# Trends in Interactive Knowledge Discovery for Personalized Medicine: Cognitive Science meets Machine Learning

Andreas Holzinger, *Member, IEEE*

*Abstract*—A grand goal of future medicine is in modelling the complexity of patients to tailor medical decisions, health practices and therapies to the individual patient. This trend towards personalized medicine produces unprecedented amounts of data, and even though the fact that human experts are excellent at pattern recognition in dimensions of $\leq 3$, the problem is that most biomedical data is in dimensions much higher than $3$, making manual analysis difficult and often impossible. Experts in daily medical routine are decreasingly capable of dealing with the complexity of such data. Moreover, they are not interested the data, they need knowledge and insight in order to support their work. Consequently, a big trend in computer science is to provide efficient, useable and useful computational methods, algorithms and tools to discover knowledge and to interactively gain insight into high-dimensional data. A synergistic combination of methodologies of two areas may be of great help here: Human–Computer Interaction (HCI) and Knowledge Discovery/Data Mining (KDD), with the goal of supporting human intelligence with machine learning. A trend in both disciplines is the acquisition and adaptation of representations that support efficient learning. Mapping higher dimensional data into lower dimensions is a major task in HCI, and a concerted effort of computational methods including recent advances from graph-theory and algebraic topology may contribute to finding solutions. Moreover, much biomedical data is sparse, noisy and time-dependent, hence entropy is also amongst promising topics. This paper provides a rough overview of the HCI-KDD approach and focuses on three future trends: graph-based mining, topological data mining and entropy-based data mining.

*Index Terms*—HCI-KDD, interactive knowledge discovery, machine learning, graph-based data mining, topological data mining, entropy-based data mining

## I. Introduction

EXPERTS in the life sciences have to deal with large amounts of complex, high-dimensional, heterogenous, noisy, and weakly structured data sets [1], [2], and large amounts of unstructured information [3].

This "Big Data" [4] in the medical domain is driven by the trend towards precision P4-medicine (Predictive, Preventive, Participatory, Personalized) [5], [6], and has resulted in an explosion in the amount of generated data sets, in particular "-omics" data, for example from genomics, proteomics, metabolomics, epigenetics, transcriptomics, lipidomics, flux-omics, phenomics, microbiomics, etc. [7], [8], [9]. The trend is in moving from a reactive to a proactive medicine and P4-medicine is closely related to systems approaches to disease

A. Holzinger is lead of the research unit HCI-KDD at the Institute for Medical Informatics, Statistics and Documentation, Medical University Graz, Austria, Web: hci-kdd.org e-mail: a.holzinger@hci-kdd.org

and content analytics tools [10], [11]. The well-known challenges with such data include the complexity of feature dimensions (scaling and mapping problems), the heterogeneity of the data (problems of data integration, data fusion), the change over time, and most of all the classic medical data problem: uncertainty of the data quality, false, incomplete data and the danger of modelling artifacts. The often mentioned problem of large amounts of data is rather an advantage with machine learning approaches: Big data actually can provide benefits, as in the biomedical domain, we look often at only a few hundred training examples, so there is the danger of random guessing. Having millions of training samples will raise the precision. The issue of large data sets connects to this question: "What constitutes predictable structures in the world?" as something might be predictable but not comprehensible [12]. Machine learning researchers study algorithms being capable of learning from data and because learning is an important aspect of intelligent behavior, machine learning has become a modern and central aspect of research in artificial intelligence. The most obvious example of learning occur in humans, so there is a natural bridge between research in machine learning and cognitive science, which is strongly related to HCI.

The paradigmatic shift, from classical science, where you first have the question and then collect the data, to data sciences, where you first have the data and then ask questions [13]. The main challenge in this new approach is to ask relevant questions so to find relevant *structural* patterns and/or *temporal* patterns ("knowledge") in such data, because those are often hidden and not directly accessible to the expert [14].

This paper is organized as follows: In section 2 some key terms are briefly explained. In section 3 the basic idea of the HCI-KDD approach is presented, along with the seven research areas involved, however, in the following we concentrate briefly on only three of them: In section 4 on graph-based data mining, in section 5 on topological data mining and in section 6 on entropy-based data mining, concluding by emphasizing that the *combination* of such approaches may bring added values. In the limited space given, such vast topics can only be touched, so the goal of this tutorial is to provide a coarse overview, to motivate and stimulate further research and to encourage to test crazy ideas.

## II. Glossary and Key Terms

- **Algebraic Topology:** is concerned with computations of homologies and homotopies in topological spaces [15].

- **Alpha Shapes:** family of piecewise linear simple curves in the Euclidean plane associated with the shape of a finite set of points [16]; i.e. $\alpha$-shapes are a generalization of the convex hull of a point set: Let $\mathbf{S}$ be a finite set in $\mathbb{R}^3$ and $\alpha$ a real number $0 \leq \alpha \leq \infty$; the u-shape of $\mathbf{S}$ is a polytope that is neither necessarily convex nor necessarily connected. For $\alpha \to \infty$ the $\alpha$-shape is identical to the convex hull of $\mathbf{S}$ [17]; important e.g. in protein-related interactions [18].

- **Betti Number:** can be used to distinguish topological spaces based on the connectivity of $n$-dimensional simplicial complexes: In dimension $k$, the rank of the $k$-th homology group is denoted $\beta_k$, useful in the presence of noisy shapes, because Betti numbers can be used as shape descriptor admitting dissimilarity distances stable under continuous shape deformations [19].

- **Graph mining:** is the application of graph-based methods to structural data sets [20], a survey on graph mining can be found here [21].

- **Homomorphism:** is a function that preserve the operators associated with the specified structure.

- **Homotopy:** Given two maps $f, g : X \to Y$ of topological spaces, $f$ and $g$ are homotopic, $f \simeq g$, if there is a continuous map $H : X \times [0, 1] \to Y$ so that $H(x, 0) = f(x)$ and $H(x, 1) = g(x)$ for all $x \in X$ [22].

- **Homology:** (and cohomology) are algebraic objects associated to a manifold, which give one measure of the number of holes of the object. Computation of the homology groups of topological spaces is a central topic in topology; if the simplicial complex is small, the homology group computations can be done manually; to solve such problems generally a classic algorithm exists [23].

- **Human–Computer Interaction:** study, design and development of the interaction between end users and computers; this classic definition goes back to the work of Alan Newell and Herbert Simon (refs), and HCI research has in the last decades focused almost exclusively on ergonomics of the user interface, while the HCI-KDD approach concentrates almost exclusively on human–data interaction.

- **Information Entropy:** is a measure of the uncertainty in a random variable. This refers to the Shannon entropy, which quantifies the expected value of the information contained in a message.

- **Manifold:** is a fundamental mathematical object which locally resembles a line, a plane, or space.

- **Network:** Synonym for a graph, which can be defined as an ordered or unordered pair $(N, E)$ of a set $N$ of nodes and a set $E$ of edges [24]. Engineers often mention: Data + Graph = Network, or call at least directed graphs as networks; however, in theory, there is no difference between a graph and a network.

- **Pattern discovery:** subsumes a plethora of machine learning methods to detect complex patterns in data sets [25]; applications thereof are, for instance, graph mining [26] and string matching [27].

- **Persistent Homology:** Persistent homology is an algebraic tool for measuring topological features of shapes and functions. It casts the multi-scale organization we frequently observe in nature into a mathematical formalism [28].

- **Simplicial Complex:** is made up of simplices, e.g. a simplicial polytope has simplices as faces and a simplicial complex is a collection of simplices pasted together in any reasonable vertex-to-vertex and edge-to-edge arrangement. A graph is a 1-dim simplicial complex.

- **Small world networks:** are generated based on certain rules with high clustering coefficient [24], [29] but the distances among the vertices are rather short in average, hence they are somewhat similar to random networks and they have been found in several classes of biological networks, see [30].

- **Topological Entropy:** is a nonnegative real number that is a measure of the complexity of a dynamical system [31].

## III. THE HCI-KDD APPROACH

The HCI-KDD approach [32] is a beneficial synergistic combination of methodologies and approaches of two areas that offer ideal conditions towards unraveling some of the "big data" problems mentioned above: Human-Computer Interaction (HCI) and Knowledge Discovery/Data Mining (KDD), with the goal of supporting human intelligence with computational intelligence - by bringing the human into the loop. This approach appreciates both what humans can do best and what computers can do best. A good example for demonstrating the strengths of humans over sophisticated computers is GO, which is a board game from China more than 2,000 years old. It still remains a challenge for computers [33], [34]. Humans are very good at pattern recognition in the low-dimensional space, although humans do not see in three spatial dimensions directly, but via sequences of planar projections. Humans spend a lot of their life time to learn how to infer three-dimensional spatial data from these paired planar projections. Years of practice have tuned a remarkable ability to extract global structures from representations in lower dimension [35]. Kernels in machine learning have a high relevance for understanding issues of generalization and similarity in cognitive science. It is very interesting that most similarity measures considered by psychologists were examples of positive definite kernels, for which a rich body of mathematical theory exists [36]. Consequently, kernel methods can be seen as a unifying theoretical tool showing how several competing and seemingly incommensurate theories in cognitive science (exemplar models versus perceptron models) can be put together [37]. Arguably, the problem of learning represents a gateway to understanding intelligence in both brains and machines, to discovering how the human brain works and to develop intelligent algorithms, which learn from data and improve their competencies - the same as children do [38].

On the other hand, computers can be very beneficial in dealing with high-dimensional data, where we can make use of the benefits of computational topology [39], e.g. by replacing
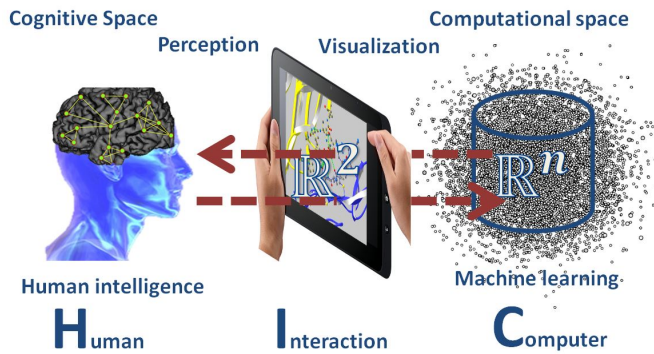
Fig. 1. This image, created originally by A. Holzinger as logo for his group hci-kdd.org, shall emphasize the importance of the manipulating data in the high-dimensional computational space in $\mathbb{R}^n$ and highlights the reality that current devices only allow data visualization in $\mathbb{R}^2$. Consequently, a major challenge for Human–Computer Interaction is to map data from high-dimensional spaces into lower-dimensional spaces.
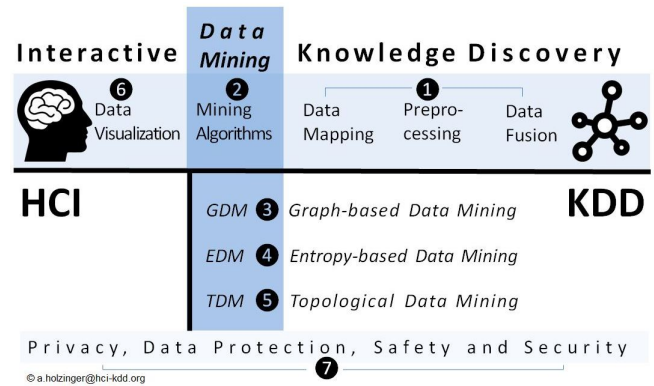


Fig. 2. The big picture of the HCI-KDD approach: KDD encompasses the whole horizontal process chain from data to information and knowledge; actually from physical aspects of raw data, to human aspects including attention, memory, vision, interaction etc. as core topics in HCI, whilst DM as a vertical subject focuses on the development of methods, algorithms and tools for data mining (Image taken from the hci-kdd.org website, as of December, 19, 2014)

a set of point cloud data with a simplicial complex, which converts the data into global topological objects. To combine the most desirable of these formidable talents might highly benefit the knowledge discovery process [32], [40] however, the most critical and most difficult part is in interaction and visualization (see Figure 1).

The original idea of the HCI-KDD [41] approach (Figure 2) is in combining aspects of the best of two worlds: Human–Computer Interaction (HCI), with emphasis on perception, cognition, interaction, reasoning, decision making, human learning and human intelligence, and Knowledge Discovery/Data Mining (KDD), dealing with data processing, computational statistics, artificial intelligence and particularly with integrative machine learning [42]. The most important aspect is the human-in-the-loop approach. Meanwhile it is acknowledged that in many domains computational approaches can not be completely automated - especially in the biomedical domain. The domain knowledge of the expert is of extreme importance and the grand goal is to enable them to interactively manipulate their data, so that they can interactively ask questions to their data sets. An early example for such an approach was given in the medical radiology domain: The clinically useful information in an image typically consists of gray level variations in highly localized regions of the x-ray image and to extract such regions automatically by standard image processing techniques is a hard problem. To bring the physician-in-the-loop means that the expert delineates the pathology bearing regions and a set of anatomical landmarks in the image. To the so marked regions, low-level computer vision tools and image processing algorithms can be applied to extract attributes related to the variations in gray scale [43]. A more recent emphasise of interaction of that kind can be found in [44] and [45].

Whilst interactive knowledge discovery encompasses the horizontal process ranging from physical aspects of data (left in Figure 2) to the human aspects of information processing (right in Figure 2), data mining can be seen vertically and deals specifically with methods, algorithms and tools for finding patterns in the data. In the HCI-KDD approach, seven (the new

magical number 7) essential research areas can be determined as outlined in Figure 2, including: Area 1: Data integration, data fusion and data mapping; Area 2: mining algorithms and Area 6: data visualization [46], [47], [48]. The remainder of this paper focuses on three hot topics, ***Area 3: Graph-based Data Mining (GDM)*** [49], [50], [51], [52]. ***Area 4: Entropy-based Data Mining (EDM)*** [53], [54], and ***Area 5: Topological Data Mining (TDM)*** [55].

In the biomedical domain  as in some other domains  issues of Area 7: privacy, data protection, safety and security are mandatory [56].

## IV. Graph-based Data Mining

Graphs have been used in the life sciences for quite a time and there is a new trend to combine graph theory, machine learning, and statistical data analysis to arrive at a new field, network analysis, to explore complex biomedical graph data. Large-scale generation of genomics, proteomics, metabolomic etc. and signaling data allows the construction of networks that provide a new framework for understanding the molecular basis of physiological and pathological states. Networks and network-based methods have been used in biology to characterize genomic and genetic mechanisms as well as protein signaling; diseases are researched as abnormal perturbations of critical cellular networks. Onset, progression, and intervention in complex diseases including cancer and diabetes can be analyzed today using network approaches. Once the system is represented by a graph = network, methods of graph theory can be applied to find novel insights, important system properties, in structure, time and function. Various statistical and machine learning methods have been developed for this purpose and have already been applied to networks [57], [58]. Graph theory provides powerful tools to map data structures and to find novel connections between single data objects [24], [59]. A mapping of already existing and in medical practice approved *knowledge spaces* as a conceptual graph (as e.g. demonstrated in [50] and a subsequent visual and

graph-theoretical analysis can bring novel insights on hidden patterns in the data, which exactly is the goal of knowledge discovery. Another benefit of a graph-based data structure is in the applicability of methods from network topology and network analysis and data mining, for example the small-world phenomenon [60], [61], and cluster analysis [62], [63].

The first question is "How to get a graph?", or simpler "How to get point sets?", because point cloud data sets (PCD) can be used as primitives for such approaches. The answer to this question is not trivial [64], apart from "naturally available" point clouds, e.g. from laser scanners [65], protein structures [66], or text mapped into a set of points (vectors) in $\mathbb{R}^n$ [67]. Looking at the last example, graphs are intuitively more informative as example words/phrase representations [68], and graphs are the best studied data structures in computer science, with a strong relation to logical languages [69]. The beginning of graph-based data mining approaches was two decades ago, some pioneering work include [70]–[72]. According to [69] there are five theoretical bases of graph-based data mining approaches such as (1) subgraph categories, (2) subgraph isomorphism, (3) graph invariants, (4) mining measures and (5) solution methods. Furthermore, there are five groups of different graph-theoretical approaches for data mining such as (1) greedy search based approach, (2) inductive logic programming based approach, (3) inductive database based approach, (4) mathematical graph theory based approach and (5) kernel function based approach [73]. However, the main disadvantage of graph-theoretical text mining is the computational complexity of the graph representation, consequently the goal of future research in the field of graph-theoretical approaches for text mining is to develop efficient graph mining algorithms which implement effective search strategies and data structures [68].

In [74] a graph-theoretical approach for text mining is used to extract relation information between terms in "free-text" electronic health care records that are semantically or syntactically related. Another field of application is the text analysis of web and social media for detecting influenza-like illnesses [75].

Moreover there can be content-rich relationship networks among biological concepts, genes, proteins and drugs developed with topological text data mining like shown in [76]. According to [77] network medicine describes the clinical application field of topological text mining due to addressing the complexity of human diseases with molecular and phenotypic network maps.

A recent example is PEGASUS, an open source graph mining library, which performs typical graph mining tasks such as computing the diameter of a graph,the radius of each node and finding connected components. PEGASUS is implemented on the HADOOP platform, the open source version of MAPRE-DUCE. Many graph mining operations (Page Rank, spectral clustering, diameter estimation, connected components etc.) are a repeated matrix-vector multiplication; in PEGASUS the authors use a primitive, called generalized iterated matrix-vector multiplication, which is optimized and achieved good performances tested with a Web graph with 6,7 billion edges [78].

## V. Topological Data Mining

Closely related to graph-based methods are topological data mining methods; for both we need point cloud data sets - or at least distances - as input. A set of such primitives forms a space, and if we have finite sets equipped with proximity or similarity measure functions $sim_q \colon S^{q+1} \to [0, 1]$, which measure how "close" or "similar" $(q+1)$-tuples of elements of $S$ are, we speak about a *topological space*. A value of 0 means totally different objects, while 1 corresponds to equivalent items. Interesting are manifolds, which can be seen as a topological space, which is locally homeomorphic (that means it has a continuous function with an inverse function) to a real $n$-dimensional space. In other words: $X$ is a $d$-manifold if every point of $X$ has a neighborhood homeomorphic to $\mathbb{B}^d$; with boundary if every point has a neighborhood homeomorphic to $\mathbb{B}$ or $\mathbb{B}^d_+$ [79].

A topological space may be viewed as an abstraction of a metric space, and similarly, manifolds generalize the connectivity of $d$-dimensional Euclidean spaces $\mathbb{B}^d$ by being locally similar, but globally different. A $d$-dimensional chart at $p \in X$ is a homeomorphism $\phi \colon U \to \mathbb{R}^d$ onto an open subset of $\mathbb{R}^d$, where $U$ is a neighborhood of $p$ and open is defined using the metric. A $d$-dimensional manifold ($d$-manifold) is a topological space $X$ with a $d$-dimensional chart at every point $x \in X$ [80].

For us also interesting are simplicial complexes ("simplicials") which are spaces described in a very particular way, the basis is in Homology. The reason is that it is not possible to represent surfaces precisely in a computer system due to limited computational storage; thus, surfaces are sampled and represented with triangulations. Such a triangulation is called a simplicial complex, and is a combinatorial space that can represent a space. With such simplicial complexes, the topology of a space from its geometry can be separated. Zomorodian [80] compares it with the separation of syntax and semantics in logic.

Topological techniques originated in pure mathematics, but have been adapted to the study and analysis of data during the past two decades. The two most popular topological techniques in the study of data are *homology* and *persistence*. The connectivity of a space is determined by its cycles of different dimensions. These cycles are organized into groups, called homology groups. Given a reasonably explicit description of a space, the homology groups can be computed with linear algebra. Homology groups have a relatively strong discriminative power and a clear meaning, while having low computational cost. In the study of persistent homology the invariants are in the form of persistence diagrams or barcodes [81].

In data mining it is important to extract significant features, and exactly for this, topological methods are useful, since they provide robust and general feature definitions with emphasis on global information, for example Alpha Shapes [17].

A recent example for topological data mining is given by [82]: Topological text mining, which builds on the well-known vector space model, which is a standard approach in text mining [83]: a collection of text documents (corpus) is mapped into points (=vectors) in $\mathbb{R}^n$. Moreover, each word

can be mapped into so-called term vectors, resulting in a very high dimensional vector space. If there are $n$ words extracted from all the documents then each document is mapped to a point (*term vector*) in $\mathbb{R}^\varkappa$ with coordinates corresponding to the weights. This way the whole corpus can be transformed into a point cloud data set. Instead of the Euclidean metric the use of a similarity (proximity) measure is sometimes more convenient; the *cosine similarity measure* is a typical example: the cosine of the angle between two vectors (points in the cloud) reflects how "similar" the underlying weighted combinations of keywords are. Amongst the many different text mining methods (for a recent overview refer to [84]); topological approaches are promising, but need a lot of further research.

Due to finding meaningful topological patterns greater information depth can be achieved from the same data input [85]. However, with increasing complexity of the data to process also the need to find a scalable shape characteristic is greater [86]. Therefore methods of the mathematical field of topology are used for complex data areas like the biomedical field [86], [81]. Topology as the mathematical study of shapes and spaces that are not rigid [86], pose a lot of possibilities for the application in knowledge discovery and data mining, as topology is the study of connectivity information and it deals with qualitative geometric properties [87].

One of the main tasks of applied topology is to find and analyse higher dimensional topological structures in lower dimensional spaces (e.g. point cloud from vector space model as discussed in [85]). A common way to describe topological spaces is to first create simplicial complexes, because a simplicial complex structure on a topological space is an expression of the space as a union of simplices such as points, intervals, triangles, and higher dimensional analogues. Simplicial complexes provide an easy combinatorial way to define certain topological spaces [87]. A simplical complex $K$ is defined as a finite collection of simplices such that $\sigma \in K$ and $\tau$, which is a face of $\sigma$, implies $\tau \in K$, and $\sigma, \sigma' \in K$ implies $\sigma \cap \sigma'$ can either be a face of both $\sigma$ and $\sigma'$ or empty [88]. One way to create a simplicial complex is to examine all subsets of points, and if any subsets of points are close enough, a p-simplex (e.g. line) is added to the complex with those points as vertices. For instance, a Vietoris-Rips complex of diameter $\epsilon$ is defined as $VR(\epsilon) = \sigma | diam(\sigma) \leq \epsilon$, where $diam(\epsilon)$ is defined as the largest distance between two points in $\sigma$ [88]. Figure 2 shows the Vietoris-Rips complex with varying $\epsilon$ for four points with coordinates (0,0), (0,1), (2,1), (2,0). A common way a analyse the topological structure is to use persistent homology, which identifies cluster, holes and voids therein. It is assumed that more robust topological structures are the one which persist with increasing $\epsilon$. For detailed information about persistent homology, it is referred to [88].

### VI. ENTROPY-BASED DATA MINING

In the real medical world, we are confronted not only with complex and high-dimensional data sets, but usually with sparse, noisy, incomplete and uncertain data, where the
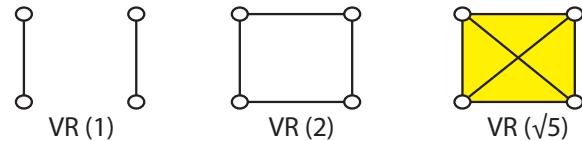


Fig. 3.  Vietoris-Rips complex of four points with varying $\epsilon$ [88].

application of traditional methods of knowledge discovery and data mining always entail the danger of modeling artifacts. Originally, information entropy was introduced by Shannon (1949), as a measure of *uncertainty in the data*. To date, there have emerged many different types of entropy methods with a large number of different purposes and applications. Here we mention only two:

**Graph Entropy** was described by [89] to measure structural information content of graphs, and a different definition, more focused on problems in information and coding theory, was introduced by Körner in [90]. Graph entropy is often used for the characterization of the structure of graph-based systems, e.g. in mathematical biochemistry, but also for any complex network [91]. In these applications the entropy of a graph is interpreted as its structural information content and serves as a complexity measure, and such a measure is associated with an equivalence relation defined on a finite graph; by application of Shannons Eq. 2.4 in [92] with the probability distribution we get a numerical value that serves as an index of the structural feature captured by the equivalence relation [92].

**Topological Entropy** (TopEn), was introduced by [93] with the purpose to introduce the notion of entropy as an invariant for continuous mappings: Let $(X, T)$ be a topological dynamical system, i.e., let $X$ be a nonempty compact Hausdorff space and $T : X \rightarrow X$ a continuous map; the TopEn is a nonnegative number which measures the complexity of the system [94].

Hornero et al. [95] performed a complexity analysis of intracranial pressure dynamics during periods of severe intracranial hypertension. For that purpose they analyzed eleven episodes of intracranial hypertension from seven patients. They measured the changes in the intracranial pressure complexity by applying ApEn, as patients progressed from a state of normal intracranial pressure to intracranial hypertension, and found that a decreased complexity of intracranial pressure coincides with periods of intracranial hypertension in brain injury. Their approach is of particular interest to us, because they proposed classification based on ApEn tendencies instead of absolute values.

Pincus et al. took in [96] heart rate recordings of 45 healthy infants with recordings of an infant one week after an aborted sudden infant death syndrom (SIDS) episode. They then calculated the ApEn of these recordings and found a significant smaller value for the aborted SIDS infant compared to the healthy ones.

Holzinger et al. (2012) [97] experimented with point cloud data sets in the two dimensional space: They developed a model of handwriting, and evaluated the performance of entropy based slant and skew correction, and compared the results to

other methods. This work is the basis for further entropy-based approaches, which are very relevant for advanced entropy-based data mining approaches.

## VII. Conclusion, Open Questions and Future Outlook

Advances in knowledge discovery in complex, high-dimensional data sets need a concerted effort of various topics, ranging from data preprocessing, data fusion, data integration and data mapping to interactive visualization within a low-dimensional space. For this reason, graph-based and topological methods are very useful, since they provide robust and general feature definitions and may support a "global information view". A promising area of future research is in graph-theoretical approaches for text mining, in particular to develop efficient graph mining algorithms which implement robust and efficient search strategies and data structures [68]. Such approaches can be combined with techniques from machine learning, e.g. multi-agents and evolutionary algorithms [98], [99], [49]. However, there remain many open questions, for example about the graph characteristics and the isomorphism complexity [69]. Not only such specific questions are challenging, there are some grand challenges directly involved, e.g. there is much work available on feature selection

As [37] pointed out, there is a large literature on feature selection in machine learning, especially in conjunction with kernel methods, but there are many more methods that could potentially be useful for identifying features, or corresponding similarity measures and in many situations in the real-world a human category learner has to learn the right features (or the right similarity measure), at the same time as he or she learns the categories [100] and machine learning methods can provide hypotheses on how a human learner might achieve this.

It is interesting that much work in cognitive science and machine learning has focused on either supervised or unsupervised learning, i.e. scenarios where either the category labels for all of the stimuli or for none of the stimuli are provided. However, in the real world semi-supervised learning can be beneficial [101].

A definitive challenge when mining high-dimensional data is in measuring distances, e.g. for clustering, outlier detection, similarity measures etc.) as interesting patterns might occur in different subspaces.

A further promising research route is to combine such methods with entropy-based approaches, which have extensively been applied for analyzing sparse and noisy time series data, but so far have not yet been applied to weakly structured data in combination with techniques from computational topology. Consequently, the inclusion of entropy measures for discovery of knowledge in high-dimensional biomedical data is a big future issue, opening a lot of challenging research routes [53].

The grand vision for the future is to effectively support human learning with machine learning. The human brain is an extremely complex organ and can perform many tasks efficiently and effective by (human) learning, particularly when humans are faced with problems that they were faced throughout human evolution (recognizing the Grizzly bear behind you), so we have to keep in mind that our brain can be seen as a statistical decison-making organ, however, only those tasks, which were most important during evolution, are handled most optimal.

The HCI-KDD network of excellence is proactively supporting this vision in bringing together experts with diverse background, but sharing a common goal. A recent output of the network can be found here [102] (for more information please refer to www.hci-kdd.org).

## References

[1] A. Holzinger, M. Dehmer, and I. Jurisica, "Knowledge discovery and interactive data mining in bioinformatics - state-of-the-art, future challenges and research directions," *BMC Bioinformatics*, vol. 15, no. Suppl 6, p. I1, 2014.

[2] A. Holzinger, *Biomedical Informatics: Discovering Knowledge in Big Data*. New York: Springer, 2014.

[3] A. Holzinger, C. Stocker, B. Ofner, G. Prohaska, A. Brabenetz, and R. Hofmann-Wellenhof, *Combining HCI, Natural Language Processing, and Knowledge Discovery - Potential of IBM Content Analytics as an assistive technology in the biomedical domain*. Heidelberg, Berlin, New York: Springer, 2013, pp. 13–24.

[4] X. D. Wu, X. Q. Zhu, G. Q. Wu, and W. Ding, "Data mining with big data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 1, pp. 97–107, 2014.

[5] A. Breso, C. Saez, J. Vicente, F. Larrinaga, M. Robles, and J. M. Garcia-Gomez, "Knowledge-based personal health system to empower outpatients of diabetes mellitus by means of p4 medicine," *Methods in molecular biology (Clifton, N.J.)*, vol. 1246, 2015.

[6] P. B. T. R. P. A. H. Klaus Donsa, Stephan Spat, *Towards personalization of diabetes therapy using computerized decision support and machine learning: some open problems and challenges*. Heidelberg, Berlin: Springer, 2015, pp. 235–260.

[7] B. Huppertz and A. Holzinger, "Biobanks a source of large biological data sets: Open problems and future challenges," in *Interactive Knowledge Discovery and Data Mining: State-of-the-Art and Future Challenges in Biomedical Informatics, Lecture Notes in Computer Science, LNCS 8401*, A. Holzinger and I. Jurisica, Eds. Heidelberg, Berlin: Springer, 2014, pp. 317–330.

[8] J. E. Sackman and M. Kuchenreuther, "Marrying big data with personalized medicine," *Biopharm International*, vol. 27, no. 8, pp. 36–38, 2014, iSI Document Delivery No.: AM9UC Times Cited: 0 Cited Reference Count: 7 Sackman, Jill E. Kuchenreuther, Michael Advanstar communications inc Duluth.

[9] E. D. Perakslis and J. Shon, "Translational informatics in personalized medicine: an update for 2014," *Personalized Medicine*, vol. 11, no. 3, pp. 339–349, 2014.

[10] L. Hood and S. H. Friend, "Predictive, personalized, preventive, participatory (p4) cancer medicine," *Nature Reviews Clinical Oncology*, vol. 8, no. 3, pp. 184–187, 2011.

[11] M. A. Hamburg and F. S. Collins, "The path to personalized medicine," *New England Journal of Medicine*, vol. 363, no. 4, pp. 301–304, 2010.

[12] O. Chapelle, B. Schoelkopf, and A. Zien, *Semi-supervised learning*. Cambridge: MIT press Cambridge, 2006.

[13] C. A. Mattmann, "Computing: A vision for data science," *Nature*, vol. 493, no. 7433, pp. 473–475, 2013.

[14] D. Otasek, C. Pastrello, A. Holzinger, and I. Jurisica, "Visual data mining: Effective exploration of the biological universe," in *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics: State-of-the-Art and Future Challenges. Lecture Notes in Computer Science LNCS 8401*, A. Holzinger and I. Jurisica, Eds. Heidelberg, Berlin: Springer, 2014, p. 1934.

[15] A. Hatcher, *Algebraic Topology*. Cambridge: Cambridge University Press, 2002.

[16] H. Edelsbrunner, D. Kirkpatrick, and R. Seidel, "On the shape of a set of points in the plane," *IEEE Transactions on Information Theory*, vol. 29, no. 4, pp. 551–559, 1983.

[17] H. Edelsbrunner and E. P. Mucke, "3-dimensional alpha-shapes," *ACM Transactions on Graphics*, vol. 13, no. 1, pp. 43–72, 1994.

[18] L. P. Albou, B. Schwarz, O. Poch, J. M. Wurtz, and D. Moras, "Defining and characterizing protein surface using alpha shapes," *Proteins-Structure Function and Bioinformatics*, vol. 76, no. 1, pp. 1–12, 2009.

[19] P. Frosini and C. Landi, "Persistent betti numbers for a noise tolerant shape-based approach to image retrieval," *Pattern Recognition Letters*, vol. 34, no. 8, pp. 863–872, 2013.

[20] D. Cook and L. B. Holder, *Mining Graph Data*. Wiley-Interscience, 2007.

[21] D. Chakrabarti and C. Faloutsos, "Graph mining: Laws, generators, and algorithms," *ACM Computing Surveys (CSUR)*, vol. 38, no. 1, p. 2, 2006.

[22] G. W. Whitehead, *Elements of homotopy theory*. Springer, 1978.

[23] J. R. Munkres, *Elements of algebraic topology*. Addison-Wesley Reading, 1984, vol. 2.

[24] S. Dorogovtsev and J. Mendes, *Evolution of networks: From biological nets to the Internet and WWW*. Oxford University Press, 2003.

[25] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification. Second Edition*. New York et al.: Wiley, 2000.

[26] D. J. Cook and L. B. Holder, "Graph-based data mining," *IEEE Intelligent Systems and their Applications*, vol. 15, no. 2, pp. 32–41, 2000.

[27] D. Gusfield, *Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology*. Cambridge University Press, 1997.

[28] H. Edelsbrunner and J. Harer, *Persistent homology - a survey*, ser. Contemporary Mathematics Series. Providence (RI): Amer Mathematical Soc, 2008, vol. 453, pp. 257–282.

[29] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, pp. 440–442, 1998.

[30] F. Emmert-Streib and M. Dehmer, "Networks for systems biology: Conceptual connection of data and function," *IET Systems Biology*, vol. 5, pp. 185–207, 2011.

[31] D. Koslicki, "Topological entropy of dna sequences," *Bioinformatics*, vol. 27, no. 8, pp. 1061–1067, 2011.

[32] A. Holzinger, "Human–computer interaction & knowledge discovery (hci-kdd): What is the benefit of bringing those two fields to work together?" in *Multidisciplinary Research and Practice for Information Systems, Springer Lecture Notes in Computer Science LNCS 8127*, A. Cuzzocrea, C. Kittl, D. E. Simos, E. Weippl, and L. Xu, Eds. Heidelberg, Berlin, New York: Springer, 2013, pp. 319–328.

[33] D. Lichtenstein and M. Sipser, "Go is polynominal space hard," *Journal of the ACM (JACM)*, vol. 27, no. 2, pp. 393–401, 1980.

[34] C. S. Lee, O. Teytaud, M. H. Wang, and S. J. Yen, "Computational intelligence meets game of go @ ieee wcci 2012," *Ieee Computational Intelligence Magazine*, vol. 7, no. 4, pp. 10–12, 2012.

[35] S. Edelman and R. Shahbazi, "Renewing the respect for similarity," *Frontiers in Computational Neuroscience*, vol. 6, 2012.

[36] B. Schoelkopf, K. Tsuda, and J.-P. Vert, *Kernel methods in computational biology*. Cambridge (MA): The MIT press, 2004.

[37] F. Jkel, B. Schlkopf, and F. A. Wichmann, "Does cognitive science need kernels?" *Trends in cognitive sciences*, vol. 13, no. 9, pp. 381–388, 2009.

[38] T. Poggio and S. Smale, "The mathematics of learning: Dealing with data," *Notices of the AMS*, vol. 50, no. 5, pp. 537–544, 2003.

[39] R. Ghrist, "Barcodes: the persistent topology of data," *Bulletin of the American Mathematical Society*, vol. 45, no. 1, pp. 61–75, 2008.

[40] A. Holzinger, *Extravaganza Tutorial on Hot Ideas for Interactive Knowledge Discovery and Data Mining in Biomedical Informatics*. Heidelberg, Berlin: Springer, 2014, pp. 502–515.

[41] ——, "On knowledge discovery and interactive intelligent visualization of biomedical data - challenges in humancomputer interaction and biomedical informatics," in *DATA 2012*. Rome, Italy: INSTICC, 2012, pp. 9–20.

[42] A. Holzinger and I. Jurisica, "Knowledge discovery and data mining in biomedical informatics: The future is in integrative, interactive machine learning solutions," in *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics: State-of-the-Art and Future Challenges. Lecture Notes in Computer Science LNCS 8401*, A. Holzinger and I. Jurisica, Eds. Heidelberg, Berlin: Springer, 2014, p. in print.

[43] C.-R. Shyu, C. E. Brodley, A. C. Kak, A. Kosaka, A. M. Aisen, and L. S. Broderick, "Assert: A physician-in-the-loop content-based retrieval system for hrct image databases," *Computer Vision and Image Understanding*, vol. 75, no. 12, pp. 111–132, 1999.

[44] C. Bauckhage, M. Hanheide, S. Wrede, T. Kaster, M. Pfeiffer, and G. Sagerer, "Vision systems with the human in the loop," *EURASIP Journal on Advances in Signal Processing*, vol. 2005, no. 14, p. 302161, 2005.

[45] G. Schirner, D. Erdogmus, K. Chowdhury, and T. Padir, "The future of human-in-the-loop cyber-physical systems," *Computer*, vol. 46, no. 1, pp. 36–45, 2013.

[46] A. Holzinger, M. Bruschi, and W. Eder, "On interactive data visualization of physiological low-cost-sensor data with focus on mental stress," in *Multidisciplinary Research and Practice for Information Systems, Springer Lecture Notes in Computer Science LNCS 8127*, A. Cuzzocrea, C. Kittl, D. E. Simos, E. Weippl, and L. Xu, Eds. Heidelberg, Berlin: Springer, 2013, p. 469480.

[47] B. L. W. Wong, K. Xu, and A. Holzinger, "Interactive visualization for information analysis in medical diagnosis," in *Information Quality in e-Health, Lecture Notes in Computer Science, LNCS 7058*, A. Holzinger and K.-M. Simonic, Eds. Springer Berlin Heidelberg, 2011, pp. 109–120.

[48] M. Wiltgen, A. Holzinger, and G. Tilz, *Interactive Analysis and Visualization of Macromolecular Interfaces Between Proteins*. Berlin, Heidelberg, New York: Springer, 2007, pp. 199–212.

[49] M. Preuss, M. Dehmer, S. Pickl, and A. Holzinger, "On terrain coverage optimization by using a network approach for universal graph-based data mining and knowledge discovery," in *Active Media Technology - 10th International Conference, AMT 2014, Warsaw, Poland, August, 11-14, 2014. Proceedings. Lecture Notes in Computer Science LNCS*. Heidelberg, Berlin: Springer, 2014, p. in print.

[50] A. Holzinger, B. Ofner, and M. Dehmer, "Multi-touch graph-based interaction for knowledge discovery on mobile devices: State-of-the-art and future challenges," in *Interactive Knowledge Discovery and Data Mining: State-of-the-Art and Future Challenges in Biomedical Informatics, Springer Lecture Notes in Computer Science LNCS 8401*, A. Holzinger and I. Jurisica, Eds. Berlin, Heidelberg: Springer, 2014, p. 241254.

[51] A. Holzinger, B. Malle, R. Aigner, and N. Giuliani, "On graph extraction from image data," in *Active Media Technology AMT 2014, Lecture Notes in Computer Science LNCS 8610*, D. Slezak, G. Schaefer, T. S. Vuong, and Y.-S. Kim, Eds. Heidelberg, Berlin: Springer, 2014, p. in print.

[52] A. Holzinger, B. Ofner, C. Stocker, A. C. Valdez, A. K. Schaar, M. Ziefle, and M. Dehmer, *On Graph Entropy Measures for Knowledge Discovery from Publication Network Data*. Heidelberg, Berlin: Springer, 2013, pp. 354–362.

[53] A. Holzinger, M. Hortenhuber, C. Mayer, M. Bachler, S. Wassertheurer, A. Pinho, and D. Koslicki, "On entropy-based data mining," in *Interactive Knowledge Discovery and Data Mining: State-of-the-Art and Future Challenges in Biomedical Informatics. Lecture Notes in Computer Science LNCS 8401*, A. Holzinger and I. Jurisica, Eds. Heidelberg, Berlin: Springer, 2014, pp. 209–226.

[54] A. Holzinger, C. Stocker, M. Bruschi, A. Auinger, H. Silva, H. Gamboa, and A. Fred, "On applying approximate entropy to ecg signals for knowledge discovery on the example of big sensor data," in *Active Media Technology, Lecture Notes in Computer Science, LNCS 7669*, R. Huang, A. Ghorbani, G. Pasi, T. Yamaguchi, N. Yen, and B. Jin, Eds., 2012.

[55] A. Holzinger, "On topological data mining," in *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics: State-of-the-Art and Future Challenges. Lecture Notes in Computer Science LNCS 8401*, A. Holzinger and I. Jurisica, Eds. Heidelberg, Berlin: Springer, 2014, pp. 331–356.

[56] P. Kieseberg, H. Hobel, S. Schrittwieser, E. Weippl, and A. Holzinger, "Protecting anonymity in the data-driven medical sciences," in *Interactive Knowledge Discovery and Data Mining: State-of-the-Art and Future Challenges in Biomedical Informatics, Springer Lecture Notes in Computer Science LNCS 8401*, A. Holzinger and I. Jurisica, Eds. Berlin, Heidelberg: Springer, 2014, pp. 303–318.

[57] M. Dehmer and S. C. Basak, *Statistical and Machine Learning Approaches for Network Analysis*. Wiley Online Library, 2012.

[58] M. Dehmer, F. Emmert-Streib, and A. Mehler, *Towards an Information Theory of Complex Networks: Statistical Methods and Applications*. Boston: Birkhaeuser, 2011.

[59] S. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268–276, 2001.

[60] A. L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[61] J. Kleinberg, "Navigation in a small world," *Nature*, vol. 406, no. 6798, pp. 845–845, 2000.

[62] W. Koontz, P. Narendra, and K. Fukunaga, "A graph-theoretic approach to nonparametric cluster analysis," *IEEE Transactions on Computers*, vol. 100, no. 9, pp. 936–944, 1976.

[63] T. Wittkop, D. Emig, A. Truss, M. Albrecht, S. Boecker, and J. Baumbach, "Comprehensive cluster analysis with transitivity clustering," *Nature protocols*, vol. 6, no. 3, pp. 285–295, 2011.

[64] A. Holzinger, B. Malle, M. Bloice, M. Wiltgen, M. Ferri, I. Stanganelli, and R. Hofmann-Wellenhof, "On the generation of point cloud data sets: the first step in the knowledge discovery process," in *Interactive Knowledge Discovery and Data Mining: State-of-the-Art and Future Challenges in Biomedical Informatics, Springer Lecture Notes in Computer Science LNCS 8401*, A. Holzinger and I. Jurisica, Eds. Berlin, Heidelberg: Springer, 2014, pp. 57–80.

[65] K. Khoshelham and S. O. Elberink, "Accuracy and resolution of kinect depth data for indoor mapping applications," *Sensors*, vol. 12, no. 2, pp. 1437–1454, 2012.

[66] A. A. Canutescu, A. A. Shelenkov, and R. L. Dunbrack, "A graph-theory algorithm for rapid protein side-chain prediction," *Protein science*, vol. 12, no. 9, pp. 2001–2014, 2003.

[67] G. Salton, A. Wong, and C. Yang, "A vector space model for automatic indexing," *Communications of the ACM*, vol. 18, no. 11, p. 620, 1975.

[68] C. Jiang, F. Coenen, R. Sanderson, and M. Zito, "Text classification using graph mining-based feature extraction," *Knowledge-Based Systems*, vol. 23, no. 4, pp. 302–308, May 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S095070510900152X

[69] T. Washio and H. Motoda, "State of the art of graph-based data mining," *ACM SIGKDD Explorations Newsletter*, vol. 5, no. 1, p. 59, Jul. 2003. [Online]. Available: http://portal.acm.org/citation.cfm?doid=959242.959249

[70] D. J. Cook and L. B. Holder, "Substructure discovery using minimum description length and background knowledge," *J. Artif. Int. Res.*, vol. 1, no. 1, pp. 231–255, Feb. 1994. [Online]. Available: http://dl.acm.org/citation.cfm?id=1618595.1618605

[71] K. Yoshida, H. Motoda, and N. Indurkhya, "Graph-based induction as a unified learning framework," *Applied Intelligence*, vol. 4, no. 3, pp. 297–316, Jul. 1994. [Online]. Available: http://link.springer.com/10.1007/BF00872095

[72] L. Dehaspe and H. Toivonen, "Discovery of frequent DATALOG patterns," *Data Mining and Knowledge Discovery*, vol. 3, no. 1, pp. 7–36, Mar. 1999. [Online]. Available: http://link.springer.com/article/10.1023/A%3A1009863704807

[73] D. Windridge and M. Bober, "A kernel-based framework for medical big-data analytics," in *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics: State-of-the-Art and Future Challenges. Lecture Notes in Computer Science LNCS 8401*, A. Holzinger and I. Jurisica, Eds. Heidelberg, Berlin: Springer, 2014, pp. 196–207.

[74] X. Zhou, H. Han, I. Chankai, A. Prestrud, and A. Brooks, "Approaches to text mining for clinical medical records," in *Proceedings of the 2006 ACM symposium on Applied computing - SAC '06*. New York, New York, USA: ACM Press, Apr. 2006, p. 235. [Online]. Available: http://dl.acm.org/citation.cfm?id=1141277.1141330

[75] C. D. Corley, D. J. Cook, A. R. Mikler, and K. P. Singh, "Text and structural data mining of influenza mentions in Web and social media." *International journal of environmental research and public health*, vol. 7, no. 2, pp. 596–615, Feb. 2010.

[76] H. Chen and B. M. Sharp, "Content-rich biological network constructed by mining PubMed abstracts." *BMC bioinformatics*, vol. 5, no. 1, p. 147, Oct. 2004. [Online]. Available: http://www.biomedcentral.com/1471-2105/5/147

[77] A. Barabási, N. Gulbahce, and J. Loscalzo, "Network medicine: a network-based approach to human disease," *Nature Reviews Genetics*, vol. 12, no. 1, pp. 56–68, 2011. [Online]. Available: http://www.nature.com/nrg/journal/v12/n1/abs/nrg2918.html

[78] U. Kang, C. E. Tsourakakis, and C. Faloutsos, "Pegasus: mining peta-scale graphs," *Knowledge and Information Systems*, vol. 27, no. 2, pp. 303–325, 2011.

[79] J. W. Cannon, "The recognition problem: what is a topological manifold?" *Bulletin of the American Mathematical Society*, vol. 84, no. 5, pp. 832–866, 1978.

[80] A. Zomorodian, *Computational Topology*, ser. Chapman & Hall/CRC Applied Algorithms and Data Structures series. Boca Raton (FL): Chapman and Hall/CRC, 2010, pp. 1–31, doi:10.1201/9781584888215-c3.

[81] C. Epstein, G. Carlsson, and H. Edelsbrunner, "Topological data analysis," *Inverse Problems*, vol. 27, no. 12, p. 120201, Dec. 2011. [Online]. Available: http://iopscience.iop.org/0266-5611/27/12/120201

[82] H. Wagner and P. Dlotko, "Towards topological analysis of high-dimensional feature spaces," *Computer Vision and Image Understanding*, vol. 121, pp. 21–26, 2014.

[83] M. Kobayashi and M. Aono, "Vector space models for search and cluster mining," in *Survey of Text Mining: Clustering, Classification, and Retrieval*, M. W. Berry, Ed. New York: Springer, 2004, pp. 103–122.

[84] A. Holzinger, J. Schantl, M. Schroettner, C. Seifert, and K. Verspoor, *Biomedical Text Mining: Open Problems and Future Challenges*. Heidelberg, Berlin: Springer, 2014, pp. 271–300.

[85] H. Wagner, P. Dlotko, and M. Mrozek, "Computational topology in text mining," in *Computational Topology in Image Context*, ser. Lecture Notes in Computer Science, M. Ferri, P. Frosini, C. Landi, A. Cerri, and B. Fabio, Eds. Springer Berlin Heidelberg, 2012, vol. 7309, pp. 68–78.

[86] M. Nicolau, A. J. Levine, and G. Carlsson, "Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 17, pp. 7265–70, Apr. 2011. [Online]. Available: http://www.pnas.org/content/108/17/7265.short

[87] G. Carlsson, "Topology and Data," *Bull. Amer. Math. Soc.*, vol. 46, pp. 255–308, 2009. [Online]. Available: http://comptop.stanford.edu/u/preprints/topologyAndData.pdf

[88] X. Zhu, "Persistent homology: An introduction and a new text representation for natural language processing." in *IJCAI*, F. Rossi, Ed. IJCAI/AAAI, 2013.

[89] A. Mowshowitz, "Entropy and the complexity of graphs: I. an index of the relative complexity of a graph," *The bulletin of mathematical biophysics*, vol. 30, no. 1, pp. 175–204, 1968.

[90] J. Körner, "Coding of an information source having ambiguous alphabet and the entropy of graphs," in *6th Prague Conference on Information Theory*, 1973, pp. 411–425.

[91] A. Holzinger, B. Ofner, C. Stocker, A. C. Valdez, A. K. Schaar, M. Ziefle, and M. Dehmer, "On graph entropy measures for knowledge discovery from publication network data," in *Multidisciplinary Research and Practice for Information Systems, Springer Lecture Notes in Computer Science LNCS 8127*, A. Cuzzocrea, C. Kittl, D. E. Simos, E. Weippl, and L. Xu, Eds. Heidelberg, Berlin: Springer, 2013, pp. 354–362.

[92] M. Dehmer and A. Mowshowitz, "A history of graph entropy measures," *Information Sciences*, vol. 181, no. 1, pp. 57–78, 2011.

[93] R. L. Adler, A. G. Konheim, and M. H. McAndrew, "Topological entropy," *Transactions of the American Mathematical Society*, vol. 114, no. 2, pp. 309–319, 1965.

[94] R. Adler, T. Downarowicz, and M. Misiurewicz, "Topological entropy," *Scholarpedia*, vol. 3, no. 2, p. 2200, 2008.

[95] R. Hornero, M. Aboy, D. Abasolo, J. McNames, W. Wakeland, and B. Goldstein, "Complex analysis of intracranial hypertension using approximate entropy." *Crit Care Med*, vol. 34, no. 1, pp. 87–95, 2006.

[96] S. M. Pincus, "Approximate entropy as a measure of system complexity." *Proceedings of the National Academy of Sciences*, vol. 88, no. 6, pp. 2297–2301, 1991.

[97] A. Holzinger, C. Stocker, B. Peischl, and K.-M. Simonic, "On using entropy for enhancing handwriting preprocessing," *Entropy*, vol. 14, no. 11, pp. 2324–2350, 2012.

[98] K. Holzinger, V. Palade, R. Rabadan, and A. Holzinger, "Darwin or lamarck? future challenges in evolutionary algorithms for knowledge discovery and data mining," in *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics: State-of-the-Art and Future Challenges. Lecture Notes in Computer Science LNCS 8401*, A. Holzinger and I. Jurisica, Eds. Heidelberg, Berlin: Springer, 2014, p. in print.

[99] A. Holzinger, D. Blanchard, M. Bloice, K. Holzinger, V. Palade, and R. Rabadan, "Darwin, lamarck, or baldwin: Applying evolutionary algorithms to machine learning techniques," in *The 2014 IEEE/WIC/ACM*

*International Conference on Web Intelligence (WI 2014)*.   IEEE, 2014, pp. 449–453.

[100] P. G. Schyns, R. L. Goldstone, and J.-P. Thibaut, "The development of features in object concepts," *Behavioral and Brain Sciences*, vol. 21, no. 1, pp. 1–17, 1998.

[101] K. Vandist, M. De Schryver, and Y. Rosseel, "Semisupervised category learning: The impact of feedback in learning the information-integration task," *Attention, Perception, and Psychophysics*, vol. 71, no. 2, pp. 328–341, 2009.

[102] A. Holzinger and I. Jurisica, *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics: State-of-the-Art and Future Challenges. Lecture Notes in Computer Science LNCS 8401*.   Heidelberg, Berlin: Springer, 2014.