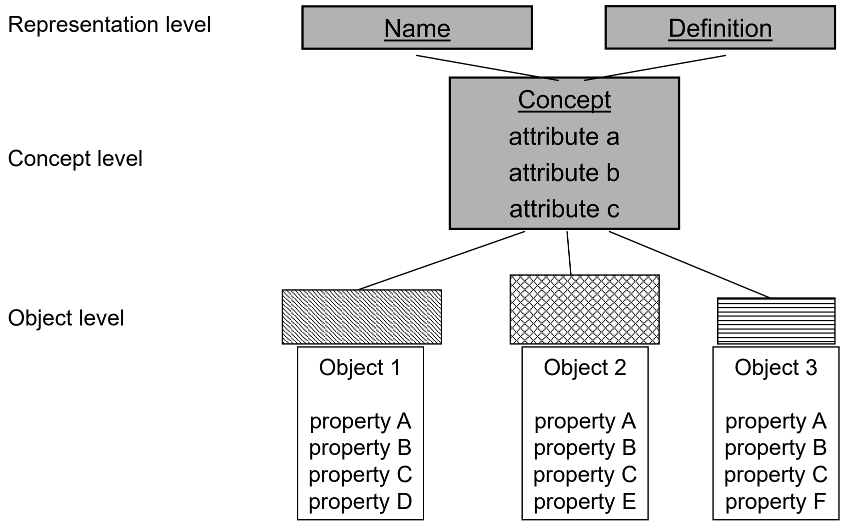

Figure 3. Object level, concept level, and representation level according to ISO 704.

3. Formal contexts and concepts as surrogates

"Knowledge Representation is most fundamentally a *surrogate*, a substitute for the thing itself, used to enable an entity to determine consequences by thinking rather than acting, i.e., by reasoning about the world rather than taking action in it. [...] Reasoning is a process that goes on internally [inside of a person or program], while most things it wishes to reason about exist only externally. [...] This unavoidable dichotomy is a fundamental rationale and role for a representation: it functions as a surrogate inside the reasoner".[40] The authors emphasize that (human or machine) reasoning

in: Foo/Goebel (Eds.), PRICAI'96: Topics in Artificial Intelligence, Berlin/Heidelberg/New York 1996, 121.

39 *Prediger/Wille*, The Lattice of Concept Graphs of a Relationally Scaled Context, in: Tepfenhart/Cyre (Eds.), Conceptual Structures: Standards and Practices, Berlin/Heidelberg/New York 1999, 401; *Wille*, Conceptual Graphs and Formal Concept Analysis, in: Lukose/Delugach/Keeler/Searle/Sowa (Eds.), Conceptual Structures: Fulfilling Peirce's Dream, Berlin/Heidelberg/New York 1997, 290; see also *Mineau/Stumme/Wille*, in: Tepfenhart/Cyre (Eds.) (Fn. 16).

40 *Davis/Shrobe/Szolovits*, AI Magazine 14.1 (1993) (Fn. 18), 17.

cannot deal directly with objects in the world, but only with an internal substitute: the knowledge representation.

The basic surrogates in FCA are formal contexts and concept lattices. The notion of *formal contexts* follows the understanding that one can analyse and argue only in restricted contexts, which are always subject to pre-knowledge and social conventions.[41] In applications, the transition from reality to the formal model (and back) is made explicit by the use of formal contexts; such that this interface between reality and model is always open to argumentation. Also, *formal concepts*, being surrogates, only consider selected aspects of concepts, excluding, for instance, fuzzyness, prototypical concepts, modification over time, and so forth. In order to overcome some of the restrictions, extensions to the formalism were developed, for instance, allowing for fuzzy concepts[42] or more expressive intensional descriptions of concepts[43].

4. FCA as fragmentary theory of intelligent reasoning

Knowledge Representation "is a *fragmentary theory of intelligent reasoning*, expressed in terms of three components: (i) the representation's fundamental conception of intelligent reasoning; (ii) the set of inferences the representation *sanctions*; and (iii) the set of inferences it *recommends*. [...] The initial conception of a representation is typically motivated by some insight indicating how people reason intelligently, or by some belief about what it means to reason intelligently at all".[44] The authors consider five fields which have provided notions of what constitutes intelligent reasoning: mathematical logic (e.g., Prolog), psychology (e.g., frames), biology (e.g., neural networks), statistics (e.g., Bayesian networks), and economics (e.g., rational agents).

As other knowledge representation formalisms, FCA is opposed to the logistic belief that reasoning intelligently necessarily means reasoning in the fashion defined by first-order logic. The roots of FCA are best described in a philosophical view (which is close to what Davis *et al* describe as "psychological view"). It emphasizes on inter-subjective communication

---

41 *Wille*, in: Lukose/Delugach/Keeler/Searle/Sowa (Eds.) (Fn. 39).

42 *Pollandt*, Fuzzy Begriffe: Formale Begriffsanalyse von unscharfen Daten, Berlin/Heidelberg 1997.

43 *Prediger/Stumme*, in: Franconi/Kifer (Eds.) (Fn. 38); *Prediger*, in: Lukose/Delugach/Keeler/Searle/Sowa (Eds.) (Fn. 38).

44 *Davis/Shrobe/Szolovits*, AI Magazine 14.1 (1993) (Fn. 18), 17.

and argumentation, as discussed in Section 3.1. Thus – in contrast to other formalisms – FCA refers the reasoning to the human user who is able to involve common sense, social conventions, views, and purposes. One of the foremost aims of FCA has always been to *support* human thinking, communication, and argumentation rather than *mechanizing* it. Wille discusses the diversity in which intelligent reasoning supported by FCA takes place through sets of real-world applications.[45] Reasoning with concepts comprises, for instance, implicational theories[46], clauses[47], and hypothesis generation[48].

5. *Efficient computation within FCA*

Knowledge Representation "is a *medium for pragmatically efficient computation*, i.e., the computational environment in which thinking is accomplished. One contribution to this pragmatic efficiency is supplied by the guidance a representation provides for organizing information so as to facilitate making the recommended inferences".[49] Davis *et al* stress the importance of having a description of a useful way to organize information which allows for suggesting reasoning mechanisms and for facilitating their execution. Even though automatic reasoning is less in the heart of FCA as it is in most other knowledge representation formalisms, the question how to organize information is important for supporting human reasoning.

In FCA, information is organized in lattices. Lattices provide a clear structure for knowledge representation, which most fundamentally com-

45 *Wille*, in: Gaul/Locarek-Junge (Eds.) (Fn. 30); *Wille*, Bedeutungen von Begriffsverbänden, in: Ganter/Wille/Wolff (Eds.), Beiträge zur Begriffsanalyse, Mannheim 1987, 161.

46 *Ganter*, in: Ganter/Wille/Wolff (Eds.) (Fn. 21); *Stumme/Taouil/Bastide/Pasquier/Lakhal*, Intelligent Structuring and Reducing of Association Rules and with Formal Concept Analysis, in: Baader/Brewker/Eiter (Eds.), KI 2001: Advances in Artificial Intelligence, Heidelberg 2001, 335; *Wild*, Computations with finite closure systems and implications, in: Du/Li (Eds.), Computing and Combinatorics, Berlin/Heidelberg/New York 1995, 111.

47 *Ganter/Wille*, Contextual Attribute Logic, in: Tepfenhart/Cyre (Eds.), Conceptual Structures: Standards and Practices, Berlin/Heidelberg/New York 1999, 377.

48 *Ganter/Kuznetsov*, Formalizing Hypotheses with Concepts, in: *Ganter/Mineau* (Eds.), Conceptual Structures: Logical, Linguistic, and Computational Issues, Berlin/Heidelberg/New York 2000, 342.

49 *Davis/Shrobe/Szolovits*, AI Magazine 14.1 (1993) (Fn. 18), 17.

prises a partial order. Unlike other partial orders (e.g., trees), they allow for multiple inheritance, which often supports a more structured representation and facilitates retrieval of the stored information. Additionally, knowledge representation in lattices is equivalent to apparently unrelated representations such as implications and closure operators. This allows to transfer knowledge into multiple formats each of which is best fit to the actual task. Last but not least, (concept) lattices are equipped with an algebraic structure (stemming from the existence of unique greatest common sub- and least common super-concepts, similar to greatest common divisors and least common multiples for natural numbers) which allows for computation within the lattice structure. As mentioned in Section 3.2, most concept lattice constructions and decompositions have as counterpart a context construction. As formal contexts are only 'logarithmic in size' compared to the concept lattice, they can be seen as a medium of efficient computation.

One can thus exploit the wealth of results of lattice theory for efficient computation. For instance, properties of closure systems are used for computing the concept lattice[50] and valid implications[51], and lattice constructions are also used for the efficient visualization by nested line diagrams[52]. Results from lattice theory have also been exploited for data mining tasks, for instance, for conceptual clustering[53] and for association rule mining[54]. There is still a huge open scientific potential in bringing together structural–mathematical aspects (here especially from FCA) and procedural–computational aspects from computer science.

---

50 *Ganter*, in: Ganter/Wille/Wolff (Eds.) (Fn. 21); *Stumme/Taouil/Bastide/Pasqier/Lakhal*, Computing Iceberg Concept Lattices with TITANIC, Data & Knowledge Engineering 42.2 (2002), 189.

51 *Ganter*, in: Ganter/Wille/Wolff (Eds.) (Fn. 21).

52 *Stumme*, Local Scaling in Conceptual Data Systems, in: Eklund/Ellis/Mann (Eds.), Conceptual Structures: Knowledge Representation as Interlingua, Berlin/Heidelberg/New York 1996, 308; *Wille*, Liniendiagramme hierarchischer Begriffssysteme, in: Bock (Ed.), Anwendungen der Klassifikation: Datenanalyse und numerische Klassifikation, Frankfurt 1984, 32 (English translation: *Wille*, Line diagrams of hierarchical concept systems, *Int. Classif.* 11 (1984), 77).

53 *Hotho*, Clustern mit Hintergrundwissen, Berlin 2004; *Mineau/Godin*, Automatic Structuring of Knowledge Bases by Conceptual Clustering, IEEE Transactions on Knowledge and Data Engineering 7.5 (1995), 824; *Strahringer/Wille*, Conceptual Clustering via Convex–Ordinal Structures, in: Opitz/Lausen/Klar (Eds.), Information and Classification, Heidelberg 1993, 85; *Stumme/Taouil/Bastide/Pasqier/Lakhal*, Data & Knowledge Engineering 42.2 (2002) (Fn. 50).

54 *Stumme/Taouil/Bastide/Pasquier*/Lakhal, in: Baader/Brewker/Eiter (Eds.) (Fn. 46).

Having discussed the adequacy of FCA as a knowledge representation method for computer science, we will study in the next section why and how mathematics-based FCA researchers got attracted by computer science.

IV. Ordinal Data Science for Socially Acceptable ICT Design

Based on the philosophical foundations mentioned in Section 3.1 and integrating several ideas from quite different domains,[55] FCA was introduced in 1979 by R. Wille as a *mathematical* theory in order to "restructure lattice theory", following Hentig's restructuring program (see Section 3.1). A consequence of the aim of restructuring lattice theory was that research in the early time of FCA (1980s and early 1990s) mainly fell into three categories: *i*) lattice theory (e.g., lattice constructions and decompositions)[56], *ii*) qualitative data analysis (e.g., a generalized measurement theory)[57], and *iii*) applications (e.g., the analysis of surveys)[58]. Of course, algorithms for computing concept lattices also were an important topic.[59]

As concepts are the most basic units of thought, it is not surprising that they became important building blocks in Artificial Intelligence (AI) research. Their appearance is prevailing in Knowledge Representation (e.g., in semantic networks, conceptual graphs, description logics), but they also appear, for instance, in Machine Learning (e.g., in conceptual clustering, concept learning). All these approaches focus on other aspects of concepts, leading to different formalizations.

However, until the beginning of the 1990s, the development in AI and in FCA went on almost independently. By then, the mutual perception increased. For instance, FCA researchers got in contact with the knowledge acquisition community, and AI researchers integrated FCA in their ap-

---

55 Barbut/Monjardet (Eds.), L'ordre et la classification. Algèbre et combinatoire, tome II, 1970; *Birkhoff*, Lattice Theory, 3rd ed., New York 1967; *Deutsches Institut für Normung*, DIN 2330: Begriffe und Benennungen – Allgemeine Grundsätze, 1993.

56 *Wille*, Subdirect Decomposition of Concept Lattices, Algebra Universalis 17 (1983), 275.

57 *Ganter/Stahl/Wille*, Conceptual measurement and many–valued contexts, in: Gaul/Schader (Eds.), Classification As a Tool of Research, Amsterdam 1986, 169.

58 *Kollewe*, Evaluation of a Survey with Methods of Formal Concept Analysis, in: Opitz (Ed.), Conceptual and Numerical Analysis of Data, Berlin/Heidelberg 1989, 123.

59 See, for instance, *Ganter*, in: Ganter/Wille/Wolff (Eds.) (Fn. 21).

proaches. This convergence led to the aim of establishing Conceptual Knowledge Processing as an extension of FCA.

In 1993, the ErnstSchröderCenter for Conceptual Knowledge Processing[60] was founded in Darmstadt to support and accompany this development. The agenda of the ErnstSchröderCenter follows the philosophical foundations as described above. It can best be understood by looking on the rich variety of topics of its long-lasting series of quarterly colloquia.[61] Here, we can mention only a few of them: The introductory colloquium was given in 1993 by Karl-Otto Apel on the topic of 'Discoursive Ethics and Semiotics'. A year later already, Artificial Intelligence became a topic, when Joseph Weizenbaum discussed 'Artificial Intelligence as an Ideology'. And only one and a half years later, the relationship between law and IT was addressed, when Alexander Roßnagel, at that time SEL-Stiftungsprofessor at TH Darmstadt, talked about 'Constitutionally Compatible Technology Design'.

Roßnagel's talk provided a long-sighted perspective on the normative role of ethical and social conventions coded in constitution and law for the process of designing information technology. The longevity of this fundamental interrelationship between law and IT design led, eight years later, to the foundation of the Research Center for Information System Design (ITeG) at the University of Kassel by Roßnagel and colleagues. ITeG's trademark is its socio-technical perspective on the design of information technology. In its course, the disciplines contributing to this work have extended far beyond law and computer science; comprising now also human machine interaction, business informatics, gender & diversity, psychology and sociology. In 2014, ITeG has been promoted by the University of Kassel to one of its four scientific centers – a clear indicator that the design of socially acceptable information technology will remain a challenging task for a large range of scientific disciplines also in the future.

60 http://www.ernstschroederzentrum.de.

61 http://www.ernstschroederzentrum.de/arch.html.