International Journal on Advances in Intelligent Systems, vol 7 no 3 & 4, year 2014, http://www.iariajournals.org/intelligent_systems/

From Multi-disciplinary Knowledge Objects to Universal Knowledge Dimensions: Creating Computational Views

Claus-Peter Rückemann
Westfälische Wilhelms-Universität Münster (WWU),
Leibniz Universität Hannover,
North-German Supercomputing Alliance (HLRN), Germany
Email: `ruckema@uni-muenster.de`

***Abstract*—Creating and deploying long-term knowledge resources requires research on content as well as on application side. This paper presents the results on high-end structure and classification, being the base for an improved long-term documentation and providing advanced computational views. Content and application context cover factual, conceptual, procedural, and metacognitive knowledge. The core elements are multi-disciplinary knowledge objects, which carry attributes allowing for the creation of various dimensions inside the knowledge resources. The attributes are accompanied by references to flexible multi-lingual universal classifications. The discussion presents the fundamental results having contributed to the knowledge resources. The organisational structure of the knowledge resources supports Big Data integration by a sustainable knowledge definition. The outcome are multi-disciplinary knowledge bases linked with any context, which have led to generate knowledge matrices and which can be used to create computational views. The case studies implemented content and structure as well as workflows for creating knowledge matrices. The implementation deploys dynamical, interactive, and batch computing and storage resources in an Integrated Information and Computing System environment exploiting High End Computing (HEC) and High Performance Computing (HPC) resources and references to the Universal Decimal Classification (UDC). The paper discusses the new practical results from the accompanying long-term projects and implementation.**

Keywords–Knowledge Resources; Computational views; Multi-disciplinary documentation; Advanced Computing; Sustainability; Information Systems; Classification; UDC; Natural Sciences.

I. INTRODUCTION

This article is build on the work performed on advanced computing for the processing of complex knowledge-based information [1]. It has been shown that knowledge processing considerably benefits from employing universal classification.

The demand for long-term sustainability of the resources increases with the complexity of content and context. The organisation and structure of the resources are getting essentially important, the more important the more the data sizes and complexity as well as their intelligent use are required [2]. The article therefore introduces and discusses the background, a comprehensive knowledge definition, including the systematics and methodologies required for an advanced long-term documentation, which can be deployed in most flexible ways. The general requirements have to consider the condition that it is not sufficient to support only an isolated or special methodology. The knowledge requires special qualities in order to be usable as well as the quantities of knowledge counts. A suitable general conceptual handling and a universal knowledge definition is required in this environment for supporting a high quality of resulting context and matrices.

The results are the outcome of the case studies conducted over the last years. The article discusses the fundamental components developed and employed for enabling a systematical processing for arbitrary disciplines. It presents examples for complementary knowledge and sectional views, which can be considered from the knowledge dimensions. Examples for generator workflows and considerations on statistics are the base for several case studies in geosciences and archaeology knowledge. The main new contribution of this research is the implementation of using a universal classification for multi-disciplinary and multi-lingual knowledge classification, for example, for dynamical and computational views.

This paper is organised as follows. Section II introduces with the qualities and quantities, especially with the implemented knowledge resources and organisational structures, data integration, and the concrete practical knowledge definition used. In this context, this section shows how others address challenges of classification, knowledge, and Big Data. Section III discusses the previous work and components employed, Section IV presents the systematics and classification used for the processing, Section V shows the conceptual base for the knowledge dimensions and examples for practical section views. Section VI introduces in the resulting matrix generators and computational consequences. Section VII delivers a discussion of the results from the extended implementation case study. Section VIII shows the results from an external knowledge integration from web resources. Sections IX and X evaluate main results, and summarise the lessons learned, conclusions and future work.

II. KNOWLEDGE QUALITIES AND QUANTITIES

The essential base for any implementation is the creation of suitable content comprising all aspects of knowledge. We have to discuss the consequences of a systematical use, the data itself as well as the basic definition resulting from this work and providing the fundaments for future developments.

A. Knowledge and organisational structures

The systematical use of knowledge resources has led to organising the resources in structures allowing objects and containers. The application of conceptual knowledge on these elements has led to the creation of knowledge dimensions, which can be used with references to the universal classification.

A practical application is the generation of computational views. Example cases studies presented for this extended research are from two different disciplines, volcanological features and meteorite impact structures, whose attributes and features provide criteria for classification, indicators, and multi-disciplinary research. Adding structures, classification, and other qualities to large knowledge resources implicitly increases the computational and storage requirements for the factual knowledge resources. But at the same time the additional conceptual, procedural, and metacognitive knowledge drastically contributes to the optimisation of workflows, to the documentation and integration of knowledge resources, and the quality of applications and results.

These knowledge dimensions effectively reduce the computational and storage requirements for complex and integrated knowledge, supporting really Big Data scenarios. At the same time the requirements for the conceptual knowledge services, the classification, are high regarding width and depth.

For the referenced classification, which has shown up being most important with complex multi-disciplinary long-term classification with practical simple and advanced applications of knowledge resources is the Universal Decimal Classification (UDC) [3]. According to Wikipedia currently about 150,000 institutions, mostly libraries and institutions handling large amounts of data and information, e.g., the ETH Library (Eidgenössische Technische Hochschule), are using basic UDC classification worldwide [4], e.g., with documentation of their resources, library content, bibliographic purposes on publications and references, for digital and realia objects. Regarding the library applications only, UDC is used in more than 144,000 institutions and 130 countries [5]. Further operational areas are author-side content classifications and museum collections.

UDC allows an efficient and effective processing of knowledge data. UDC provides facilities to obtain a universal and systematical view on the classified objects. UDC in combination with statistical methods can be used for analysing knowledge data for many purposes and in a multitude of ways.

With the knowledge resources in this research handling 70,000 classes, for 100,000 objects and several millions of referenced data then simple workflows can be linear but the more complex the algorithms get the workflows will mostly become non-linear. The workflows allow interactive use, dynamical communication, computing, decision support, management, and pre- and postprocessing, e.g., visualisation.

Besides the content there are reasons from the application side, which introduce non-linear behaviour. If going into real application cases, as the ones presented here, the linearity depends on the workflow and data. Any real workflow is subject of limitation, e.g., time, resources, generation of excerpts, and amount of ergonomically practical matrix elements. Creating and exploiting growing knowledge resources regulary goes far beyond the growth of many end-user scenarios. Nevertheless, the growing resources provide excellent means for improving workflows end results. Therefore, the source matrices and the number of references grow much faster than the result matrices, which inevitably make these processes non-linear.

The meaning of knowledge management changed with the extended use, especially when covering the development from “information society” to “knowledge society”. The reasons for the change are resulting from advances in quality, efficiency, precision, consistency and a systematical long-term use, which in combination can contribute reducing digital gaps. In consequence this also promotes new technologies with used resources, e.g., in knowledge management [6] and Library Information Systems (LIS) [7]. The advances in creating new resources pushes the features for integration of resources and for the development of advanced views.

B. Knowledge and Big Data

Many critical reflections on Big Data activities [8] argue that most of the data [9] is unstructured, which is a central issue against an efficient handling of the content. On the other hand, examples of the fundamental aspects of classification in traditional library-focussed context show large benefits of applying even a small extend of knowledge in a well defined environment [10]. The multi-disciplinary aspects from science, economy, society [11] as well as the attribution of “Big Wealth” [12] have triggered new perspectives.

On the one hand, new visions in research [13] and education [14] focus on intelligent and smart environments and components [15]. Not surprisingly, the industry view on the optimisation of “cognition” from Big Data is more or less short term and application centric [16].

On the other hand, the wealth of data implies consequences for anyone handling data and developing applications, which especially results in increased potential [17] and also in challenges [18]. with technologies, methodologies, and systematics.

The overview reveals that there are differences when working with “Big Data” for various academic research on the one hand and use in industry and economy on the other hand. Besides science and industry, assessing knowledge loss risks [19] resulting from departing personnel [20] can be summarised by the risk of knowledge loss, the probability for loss of employees, the consequences of human knowledge loss, and the quality of knowledge resources.

Quality and risks of knowledge loss are correlated with the assessment of management positions [21]. Especially the desintegration of knowledge is hand in hand with the desintegration of workflow and system components. The importance of external auditing for the casting of management and decisions increases with the size of centers and with mission critical services. Exactly the knowledge loss with Big Data then means “big loss of knowledge”, especially as the loss in mostly quality, which can even be worse as the quantity of less quality data is increasing at the same time.

The overall big data challenges, data intensive volume, variability, velocity and for future scenarios especially data vitality, meaning long-term documentation, usability, and accessibility can be handled in a scalable, modular way. The often publicly proposed three "V" for Big Data, namely volume, variability, velocity have been exended by volatility, and veracity in the last years. All of these a primarily describing technical, non-content related issues. For long-term relevant Big Data, especially if the share in structure is significant and even increasing, then the vitality is of critical value.

The approach for coping with Big Data starts with the application of conceptual knowledge. Hence, conceptual knowledge as with an implementation on universal classification provides a "Knowledge as an Infrastructure" solution.

Practical examples are knowledge integration for scientific classification and computation [22], flexible general object envelopes [23], and the management of knowledge and resources, e.g., for environmental information and computation, which has become a focus application [24] in environmental and disaster management. Research conducted on environmental protection and climate change, integrating multi-disciplinary and multi-lingual knowledge resources [25] on an international and trans-national base has shown that sustainable collaboration and governance requires long-term knowledge, classification, and standards as well as management system components.

C. Knowledge definition and understanding

The World Social Science Report 2013 [26] defines knowledge as "The way society and individuals apply meaning to experience ...". Accordingly, the report proposes that "New media and new forms of public participation and greater access to information, are crucial" for open knowledge systems.

In general, we can have an understanding, where knowledge is: Knowledge is created from a subjective combination of different attainments as there are intuition, experience, information, education, decision, power of persuasion and so on, which are selected, compared and balanced against each other, which are transformed and interpreted.

The consequences are: Authentic knowledge therefore does not exist, it always has to be enlived again. Knowledge must not be confused with information or data, which can be stored. Knowledge cannot be stored nor can it simply exist, neither in the Internet, nor in computers, databases, programs or books. Therefore, the demands for knowledge resources in support of the knowledge creation process are complex and multifold.

There is no universal "definition" of the term "knowledge", but UDC provides a good overview of the possible width, depth, and facets. For this research the classification references of UDC:0 (Science and knowledge) define the view on universal knowledge.

1) Big Data: Knowledge Top Level: The question "What is knowledge?" in the conceptual knowledge dimension is best answered by the appropriate classification used with the knowledge resources' application scenarios (Table I). For this case the table shows an excerpt of the knowledge top level classification (Universal Decimal Classification, UDC) used with the knowledge resources.

TABLE I. UNIVERSAL DECIMAL CLASSIFICATION: KNOWLEDGE TOP LEVEL CLASSIFICATION WITH KNOWLEDGE RESOURCES.

UDC Code	*Description (English, excerpt)*
UDC:0	Science and knowledge. Organization. Computer science. ...
UDC:00	Prolegomena. Fundamentals of knowledge and culture ...
UDC:001	Science and knowledge in general. Organization ...
UDC:002	Documentation. Books. Writings. Authorship
UDC:003	Writing systems and scripts
UDC:004	Computer science and technology. Computing
UDC:004.2	Computer architecture
UDC:004.3	Computer hardware
UDC:004.4	Software
UDC:004.5	Human-computer interaction
UDC:004.6	Data
UDC:004.7	Computer communication
UDC:004.8	Artificial intelligence
UDC:004.9	Application-oriented computer-based techniques
UDC:005	Management
UDC:005.1	Management Theory
UDC:005.2	Management agents. Mechanisms. Measures
UDC:005.3	Management activities
UDC:005.5	Management operations. Direction
UDC:005.6	Quality management. Total quality management (TQM)

This classification reflects the conceptual dimension and is intended to be used with the full bandwidth of knowledge and knowledge resources.

2) Knowledge processing and application: Geoscientific knowledge processing is traditionally focussed on processing and analysis of data resulting from geophysical or geological measurements. Examples are processing based on seismological, seismic, magnetic, or gravimetric data. The amount of information and documentation from geosciences and natural sciences based methods and features as well as their complexity has steadily increased for decades. Efficiency and economical practice forces to long-term document and exploit this pool of multi-disciplinary information. Spatial and chronological data and classification are an indispensable component. It is becoming increasingly important that with most professional analysis different geophysical methods and results have to be used in combination.

Common means of application and knowledge discovery, e.g., isolated batch or interactive application scenarios or string based search routines on plain data cannot even approximately integrate the required higher complexity of real environments.

The knowledge gathered during generations should be considered the most valuable component, the more important for long-term results from geosciences. The universal knowledge resources require long-term documentation as well as universal classification and structuring, beyond traditional collections [27], digital libraries [28] and isolated content. With the long-term multi-disciplinary resources the high end processing and computing aspects are essential for sustainability and discovery. Therefore, it is recommended to implement scientific supercomputing resources supporting advanced information systems and creating and improving workflows as recommended [29] with Integrated Information and Computing System (IICS) components [30] and High End Computing (HEC) [31]. This paper presents the results from creating and managing long-term knowledge resources for knowledge processing by employing a universal classification like the Universal Decimal Classification (UDC) [3]. It discusses the experiences handling systematics and classification as well as the methodological use of "Object Carousels". The paper points out the demands and challenges as resulting from the case studies within the GEXI collaborations [32] concentrating on integrating knowledge from geosciences, volcanology, and spatial sciences disciplines.

III. COMPONENTS EMPLOYED

The data used here is based on the content and context from the LX Foundation Scientific Resources [33] and corresponding case studies [34]. The LX structure and UDC [3] with its features [35] are an essential means for the processing workflows and evaluation of the knowledge objects and containers. The applied workflows and processing are based on the data and extended features developed for the Gottfried Wilhelm Leibniz resources [36]. Although shown in detail with previous publications, the following terms may be useful when discussing the knowledge resources and application components.

- Object: An entity of knowledge data being part of knowledge resources. An object can contain any documentation, references, and other data. Objects can have an arbitrary number of sub-objects. Example: Description and location of an archaeological site, locations being part of the location may be handled as sub-objects.
- Container: A collection of knowledge objects in a conjoint format. Example: Volcanological features database.
- Matrix: A subset of the entirety, the "universe", of knowledge. A workflow can consist of many subworkflows each of which can be based on an arbitrary number of knowledge matrices. The output of any subworkflow or workflow can be seen as an intermediate or final result matrix. Example: The output elements of a discovery or search request.
- Qualities: The entirety of documentation including attributes and data being part of knowledge objects.
- Quantities: The number of objects available.
- Systematics: The systematics, a plan based strategy, used for creating knowledge resources for disciplines as well as the systematical use of knowledge resources, e.g., with conceptual knowledge.

The classification, the result of assigning and arranging in classes, is state-of-the-art within the knowledge resources, which implicitly means that the classification is not created statically or even fixed. It can be used and dynamically be modified on the fly, e.g., when required by a knowledge discovery workflow description. Representations and references can be handled dynamically with the context of a discovery process. So, the classification can be dynamically modelled with the context of the workflow.

The context is made up from all the properties and qualities, which have been added to a resource. These properties and qualities can be references, text, bibliographic data, content of publications, links, classification, keywords, factual, conceptual or procedural knowledge and so on.

The granularity must be floating, it depends on the efforts a research community wants to invest for a specific documentation and application. This repeatedly depends on the type of data, the purpose, the actuality and so on. Any of these objects can be evaluated. It depends on the decision of a group or service, which criteria or automatism to take into account.

The LX resources can provide any knowledge documentation and additional information on objects as well as, e.g., geo- and knowledge references. The volcanological data used in the examples is embedded into millions of multi-disciplinary objects, dynamical and spatial information and data files.

The knowledge objects are under continuous development for more than twenty-five years. The classification information has been added in order to describe the objects with the ongoing research and in order to enable more detailed documentation in a multi-disciplinary context. The knowledge resources can make sustainable and vital use of Object Carousels [37] in order to create knowledge object references and to modularise the required algorithms [38]. This provides a universal means for improving coverage, e.g., dark data, and quality within the workflow.

This research focusses on the organisation of knowledge resources and computational views. The general aspects of components and algorithms are focus of different research studies as they mostly belong into the task of the different services and disciplines. The architecture of the components for the purpose of advanced scientific computing and multi-disciplinary documentation is described (starting with Figure 1) in [39]. An example for a workflow, algorithm, and result matrix scenario on knowledge integration for scientific classification and computation is given in [22].

Therefore, for the cases presented in this research paper we had to concentrate on the structure of the knowledge objects, georeferencing, and references data. As secondary components, besides IICS applications and interfaces are available, allowing parallel workflows [40] and intelligent components [41] on HEC and HPC resources. With the IICS, the Generic Mapping Tools (GMT) [42] have been used to visualise georeferenced data wherever a spatial representation is reasonable.

IV. DISCIPLINES, SYSTEMATICS, AND PROCESSING

In geosciences, there is no globally unique stratigraphy. Different continents and regions require different and detailed stratigraphies. Therefore, it is not practicable to have a flat unique global standard due to the regional differences in geological development. Present common stratigraphy concepts [43] fail on general use as well as on a consistent universal classification required. Implementing a universal long-term use we further need to consider appropriate systematics, e.g., lithostratigraphical, chronological, biostratigraphical, chronometrical, chronostratigraphical systematics. For example when it comes to plants, animals, and genotype "-zoic", "-phytic" or "-gen" often mix without distinction. This is the case with many languages, for example, "Mesozoic / Mesophytic" respective "Mesozoikum / Mesophytikum".

Instead of a mix-up of terminology, for systematical use the alignment of the Eonothems / Eons, Erathems / Eras, Systems / Periods, Series / Epochs, and Stages / Ages and so on should be handled consistently and consequently. In addition, the multi-regional dimension should be available for these, showing correspondence with the appropriate absolute ages, as available on-site.

For an efficient and effective processing the knowledge data requires a flexible structure and a universal systematic classification. Any knowledge resources documenting complex multi-disciplinary reality for discovery applications require features for exact documentation on the one hand and they require soft criteria on the other hand.

UDC is a classification complying with the classification criteria. Together with the content, which may deliver more detail or differing perspectives UDC provides a universal view on the classified objects. When requiring facetted classification for multi-disciplinary knowledge the universal UDC cannot be ignored as it is the most comprehensive and flexible means available and supported.

The classification deployed for documentation [44] is able to document any object with any relation, structure, and level of detail as well as intelligently selected nearby hits and references. Objects include any media, textual documents, illustrations, photos, maps, videos, sound recordings, as well as realia, physical objects such as museum objects. UDC is a suitable background classification, for example:

The objects use preliminary classifications for multi-disciplinary content. Standardised operations used with UDC are coordination and addition ("+"), consecutive extension ("/"), relation (":"), order-fixing ("::"), subgrouping ("[]"), non-UDC notation ("*"), alphabetic extension ("A-Z"), besides place, time, nationality, language, form, and characteristics.

V. DIMENSIONS AND CREATING COMPUTATIONAL VIEWS

The dimensions available mostly depend on the features and complexity of the available data. Therefore, a number of essential aspects have been considered when creating content with the knowledge resources. Regarding the views many new arrangements and visualisations are possible.

Nevertheless, it can be quite challenging for application developers to create representations, which can be visualised and to implement suitable components. In many cases so called "Section Views" can be computed, which use n-dimensional sections from 'n+m'-dimensional knowledge context.

A. Implementation of knowledge dimensions

The implemented knowledge resources integrate factual, conceptual, procedural, and metacognitive knowledge. Table II shows the major types of knowledge as complementary parts of the knowledge resources. The table shows some practical examples, which illustrate the benefits of the integration. The data itself is represented in knowledge objects containing any kind of information and collections, including content, classifications, and references.

TABLE II. COMPLEMENTARY KNOWLEDGE IN THE KNOWLEDGE RESOURCES, TYPES AND EXAMPLES.

Knowledge	*Examples*
Factual knowledge	Terminologies Factual details
Conceptual knowledge	Classifications, categorisations Principles, generalisations Theories, models, structures
Procedural knowledge	Algorithms, workflows, skills Methods, techniques Determination on procedures, decision making
Metacognitive knowledge	Strategies Self-knowledge Cognitive tasks, contexts, conditions

With the content and context documentation the knowledge objects describe the integrated knowledge space, for which any dimensions can be interconnected.

B. Section views

The implemented structure and content enable to create section views based on the knowledge dimensions, which can, e.g., be physical or contextual dimensions. Table III shows some section views, which can be based on the combination of contextual and factual knowledge.

TABLE III. SECTION VIEWS BASED ON THE KNOWLEDGE DIMENSIONS. SECTION VIEWS AND EXAMPLES IN PRACTICE.

Section Views	*Examples*
Attributes	colour, size, ... extremes ...
Space and location	spatial distributions geo-references depth distribution
Time	timelines time index
Cultures	context history time location society inventions art
Disciplines	physics geophysics archaeology
Multi-disciplines	geosciences natural sciences – humanities
Multi-lingual	English German Romanic language
Combinations	depth distribution - timelines location-fixed: Objects over time time-fixed: Objects over space/locations culture-fixed: Objects over space and time

Generators can access the knowledge resources and their workflows can apply any appropriate components and algorithms, e.g., classifications, phonetics [45], associations, references, keywords, and statistics.

VI. GENERATORS BASED ON THE DIMENSIONS

The knowledge resources are the base for extended modular matrix generators. Examples are report generators. Nevertheless, report generators can become arbitrarily complex, they are only simple compared with matrix generators in complex environments. A generator can access any data from any requested knowledge resources available. An implementation can integrate different sources as well as applications for various purposes. An example for a simple standard generator can be summarised as

1) Workflow description,
2) Knowledge resources interface,
3) Context selection,
4) Regular expressions,
5) Matrix generation,
6) Result ⇒ Term-based matrix.

The knowledge dimensions are considered in 1), their resources are chosen in 2), the context of the dimensions is selected in 3), and the objects are selected in 4). A simple advanced generator can integrate external core interfaces for extended capabilities:

1) Workflow description,
2) Knowledge resources interface,
3) Context selection,
4) Container interface,
5) Section view selection,
6) LaTeX core interface,
7) Makeindex generation,
8) lxchaidx and further special configuration,
9) Index formatting / index style,
10) Result ⇒ Index-structured matrix.

This example workflow integrates a core interface for the creation of index-structured result matrices. The result matrices contain cross referenced objects integrated from the resources and the selected containers.

The structure of the content is important for the interfaces and for the generation of the results. Besides other factors, the structure and organisation of the content is responsible for the universal deployment of features and therefore an important factor for the quality of result matrices. The workflows can integrate plausibility checks depending on the scenario. These include consideration of structure, content, and knowledge.

Currently, one Central Processing Unit (CPU) can handle sizes of reasonably comprehensive atomic context. A 100,000 object base requires about 1 hour per run per index, including formatting and references. For interactive and dynamical applications where no huge index-structured matrices may be required the object base for an atomic context can be about 100 objects. Any of these scenarios can require parallel instances on available resources instead of consecutive runs before the matrix or index generation is done.

The data herewith can be extended with these steps referencing different sources, creating synopsis matrices, collecting registers, and processing of source data.

Proceeding this way, long-term benefits can accumulate by a sustainable re-use of referenced and source data. Generators based on core knowledge resources, which are being evolved can deliver results-at-a-time. This means that the results can depend on the actual state of the resources (content, context, references and so on).

For a re-use, these results-at-a-time can be preserved. It depends on the scenario, which context data have to be kept for intelligent reports based on the preserved results. For example, the results might be usable with the original sources or snapshot of the objects so that these might be reasonable to be preserved, too.

If knowledge resources are extended this regulary goes ahead with the wish to increase the computing capacities. Contributing conceptual knowledge delivers essential benefits for the knowledge resources but it is also an add-on to the overall data sizes and requirements [46]. This is a classical case for using High End Computing resources, improving results by deploying advanced complex knowledge resources and still reducing computing times.

Even more important is the fact that the use of classification can support the creation of very scalable solutions. The reason for this is that rigid structures are replaced by a matrix of flexible and elastic objects allowing references and a multitude of methods.

The main reasons for the implementation is that the knowledge resources allow

- flexible and dynamical matrix formation, e.g., groups and subgroups and
- to displace the implementation focus from a relatively primitive increase of computational power into the direction of the improvement of structures and workflows.

Example: Adding some of the simplest classification (discipline A and discipline B) to 500 knowledge objects of discipline A and B each adds 1,000 classification entries but at the same time searching for A reduces the number of objects to be handled with their content by half. Even in a slightly more complex workflow the efficiency improvement can easily be in the range of thousands of percents.

Computing result matrices is an arbitrary complex task, which can depend on various factors. Applying statistics and classification to knowledge resources has successfully provided excellent solutions, which can be used for optimising result matrices in context of natural sciences, e.g., geosciences, archaeology, volcanology or with spatial disciplines, as well as for universal knowledge.

The method and application types used for optimisation imply some general characteristics when putting discovery workflows into practice regarding components like terms, media, and other context (Table IV). Regarding the computational views the table lists some representative numbers for Section Views (SV) in addition.

TABLE IV. RESULTING PER-INSTANCE-CALLS FOR METHOD AND APPLICATION TYPES ON OPTIMISATION WITH KNOWLEDGE DISCOVERY.

Type	*Terms*	*Media*	*Workflow*	*Algorithm*	*Combination*
Mean	500	20	20	50,000	3,000
Median	10	5	2	5,000	50
Deviation	30	5	5	200	20
Distribution	90	40	15	20	120
Correlation	15	10	5	20	90
Probability	140	15	20	50	150
Phonetics	50	5	10	20	50
Regular expr.	920	100	50	40	1,500
References	720	120	30	5	900
Association	610	60	10	5	420
UDC	530	120	20	5	660
Keywords	820	100	10	5	600
Translations	245	20	5	5	650
Corrections	60	10	5	5	150
External res.	40	30	5	5	40
SV time	1,100	25	6,000	15	2,400
SV space	850	10	2,500	15	1,800
SV attribute	20	4	70	15	1,200
SV discipline	55	8	5	15	1,500
SV culture	5	4	5	15	760

Statistics methods have shown to be an important means for successfully optimising result matrices. The most widely implemented methods for the creation of result matrices are intermediate result matrices based on regular expressions and intermediate result matrices based on combined regular expressions, classification, and statistics, giving their numbers special weight.

Based on these per-instance numbers this results in demanding requirements for complex applications –

- On numerical data: Millions of calls are done per algorithm and dataset, hundreds in parallel/compact numeric routines.
- On "terms": Hundred thousands of calls are done per sub-workflow, thousands in parallel/complex routines, are done.

Most resources are created for one application scenario or are used for one application scenario only. Only 5–10 percent overlap between disciplines – due to mostly isolated use. Large benefits result from multi-disciplinary multi-lingual integration, which is a major contribution enabling to create SV from resources.

The multi-lingual application adds an additional dimension to the knowledge matrix, which can be used by most discovery processes. As this implemented dimension is of very high quality the matrix space can benefit vastly from content and references.

Section Views show a section through any of the dimensions in the knowledge resources. However, a limited number of possible views is commonly used. The table lists some more regulary used SV based on time, space, attribute, discipline, and culture. Section View reduces the dimensions to a number that can be further used, for example, analysed or visualised.

The interesting aspect is that the means of time and space are most widely used in workflows and combined methods and they are mostly associated with terms, which implies the fact that terms are handled with means like regular expressions. Other types like attributes, discipline, and culture are more or less represented, which results from the fact that the awareness of this context is still not widely spread.

VII. CASE STUDY IMPLEMENTATION AND RESULTS

The following sections discuss the work done for using knowledge resource objects with processing and computing from within IICS. For knowledge resources it is necessary that any classification can be added while the content is developed, over long period of time, more than decades. With the cases presented the content has been created over more than twenty-five years.

- Methodologically, in a first phase, objects have been documented without classification.
- In a second phase, all objects describing volcanic features have been classified as volcanic features.
- In a third phase, volcanic features' objects have been classified into separate classes as required with ongoing extended description of objects in a multi-disciplinary context.

The case study presents a state-of-the-art selection of volcanological and geological features. An evaluation of the association that users have, showed that the criteria "date" and "location" are most prominent with objects if the workflow approaches from the "surface (of the earth)" view [32]. Mapping and timelining with all the respective views will be the natural result.

The small unsorted excerpts of the knowledge resources objects only refer to main UDC-based classes, which for this part of the publication are taken from the Multilingual Universal Decimal Classification Summary (UDCC Publication No. 088) [47] released by the UDC Consortium under the Creative Commons Attribution Share Alike 3.0 license [48] (first release 2009, subsequent update 2012).

As with any object, algorithms, phonetics, regular expressions, statistics, complementary translation or transliteration objects, and other features and discovery processes can be combined for facets and views for any classification subject. Complementary, objects on arbitary algorithms, e.g., processing or statistics, can be included in the knowledge resources and provisioned and applied for further use.

1) Time: Table V shows an excerpt of the resulting UDC classification of spaces of time practically used with the knowledge resources. Instead of the earlier UDC editions the classifications are composite UDC:551.7 mappings referring to historical geology and stratigraphy for all the spaces of time.

TABLE V. UNIVERSAL DECIMAL CLASSIFICATION OF SPACES OF TIME USED WITH THE KNOWLEDGE RESOURCES (EXCERPT).

UDC Code	*Description (English)*
UDC:”0/2”	Dates and ranges of time (CE or AD) ...
UDC:”0”	First millennium CE
UDC:”1”	Second millennium CE
UDC:”2”	Third millennium CE
UDC:”3/7”	Time divisions other than dates in Christian ...
UDC:”3”	Conventional time divisions and subdivisions ...
UDC:”4”	Duration. Time-span. Period. Term. Ages ...
UDC:”5”	Periodicity. Frequency. Recurrence at ...
UDC:”6”	Geological, archaeological and cultural time divisions
UDC:”61/62”	Geological time division
UDC:”63”	Archaeological, prehistoric, protohistoric periods ...
UDC:”67/69”	Time reckonings: universal, secular, non-Christian ...
UDC:”67”	Universal time reckoning. Before Present
UDC:”68”	Secular time reckonings other than universal and ...
UDC:”69”	Dates and time units in non-Christian ...
UDC:”7”	Phenomena in time. Phenomenology of time
UDC:551.7+”61”	Cryptozoic aeon. Precambrian. 600+ MYBP ...
UDC:551.7+”616”	Archaean. Ur-gneiss formation. Ur-schiefer formation
UDC:551.7+”618”	Eozoic. Algonkian
UDC:551.7+”62”	Phanerozoic aeon. 600 MYBP - Present
UDC:551.7+”621”	Palaeozoic. 600-180 MYBP
UDC:551.7+”621.2”	Cambrian. 600-490 MYBP
UDC:551.7+”621.3”	Ordovician. 490-430 MYBP
UDC:551.7+”621.4”	Silurian. Gothlandian. 430-400 MYBP
UDC:551.7+”621.5”	Devonian. 400-350 MYBP
UDC:551.7+”621.6”	Carboniferous. 350-270 MYBP
UDC:551.7+”621.7”	Permian. 270-220 MYBP
UDC:551.7+”622.2”	Triassic. 220-180 MYBP
UDC:551.7+”622.4”	Jurassic. 180-135 MYBP
UDC:551.7+”622.6”	Cretaceous. 135-70 MYBP
UDC:551.7+”628”	Cenozoic (Cainozoic). Neozoic
UDC:551.7+”628”	Tertiary. 70-1 MYBP
UDC:551.7+”628.2”	Palaeogenic. Nummulitic
UDC:551.7+”628.22”	Palaeocene
UDC:551.7+”628.24”	Eocene
UDC:551.7+”628.26”	Oligocene
UDC:551.7+”628.4”	Neogene
UDC:551.7+”628.42”	Miocene
UDC:551.7+”628.44”	Pliocene
UDC:551.7+”628.6”	Quaternary. 1 MYBP - Present
UDC:551.7+”628.62”	Pleistocene in general. Diluvium
UDC:551.7+”628.64”	Holocene. Postglacial in general

2) Space: Table VI shows an excerpt of the resulting UDC classification practically used for spatial features and place.

TABLE VI. UNIVERSAL DECIMAL CLASSIFICATION OF SPATIAL FEATURES AND PLACE USED WITH THE KNOWLEDGE RESOURCES (EXCERPT).

UDC Code	*Description (English)*
UDC:(1)	Place and space in general. Localization. Orientation
UDC:(1-0/-9)	Special auxiliary subdivision for boundaries and spatial ...
UDC:(1-0)	Zones
UDC:(1-1)	Orientation. Points of the compass. Relative position
UDC:(1-19)	Relative location, direction and orientation
UDC:(1-2)	Lowest administrative units. Localities
UDC:(1-5)	Dependent or semi-dependent territories
UDC:(1-6)	States or groupings of states from various points of view
UDC:(1-7)	Places and areas according to privacy, publicness ...
UDC:(1-8)	Location. Source. Transit. Destination
UDC:(1-9)	Regionalization according to specialized points of view
UDC:(2)	Physiographic designation
UDC:(20)	Ecosphere
UDC:(21)	Surface of the Earth in general. Land areas in particular. ...
UDC:(23)	Above sea level. Surface relief. Above ground generally. ...
UDC:(24)	Below sea level. Underground. Subterranean
UDC:(25)	Natural flat ground (at, above or below sea level). ...
UDC:(26)	Oceans, seas and interconnections
UDC:(28)	Inland waters
UDC:(29)	The world according to physiographic features
UDC:(3)	Places of the ancient and mediaeval world
UDC:(32)	Ancient Egypt
UDC:(36)	Regions of the so-called barbarians
UDC:(37)	Italia. Ancient Rome and Italy
UDC:(38)	Ancient Greece
UDC:(4/9)	Countries and places of the modern world

Classification references of that kind have been implemented with knowledge resources and geo-coordinates.

Any of the classification can be mapped to specific content data. The workflows and processing handle different dates and specification between classification and content as well as using equal classification elements for different absolute dates, e.g., as required for different regions or cultures.

3) Creation of knowledge objects: The following examples are taken from astrophysical, volcanological, and geoscientific context, all referring to a large number of natural sciences' disciplines and humanities. Especially, types of meteorites and types of volcanoes do have different origin but the context links them in arbitrary ways in space and time. Examples are their linkage in geological time scale, location, geological attributes as well as their association with cultural events and secondary use of their material in many objects or buildings. An object carousel generated for impact craters, shows the different types present in the knowledge resources groups and their crater categories (Figure 1).

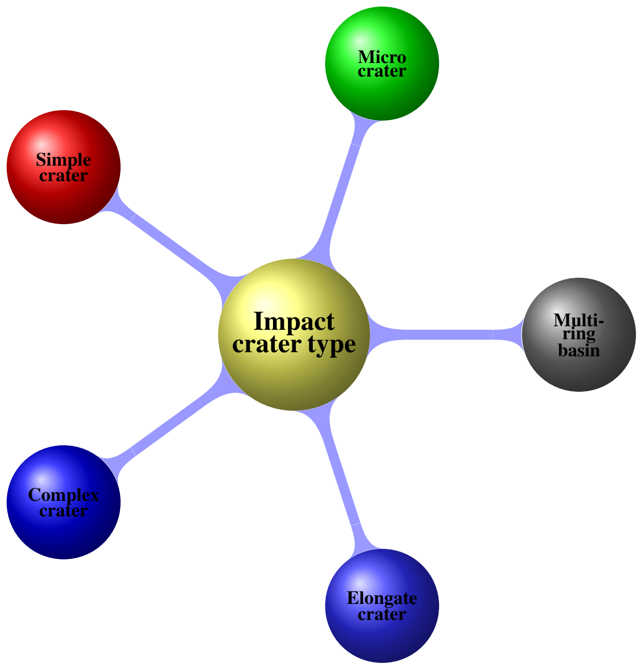

Figure 1. Object Carousel computed for impact crater categories.

Any objects in the categories can carry attributes like time and space as well as objects in other categories, which allows to have dimensions across disciplines like in the following example with impact craters and volcanoes.

4) Results of systematical use: Suitable views for volcanic features are: Type (of volcano, coarse categories), date on timeline, size (height). For craters respective views are: Type (of crater, fragmentary), date on timeline, size (diameter). Two Object Carousels have been computed. Figure 2 shows the knowledge resources groups for volcano types, and Figure 3 provides the geological spaces of time references.

Figure 2. Object Carousel for volcano and type references computed for terrestrial volcanism, providing volcano type references.

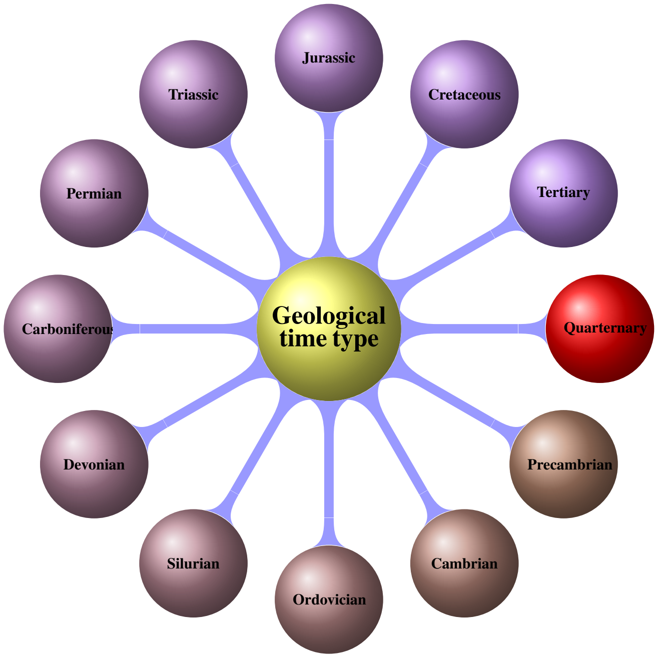

Figure 3. Object Carousel on geological spaces of time for computed references (terrestrial volcanoes, impact craters, and geological processes).

For simplicity only the main groups are shown, subgroups like for Quarternary "Holocene" and "Pleistocene" create separate carousels (Figure 4). Most geological objects have references into some instance of these carousels. This enables to create numberless links to additional information.

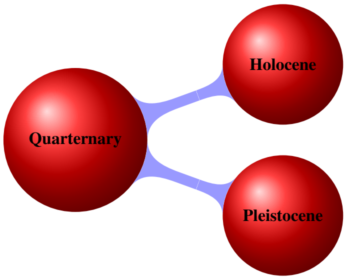

Figure 4. Object Carousel "Quarternary".

The colour coding for Carousels is symbolic and can be defined to represent any grouping as decided within the workflow. It can result from the grade of detail required for the description. In this case, the colour red links the three shown Object Carousels with the information referring to a requested object like "Vesuvius". The subgroup Object Carousels, e.g., "Quarternary" (Figure 4), opens additional references to volcanological feature objects. The listing in Figure 5 shows context replacement definitions and corrections.

```
Cretacious :: Cretaceous
Kreide :: Cretaceous
Trias :: Triassic
Carbon :: Carboniferous
Karbon :: Carboniferous
Silurium :: Silurian
Silur :: Silurian
Ordovicium :: Ordovician
Ordovizium :: Ordovician
Cambrium :: Cambrian
Kambrium :: Cambrian
Precambrium :: Precambrian
Präkambrium :: Precambrian
```

Figure 5. Replacement definition for relevant terms (LX resources).

The example lists an excerpt of relevant terms and types of notation that can be considered equal for the target context.

5) Processing media citation references: Figure 6 shows an excerpt of a media citation set used with UDC classified knowledge objects, here with a Vesuvius reference.

```
cite: YES 20070000 {LXK:Pompeii; Vesuvius; reconstruction
 ; 3D; animation; Holocene} {UDC:...} {PAGE:----..----}
 LXCITE://Bonaventura:2007:My_DVD
cite: YES 20130000 {LXK:Pompeii; Vesuvius; Vesuvio;
 Holocene; postcard} {UDC:...} {PAGE:----..----} LXCITE:
 //Guardasole:2013:Vesuvio_1270m
cite: YES 20070000 {LXK:Pompeii; Vesuvius; reconstruction
 ; diorama} {UDC:...} {PAGE:----..----} LXCITE://
 Bonaventura:2007:Pompeii
cite: YES 20070000 {LXK:Pompeii; Vesuvius; bakery; mill
 stones; material; stone; volcanic lava; basalt; Holocene
 ; diorama} {UDC:...} {PAGE:--56..--59} LXCITE://
 Bonaventura:2007:Pompeii
```

Figure 6. Media citation set excerpt used with the UDC classified knowledge object "Vesuvius" (LX resources).

The examples are part of the "Vesuvius" and "volcanic mill stone" object references. The media citations refer to 3D video animations and dioramic reconstructions as well as even to postcards. These references resolve to [49] (animation), [50] (postcard), [51] (diorama).

Objects can carry an arbitrary number of classifications views, e.g., from automated classification to individual classification by different groups of experts. The facets itself are to be built from a base of universal classification, which is under continuous development and fully consistent in its editions.

For example, knowledge can be created by a single researcher, a research institute, a collaborative effort or any other process. Knowledge resources can therefore be created by a single group or they may be created by larger organisations. Taking library and museum scenarios as examples, then the practice is to have editions representing classification states as well as instances of objects. The resources and objects can use any number of these editions and instances.

6) Classification development: All classifications are subject of a continuous development, review, and auditing process. This is also true for the UDC itself, independent from its use for different disciplines or scenarios. Table VII shows an example in different UDC editions.

Table VII. Development of “Tertiary” classification with UDC editions and knowledge resources (excerpt).

UDC Code (a)	*UDC Code (b)*	*Description*
UDC:”623”	UDC:”628”	Tertiary (70-1 MYBP)
UDC:”623.1”	UDC:”628.2”	Palaeogene (70-25 MYBP)
UDC:”623.5”	UDC:”628.4”	Neocene (25-1 MYBP)
UDC:551.77	UDC:551.7+”628”	Cenozoic (Cainozoic). Neozoic
UDC:551.78	UDC:551.7+”628”	Tertiary. 70-1 MYBP
UDC:551.781	UDC:551.7+”628.2”	Palaeogenic. Nummulitic
UDC:551.781.3	UDC:551.7+”628.22”	Palaeocene
UDC:551.781.4	UDC:551.7+”628.24”	Eocene
UDC:551.781.5	UDC:551.7+”628.26”	Oligocene
UDC:551.782	UDC:551.7+”628.4”	Neogene
UDC:551.782.1	UDC:551.7+”628.42”	Miocene
UDC:551.782.2	UDC:551.7+”628.44”	Pliocene

The example is the “Tertiary” classification development within different UDC editions. The table shows that the target not only moved $(a) \rightarrow (b)$ within the classification but was also adapted to a new subgrouping (lower block). The currently final result is a composite classification, composing from geology and time, holding both Tertiary and Cenozoic.

UDC still not considers different stratigraphies in plain. Further, respective editions can be used or references to respective editions and entries. The editions are fully consistent in themselves, so it is natural that the overall consistency of workflows using different editions has to be cared for by the disciplines or providers.

Figure 7 shows Object Carousels computed for a complete common system (top) as well as for an alternative system (below) used for some purposes [43] after the year 2000, missing “Tertiary”. The colours represent the term levels within the respective system.

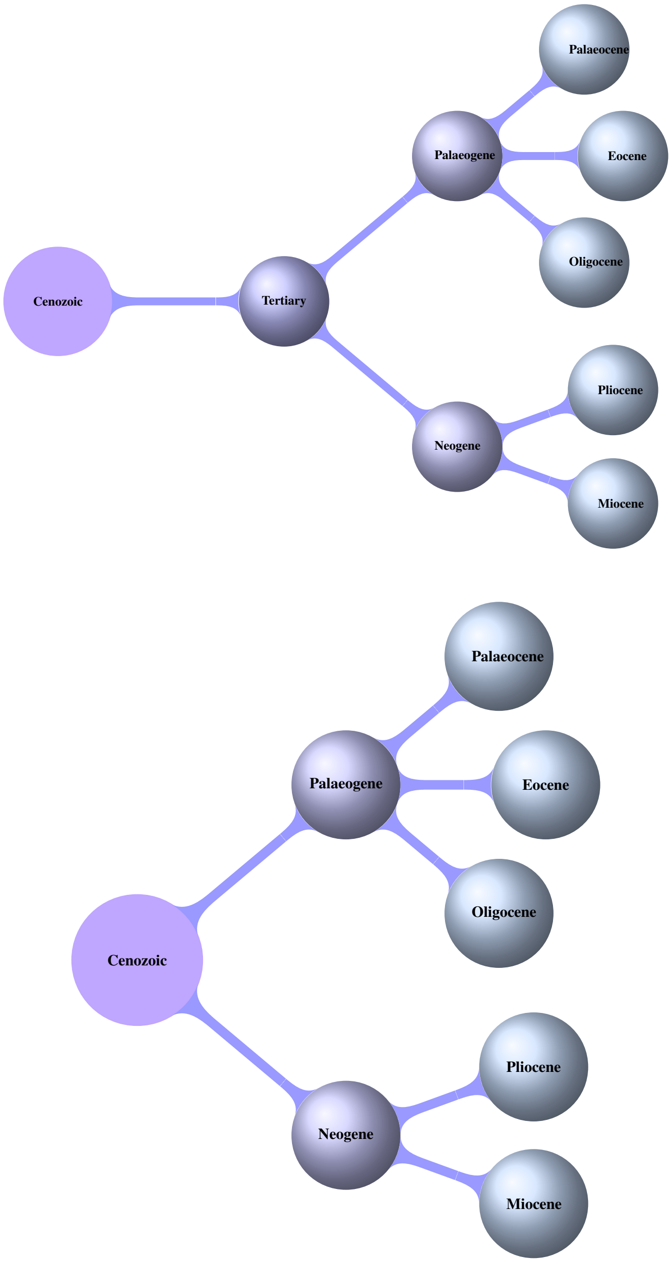

Figure 7. Object Carousel “Tertiary”: Common (top) and alternative (below).

Moved items have to be considered “persistent” within long-term knowledge resources appropriately with all consequences. It is possible to support any number of instances or versions within the knowledge resources as long as each is handled consistently. In this case, the consistency can refer to the classification, the content, as well as to the references. The UDC can reflect the developments in classification with editions, which can be used consistently with the resources.

7) Result matrix: Table VIII shows the results from the computation of a systematical classification of volcanological features, short “volcano types”.

TABLE VIII. COMPUTED SYSTEMATICAL CLASSIFICATION OF VOLCANOLOGICAL FEATURES FROM THE KNOWLEDGE RESOURCES.

Volcano Type	*Group*	*References Data Examples*
Complex volcano	A	Vesuvius VNUM:0101-02= UDC:[551.21+911.2+55]:[902]"63"(4+23+24)... GPS:40.821N14.426E Quarternary VEI:VEI5
Compound volcano	A	Cayambe VNUM:1502-004 UDC:[551.21+911.2+55]:(8+23+24)... GPS:... Holocene ...
Somma volcano	A	Ebeko VNUM:0900-38= UDC:... GPS:... Quarternary ...
Submarine volcano	A	Campi Flegrei Mar Sicilia VNUM:0101-07= UDC:... GPS:... Quarternary ...
Subglacial volcano	A	Katla VNUM:1702-03= UDC:... GPS:... Quarternary ...
Unspecified type	A	– VNUM:– GPS:... – ...
Strato volcano	B	Vulcano VNUM:0101-05= UDC:... GPS:... Quarternary ...
Shield volcano	C	Etna VNUM:0101-06= UDC:... GPS:... Quarternary ...
Explosion crater	D	Larderello VNUM:0101-001 UDC:... GPS:... Quarternary ...
Caldera	D	Campi Flegrei VNUM:0101-01= UDC:... GPS:... Quarternary ...
Tuff cone	E	Tutuila VNUM:0404-02- UDC:... GPS:... Holocene ...
Scoria cone	E	Antofagasta de la Sierra VNUM:1505-124 UDC:... GPS:... Holocene ...
Pyroclastic cone	E	Anunciacion, Cerro VNUM:1405-032 UDC:... GPS:... Holocene ...
Cinder cone	E	Chiquimula Field VNUM:1402-20- UDC:... GPS:... Holocene ...
Lava dome	E	El Chichon VNUM:1401-12 UDC:... GPS:... Quarternary ...
Volcanic field	F	Holotepec VNUM:1401-07- UDC:... GPS:... Quarternary ...
Hydrothermal field	F	Musa River VNUM:0503-02= UDC:... GPS:... Quarternary ...
Fumarole field	F	Kos VNUM:0102-06= UDC:... GPS:... Pleistocene ...
Maar	F	West Eifel Volcanic Field VNUM:0100-01- UDC:... GPS:... Quarternary ...
Fissure vent	F	Quetena VNUM:1505-074 UDC:... GPS:... Holocene ...

It compiles a small excerpt of computed data from the LX resources [33]. The table delivers comprehensive information for the volcanological topics integrated here: Volcanic feature types, computed groups, UDC mappings, and examples of computed references, e.g., Volcano Number (VNUM) the volcanic reference file number, geo-coordinates and spatial data, and spaces of time, as well as referenced data, e.g., the Volcanic Explosivity Index (VEI) [52]. The full result matrix for this request contains several hundreds of computed objects with tenthousands of references.

The selection of the most relevant objects is not an issue of the documentation or the view itself. The selection must be handled by the algorithms and workflows from the disciplines handling the specific resources, e.g., for providing a ranking. Disciplines can be any, therefore the algorithms can comprise statistics as well as ranking algorithms or image processing techniques, depending on the objects and their features.

In this case the most relevant elements are defined by the objects in the volcanological feature container. In addition, in-depth completion within object containers has been enabled for the case of volcanological features. A container represents a collection of equally structured groups of related objects on a certain topic. Examples for available containers are:

- Volcanological object container,
- Earthquake object container,
- Meteoric impact object container,
- Astronomy object container,
- Mineral object container.

The in-depth completion allows additional data, e.g., spatial data, processing, and media object references.

The resources further allow for a flexible mapping of attributes, e.g., container relations, classification, keywords, numbers, references, media samples, material samples, spatial data, and geological spaces of time. With these references the volcanological features can be associated with a VEI, e.g., Vesuvius (Pompeii) VEI5, Krakatau VEI6, Tambora VEI7, Thera (Santorini) VEI7, Toba (Sumatra) VEI8, whereas a "Caldera" object itself being a crater does not have a VEI.

With existing models used in simulation and modelling there is no consideration of references between disciplines, e.g., volcanoes and weather. With the knowledge resources, volcanological features can be referred to volcanological events, seismological events, and weather phenomenon events or biology. The larger the data base is the more correlatable events get available in space and time. In comparison to mono-disciplinary information the multi-disciplinary context of the knowledge resources supports an improved knowledge description. Further, even indirect correlation, e.g., in the above case between volcanic features and meteorite impact features can be investigated.

8) Knowledge generation, combination, and visualisation: The following visualisation (Figure 8) paradigmatically illustrates the results from the compute requests. An on-location attribute has been chosen for the relations in order to compute a distribution map for volcanic features using the `lxlocation` workflow. The location attribute is suitable for referring to an unlimited number of multi-disciplinary information in this case.

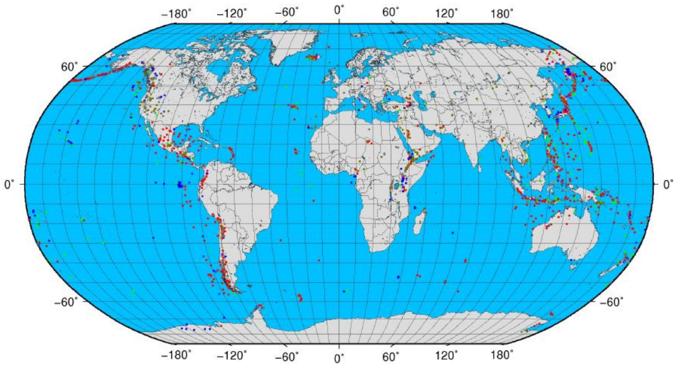

Figure 8. Volcanomap – computed worldwide spatial distribution of classified terrestrial volcanological features from resulting object entries.

The distribution is computed from the result matrix of related object context of several hundred classified terrestrial volcanic features via the knowledge resources research database. The result matrix is the result of the present content, references, and workflow. In all examples only an excerpt of these can be shown. Several modules have been used for this example: `select_knowledge_environment`, `lxgrep_udc`, `lxkwgrep`, `generateCarousel`, `lxvolcanoes2gmt`, `cprgmt_world_cprvolcanoes_separated`, as well as `pscoast`, `pstoraster`, and `psxy`. System interfaces can be created via instructions, programming interfaces, or any kind of interface the disciplines working on implementations and suggested workflows want to built on top of the knowledge resources. The workflow allows any feature supported by the deployed components, e.g.,

- Association by classification weighting,
- Association by grouping,
- Association by colourisation,
- Association and by symbolisation.

Any association and context that can be described and expressed for any objects can be realised with the knowledge resources. In this example, colour groups have been computed via the result matrix (Table VIII): A: green, B: red, C: blue, D: lighter blue, E: grey, F: dark green. The volcanic features are classified and several classification groups have been chosen for the result. The map shows the present situation according to the present state of the available volcanological data.

If we want to generate an according section view on data from the geo-related knowledge resources, we can choose from a number of topics and disciplines. A comparable section view regarding the dimensions and attributes can be described on content base (object classification) by a set of

- Planetary surface objects,
- Geological / geophysical context,
- Geo-coordinates,
- Spatial distribution.

The creator of the overall workflow can define what to do with the data and how to present it. This includes adding a suitable application component for representing the knowledge for the the intended purpose. From application component base, in this case GMT, it can be described by

- Mapping,
- Projection,
- Grid parameters,
- Label attributes / size,
- Colours.

The component specific description will most probably be placed with the workflow implementation but if required it might also be integrated with the objects. The result is shown in the associated sample distribution of terrestrial impact features (meteorite) as depicted in Figure 9.

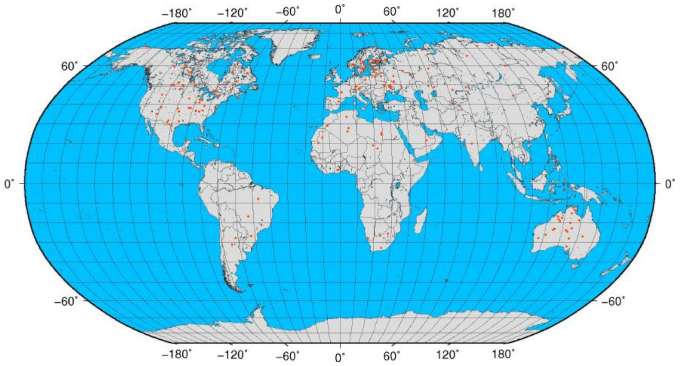

Figure 9. Impactmap – computed worldwide spatial distribution of classified terrestrial impact features (meteorite) from resulting object entries.

The map is computed from the related object context (`lxlocation` workflow) contained in a meteorite impacts features research database of the knowledge resources. It is possible to combine any information, e.g., computing a map animation varying in time, showing the development of volcanic features. Therefore, a temporal sequence of spatial distributions can be used as a simple example for variations over time.

On base of a physical view (criteria classification) the criteria for impact crater classification are:

- Size of the impacting object,
- Speed of the impacting object,
- Material of the impacting object,
- Composition and structure of the target rock,
- Angle that the impacting object hits the target,
- Gravity of the target object respective planet,
- Porosity and other ... of the impacting object,
- Age of the impact,
- Size of the impact,
- Structure of the crater.

Further associated phenomena (indicator classification) are impact crater indicators on the other hand, which are:

- Planar fractures in quartz,
- Shocked quartz,
- Glass fragments.

When approaching from the “catastrophe” view it has shown that the most prominent relation is the “size”. This mostly correlates with “diameter” and still mapping and timelining will come natural.

9) Processing and computational numbers and issues: Table IX shows the processing and computational demands per instance resulting from the presented scenarios.

TABLE IX. PROCESSING AND COMPUTATIONAL DEMANDS.

Item	*Value / Description*
UDC, number of classification items	70,000
Number of classification languages	50
Number of classification variations (50×70,000)	3,500,000
Knowledge object subset, number of items	100,000
Number of terms	10,000,000
Number of object languages	2
Operations, number per subset result entry	50,000,000
Number per subset result entry, incl. keywords	500,000,000
Parallelisation (subset), wall time / num. of nodes	7,500 s / 1
Wall time / number of nodes	1,300 s / 10
Wall time / number of nodes	220 s / 100
Wall time / number of nodes (extrapolated)	4 s / 10,000

Besides the large requirements per instance with most workflows there are significant effects by parallelising even within single instances. The following issues have shown to lead to advanced challenges and increased processing and computational times. Nomenclature, terms, and attributes tend to be at least partially different in different cultures and languages. For many discovery workflows as well as efforts to increase the quality of the result matrices it is necessary to consider more than one culture and language. Processing a classification numbering in decreasing numbering with increasing age or following in different directions is less consequent. For example, in geosciences it is natural to start spaces of time with Quarternary, followed by older stratigraphy. In addition to the existing singular spaces of time mapping most objects require appropriate different mappings to absolute dates, e.g., with Bronze Age having different absolute dates for different regions or cultures. The calculation with extensive composite classifications, facets, and respective ranges instead of native classifications can increase the computational requirements drastically as has been shown with the knowledge content from the Gottfried Wilhelm Leibniz resources [36].

VIII. KNOWLEDGE INTEGRATION: CONCEPTUAL CONTRIBUTION

Integrating referenced knowledge resources can essentially contribute valuable content, both on quantity and depth. Knowledge dimensions as well as computational views can benefit from these contributions.

Nevertheless, references to unstructured or differently structured knowledge objects will contain a number of deviations, which may need to be unified, either for contributing to manifoldness or to intensification of resulting statements.

Utilising the classifications and facets for collecting references to material can vastly enrich the matrices and results, which has already been proposed [53] for solutions coping with Big Data resources [54]. This procedure can help to overcome conceptual deficits of unstructured material as well as language aspects. The procedure delivers a high percentage of relevant otherwise unstructured conceptual knowledge and allows for an efficient unification and integration of the results.

One of the largest resource of references to unstructured, heterogeneous, multi-lingual data are services like Google [55]. Table X shows the results resolved from references of a Google search done for the topics “volcano, udc, classification”. This means references for classifications and other conceptual knowledge can be used, in any combination, in order to trigger search requests. The results contain the UDC classification found with the request as well as the terms associated with this in the text.

TABLE X. VOLCANO RESULTS FROM PUBLICLY AVAILABLE INFORMATION, GOOGLE SEARCH, STATUS OF JANUARY 2013 (EXCERPT).

UDC Classification	*In-text Terms*
UDC:551.442(437.6)	Volcano phreatic
UDC:631.4	Volcano
UDC:553.405	Uranium deposit volcano
UDC:551.31:551.44(532)	Volcano
UDC:(*764)	Volcano
UDC:(*7)	Volcano

Table XI shows the results resolved from references of a Google search done for the topics “cenote, udc, classification”.

TABLE XI. CENOTE RESULTS FROM PUBLICLY AVAILABLE INFORMATION, GOOGLE SEARCH, STATUS OF JANUARY 2013 (EXCERPT).

UDC Classification	*In-text Terms*
UDC:930.85(726.6) 551.435.8:528.9	Sinkhole cenote maya
UDC:551.44	Doline sinkhole
UDC:556.34:519.216	Sinkhole drainage
UDC:551.435.82(234.41)	Sinkhole collapse
UDC:624.153.6:699.8:551.448	Sinkhole collapse
UDC:551.44(450.75)	Karst sinkhole collapse
UDC:551.44(045)=20	Groundwater surfacewater
UDC:551.44:001.4	Grotte Höhle
UDC:551.44(450.75)	Karst Apulia
UDC:551.44(437.2)	Geology karst

All documents found from public external sources have been identified to contain academical and scientific content. Even as this example is intended to provide a lower depth of knowledge mapping than available in specialised knowledge resources, it provides an excellent spectrum of related information for the respective disciplines. First, the unstructured “Big Data” Google is using is currently not classificed and, second, Google cannot search using a universal classification up to now. This means that up to now most conceptual information request cannot be expressed directly via search engines.

Digital sources can be consistently classified in groups as well as single objects. Groups or objects and their associated references can refer to a classification. The examples show the possible bandwidth within a topic. They show that there is much more in-depth knowledge in the data. The classification support can extract focussed knowledge even if the data is only available in different languages as in this case. Therefore, internal as well as external sources have been used for the examples.

IX. DISCUSSION AND EVALUATION

The results presented here have several scientific and technical aspects. The following passages discuss some major contributions to the content and scientific results as well as to the knowledge resources and features.

A. Knowledge resources

The Knowledge Oriented Architecture (KOA) of the resources is based on a flexible integration of the documentation and development architecture utilising the Collaboration house framework for disciplines, services, and resources [34]. The knowledge objects, here the geological and volcanic feature objects, can be used with any of their attributes. Therefore, any references to objects belonging or referring to any other objects can be computed from this. For an object referred to a timescale of periods other objects can be associated with the respective object, even beyond direct references. For example, "geological time type" can refer from "volcano type" to any other suitable for a geological or comparable spaces of time classification. For example, this will be true for geophysical, palaeontological or archaeological objects. Further, volcanic objects from the Quarternary can be associated to meteorite impact events from the Quarternary. The more, they can be restricted to associated objects of a certain attribute, e.g., from the same region. With secondary steps further information can be integrated. This can include geophysical data, media data or associated objects. The resulting quality depends on an intelligent use of context and classification. A strong classification support is essential, the more as object and even many citations, media, and publications are not explicitly aware of the nomenclature of spaces of time used with specific content can, e.g., to express that the spaces of time refer to plants or animals. Employing a universal classification with multidisciplinary content this way, e.g., with volcanological content, expedites knowledge discovery as well as it targets on scientific discovery.

Regarding methodology it further allows to

- Support a systematic documentation,
- Define a normative classification,
- Define cognitive interfaces.

Regarding architecture and implementation it allows to

- Support decision making in complex systems,
- Implement learning system components and
- Support components by intelligent systems.

Creating classified knowledge resources objects has proven to be most sustainable for a significant period of operation and implementation. It has been efficient and portable with all application scenarios and environments for more than two decades, used with ten different operating system environments, with different editing components, processing languages, and compute and storage resources. From classification side it is suggested to have advanced computing support, e.g., for spaces of time as well as for the complementary systematics for disciplines. In addition, a methodological framework for UDC supporting the required processing and computation would add immense benefits to its universal applicability. Some new types of stratigraphies have not widely been adopted and should again become subject of modification regarding a long-term use. In many cases, the consequence of claims on consistency has been to use one dedicated edition of the classification. This shall ensure consistency within the application. Using a small subset of classification can help to reduce the appearent work that has to be done for classification but it cannot ensure to avoid variances in different editions. Consistent version management support for the classification has shown to become necessary as soon as knowledge resources are using modified classifications over time.

B. Content: Case lessons learned

Besides the detailed results and references, the overall results for the discussed cases are:

- Volcanic features are well known above and below sea-level and are more often long-term processes.
- Known impact features show a concentration in highly populated and industrialised areas.
- Impact features have been reduced by morphological processes and are mostly only available above sea-level.
- Both impact and volcanic features are related to social and archaeological findings.
- Both impact and volcanic features are publicly known.
- Compared with impacts and volcanic features, archaeological sites and results are not known to the same amount in order to protect the sites.

C. Classification expenses and benefits

Classification support, e.g., via UDC, does require intensive work and can be expensive if used in non automated ways. Nevertheless, this can make a difference as classification views very much profit from professional experiences. The application of UDC with complex knowledge context requires flexibility of the resources as wells as a flexible handling regarding extendability. The challenges with the distributed use of UDC are, e.g., the use of private catalogues, like external codes or author abbreviations. In addition, the sustainability of knowledge objects benefits from the use of methods like facetted versioning, universal dates (e.g., ISO dates), and georeferencing.

D. Complementary information and classification

Text information and classification information are complementary. This is important for knowledge resources as well a for application scenarios, e.g., search algorithms. Using classification supported search algorithms can improve the result drastically. The quality of results improves from below five percent to up to over ninety percent.

With the presented Object Carousels an undefined number of practical workflows can be created on the knowledge resources. Examples, which have been investigated for gathering complementary results are regular expression and string search, classification search (UDC), keyword search, sort support search, references search or phonetic search (Soundex).

E. Knowledge and views

UDC can be used with any object in any context and can help to reduce compute requirements with knowledge discovery.

The application of a universal classification and knowledge resources can drastically reduce the computational requirements, as well as it supports the parallelisation of instances within workflows due to the large numbers of representations in common per-instance calls (Table IV). The classification views and facets enable to reduce the amount of object analysis required for discovery and reuse workflows. The algorithms benefitting from this are on the algorithm side (for example, object and references search, content string comparisons, and knowledge based associations) on the hardware and resources side (for example, input/output requests including read/write processes, compute resources' and memory requirements, and communication requirements).

X. CONCLUSION AND FUTURE WORK

It has been demonstrated that multi-disciplinary knowledge-based objects can be successfully used in order to create computational views. The knowledge processing employing UDC classification has shown to be a universal and most flexible solution for creating long-term multi-disciplinary knowledge resources and providing a base for universal knowledge dimensions.

The implemented knowledge resources as well as deployed conceptual knowledge, demands on processing and resources, and examples for the computational views have been presented, delivering components that can be used with future developments. In addition, the examples present content and references on about more than 1,000 objects (volcanological features and terrestrial impact features), which delivered a base for new discovery after having been integrated with the knowledge resources.

The resources and framework can be used even with basic attributes and cross-references, and assure support for subsequent use and knowledge procurement processes. Structuring and classification with long-term knowledge resources and UDC support have successfully provided excellent solutions, which can be used for natural sciences, e.g., geosciences, volcanology or with spatial disciplines as well as for universal knowledge. The knowledge resources can provide any kind of Object Carousel and object references. Decisions can be computed with support of the UDC classification. Due to the universal long-term multi-disciplinary knowledge gathering, the knowledge resources are a general universal decision support base.

Besides these, a major benefit of the extensive support of UDC language translations is that regarding discovery workflows it can also be used for improving the quality as well as the quantity of elements and references in the result matrices. Employing a universal classification when creating knowledge resources has provided substantial benefit for both. The workflow procedures build for special purposes are property of the researchers and disciplines creating, developing, and operating their implementations. The data used by them is intended to be part of the respective collaboration. Currently, if someone creates data, he or she can use the data and share it with others, creating agreements and policies.

As knowledge resources have been proven to be a valuable means for research in many disciplines, components are candidates for community tasks as well as for open access development and licensing models. Currently, the policies with many collaborations, funding, and services (as comparable with the UDC model) do not allow to make sources and content of the knowledge resources public.

Because the process of creating long-term sustainable content is quite pretentious and will never be completed there might be support by a sustainable funding in the future, too. No discipline should commonly be funded because it is possible to increase the disciplines' requirements for associated tasks. Generally, the priority should be on knowledge "content and result" and not merely on the quantity of background resources required.

On the other end, the operation for disciplines, services, and resources providers can be accompanied by licensing models, supporting a sustainable long-term development and operation on all three sections. Factual and conceptual knowledge, e.g., information collections and classification editions, web services and interfaces, e.g., discovery and section views, as well as compute and storage services can be provided, developed, and priced that way.

As presented, the knowledge processing can base on a solid and sustainable long-term resource, which allows to create any kind of workflows, dynamical discovery, and IICS components and facilitate the use of High End Computing resources. Based on this research, in the future further features for a high end technical integration of data interfaces and resources and more intelligent learning components can be developed.

ACKNOWLEDGEMENTS

We are grateful to all national and international partners in the GEXI cooperations for the innovative constructive work. We thank the Science and High Performance Supercomputing Centre (SHPSC) for long-term support of collaborative research since 1997, including the GEXI developments and case studies on archaeological and geoscientific information systems. Special thanks go to the scientific colleagues at the Gottfried Wilhelm Leibniz Bibliothek (GWLB) Hannover, especially Dr. Friedrich Hülsmann, for prolific discussion within the "Knowledge in Motion" project, for inspiration, and practical case studies. Many thanks go to the scientific colleagues at the Leibniz Universität Hannover, especially Mrs. Birgit Gersbeck-Schierholz, to the Institute for Legal Informatics (IRI), Leibniz Universität Hannover, and to the Westfälische Wilhelms-Universität (WWU), for discussion, support, and sharing experiences on collaborative computing and knowledge resources and for participating in fruitful case studies as well as to the participants of the postgraduate European Legal Informatics Study Programme (EULISP) for prolific discussion of scientific, legal, and technical aspects over the last years.

REFERENCES

[1] C.-P. Rückemann, “Knowledge Processing for Geosciences, Volcanology, and Spatial Sciences Employing Universal Classification,” in Proceedings of The Sixth International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2014), March 23–27, 2014, Barcelona, Spain. XPS Press, 2014, pp. 76–82, ISSN: 2308-393X, ISBN: 978-1-61208-326-1, URL: http://www.thinkmind.org/download.php?articleid=geoprocessing_2014_4_10_30044 [accessed: 2014-11-30].

[2] A. Woodie, “Forget the Algorithms and Start Cleaning Your Data,” Datanami, 2014, March 26, 2014, URL: http://www.datanami.com/datanami/2014-03-26/forget_the_algorithms_and_start_cleaning_your_data.html [accessed: 2014-11-30].

[3] “Universal Decimal Classification Consortium (UDCC),” 2014, URL: http://www.udcc.org [accessed: 2014-11-30].

[4] “Universal Decimal Classification (UDC),” 2014, Wikipedia, URL: http://en.wikipedia.org/wiki/Universal_Decimal_Classification [accessed: 2014-11-30].

[5] A. Slavic, “UDC libraries in the world - 2012 study,” universaldecimalclassification.blogspot.de, 2012, Monday, 20 August 2012, URL: http://universaldecimalclassification.blogspot.de/2012/08/udc-libraries-in-world-2012-study.html [accessed: 2014-11-30].

[6] “Wissensmanagement in Bibliotheken,” library essentials, LE_Informationsdienst, April 2014, 2014, pp. 9–11, ISSN: 2194-0126, URL: http://www.libess.de [accessed: 2014-11-30].

[7] S. Yadagiri and T. K. Gireesh Kumar, Knowledge Management: Changing Role of LIS in the Digital Environment. B. R. Publishing Corporation, Delhi, 2014, in: Libraries in the Changing Dimensions of Digital Technology, Commemorative Publication in Honour of Prof. D. Chandran, 2013, pp. 476–481, URL: http://eprints.rclis.org/22813 [accessed: 2014-11-30].

[8] C. Sherman, “What’s the Big Deal About Big Data?” Online Searcher, 2014, pp. 10–16, March/April, 2014.

[9] “Was ist so „Big“ an Big Data?” library essentials, LE_Informationsdienst, April 2014, 2014, pp. 6–8, ISSN: 2194-0126, URL: http://www.libess.de [accessed: 2014-11-30].

[10] Fundamentals of Library of Congress Classification, Developed by the ALCTS/CCS-PCC Task Force on Library of Congress Classification Training, 2007, Robare, L., Arakawa, S., Frank, P., and Trumble, B. (eds.), ISBN: 0-8444-1186-8 (Instructor Manual), ISBN: 0-8444-1191-4 (Trainee Manual), URL: http://www.loc.gov/catworkshop/courses/fundamentalslcc/pdf/classify-trnee-manual.pdf [accessed: 2014-10-30].

[11] O. Günther, “Big Data: Was ist das? Und was bedeutet es für Wissenschaft, Wirtschaft und Gesellschaft?” Informatik Spektrum, “Big Data”, Springer Verlag, Berlin, Heidelberg, vol. 37, no. 2, Apr. 2014, pp. 85–86, DOI: 10.1007/s00287-014-0783-7.

[12] C. Leng, “Die Vorstandsperspektive: Big Data = Big Wealth?” Informatik Spektrum, “Big Data”, Springer Verlag, Berlin, Heidelberg, vol. 37, no. 2, Apr. 2014, pp. 88–89, DOI: 10.1007/s00287-014-0780-x.

[13] J.-C. Freytag, “Grundlagen und Visionen großer Forschungsfragen im Bereich Big Data,” Informatik Spektrum, “Big Data”, Springer Verlag, Berlin, Heidelberg, vol. 37, no. 2, Apr. 2014, pp. 97–104, DOI: 10.1007/s00287-014-0771-y.

[14] C. Meinel, “Big Data in Forschung und Lehre am HPI,” Informatik Spektrum, “Big Data”, Springer Verlag, Berlin, Heidelberg, vol. 37, no. 2, Apr. 2014, pp. 92–96, DOI: 10.1007/s00287-014-0773-9.

[15] P. Liggesmeyer, J. Dörr, and J. Heidrich, “Big Data in Smart Ecosystems,” Informatik Spektrum, “Big Data”, Springer Verlag, Berlin, Heidelberg, vol. 37, no. 2, Apr. 2014, pp. 105–111, DOI: 10.1007/s00287-014-0772-x.

[16] IBM, “Wie man Erkenntnisse aus Big Data optimiert,” 2014, URL: http://www.ibm.com/smarterplanet/de/de/madewithibm/stories/#!story/18?ref=home&cmp=333ab&ct=333ab02w&cr=google&cm=k&csr=41429analytics_and_big_data-astron&ccy=de&ck=big_data_storage&cs=broad&S_PKG=-&S_TACT=333ab02w&mkwid=sRUgxqgp6-dc_41984993492_432i044571 [accessed: 2014-11-30].

[17] S. Fischer, “Big Data: Herausforderungen und Potenziale für deutsche Softwareunternehmen,” Informatik Spektrum, “Big Data”, Springer Verlag, Berlin, Heidelberg, vol. 37, no. 2, Apr. 2014, pp. 112–119, DOI: 10.1007/s00287-014-0770-z.

[18] G. Dueck, “Bigger Data – Bigger Trouble?” Informatik Spektrum, “Big Data”, Springer Verlag, Berlin, Heidelberg, vol. 37, no. 2, Apr. 2014, pp. 139–143, DOI: 10.1007/s00287-014-0774-8.

[19] M. E. Jennex, “A Proposed Method for Assessing Knowledge Loss Risk with Departing Personnel,” vol. 44, no. 2, 2014.

[20] “Wissensverlust vermeiden beim Abgang von Wissensarbeitern,” library essentials, LE_Informationsdienst, Juni/Juli 2014, 2014, pp. 9–11, ISSN: 2194-0126, URL: http://www.libess.de [accessed: 2014-11-30].

[21] “International Expert Panel on Quality and Risk Assessment in Telecommunications Services, July 22, 2014, The Tenth Advanced International Conference on Telecommunications (AICT 2014), The Tenth International Conference on Internet Monitoring and Protection (ICIMP 2014), July 20–24, 2014, Paris, France,” 2014, URL: http://www.iaria.org/conferences2014/ProgramAICT14.html [accessed: 2014-11-23], URL: http://www.iaria.org/conferences2014/filesAICT14/AICT2014_EXPERT_PANEL.pdf [accessed: 2014-11-23].

[22] C.-P. Rückemann, “Knowledge Integration for Scientific Classification and Computation,” in The Fourth Symposium on Advanced Computation and Information in Natural and Applied Sciences, Proceedings of The 12th International Conference of Numerical Analysis and Applied Mathematics (ICNAAM), September 22–28, 2014, Rhodes, Greece, Proceedings of the American Institute of Physics (AIP), AIP Conference Proceedings. AIP Press, 2014, ISSN: 0094-243X.

[23] C.-P. Rückemann and B. F. S. Gersbeck-Schierholz, “Object Security and Verification for Integrated Information and Computing Systems,” in Proceedings of the Fifth International Conference on Digital Society (ICDS 2011), Proceedings of the International Conference on Technical and Legal Aspects of the e-Society (CYBERLAWS 2011), February 23–28, 2011, Gosier, Guadeloupe, France / DigitalWorld 2011. XPS, 2011, pp. 1–6, ISBN: 978-1-61208-003-1, URL: http://www.thinkmind.org/download.php?articleid=cyberlaws_2011_1_10_70008 [accessed: 2014-11-30].

[24] C.-P. Rückemann, Sustainable Knowledge and Resources Management for Environmental Information and Computation. MacMillan, 2014.

[25] C.-P. Rückemann, “Fostering Environmental Protection via Sciences and Society: Key Knowledge and Climate Change,” May, 2014, Contribution to ISSC, 2014.

[26] World Social Science Report 2013, Changing Global Environments, 1st ed. Published jointly by the United Nations Educational, Scientific and Cultural Organization (UNESCO), the International Social Science Council (ISSC), the Organisation for Economic Co-operation and Development (OECD), 2013, DOI: 10.1787/9789264203419-en, OECD ISBN 978-92-64-20340-2 (print), OECD ISBN 978-92-64-20341-9 (PDF), UNESCO ISBN 978-92-3-104254-6 (PDF and print).

[27] “The Digital Archaeological Record (tDAR),” 2014, URL: http://www.tdar.org [accessed: 2014-11-30].

[28] “WDL, World Digital Library,” 2014, URL: http://www.wdl.org [accessed: 2014-11-30].

[29] Wissenschaftsrat, “Übergreifende Empfehlungen zu Informationsinfrastrukturen, (English: Spanning Recommendations for Information Infrastructures),” Wissenschaftsrat, Deutschland, (English: Science Council, Germany), Drs. 10466-11, Berlin, 28.01.2011, 2011, URL: http://www.wissenschaftsrat.de/download/archiv/10466-11.pdf [accessed: 2014-11-30].

[30] C.-P. Rückemann, “Implementation of Integrated Systems and Re-

sources for Information and Computing," in Proceedings of the International Conference on Advanced Communications and Computation (INFOCOMP 2011), October 23–29, 2011, Barcelona, Spain, 2011, pp. 1–7, ISBN: 978-1-61208-009-3, URL: http://www.thinkmind.org/download.php?articleid=infocomp_2011_1_10_10002 [accessed: 2014-11-30].

[31] C.-P. Rückemann, Queueing Aspects of Integrated Information and Computing Systems in Geosciences and Natural Sciences. InTech, 2011, pp. 1–26, Chapter 1, ISBN-13: 978-953-307-737-6, DOI: 10.5772/29337.

[32] "Geo Exploration and Information (GEXI)," 1996, 1999, 2010, 2014, URL: http://www.user.uni-hannover.de/cpr/x/rprojs/en/index.html#GEXI [accessed: 2014-10-26].

[33] "LX-Project," 2014, URL: http://www.user.uni-hannover.de/cpr/x/rprojs/en/#LX (Information) [accessed: 2014-10-26].

[34] C.-P. Rückemann, "Enabling Dynamical Use of Integrated Systems and Scientific Supercomputing Resources for Archaeological Information Systems," in Proceedings of The International Conference on Advanced Communications and Computation (INFOCOMP 2012), October 21–26, 2012, Venice, Italy. XPS, 2012, pp. 36–41, ISBN: 978-1-61208-226-4, URL: http://www.thinkmind.org/download.php?articleid=infocomp_2012_3_10_10012 [accessed: 2014-11-30].

[35] "UDC Online," 2014, URL: http://www.udc-hub.com/ [accessed: 2014-11-30].

[36] C.-P. Rückemann, "Archaeological and Geoscientific Objects used with Integrated Systems and Scientific Supercomputing Resources," International Journal on Advances in Systems and Measurements, vol. 6, no. 1&2, 2013, pp. 200–213, ISSN: 1942-261x, LCCN: 2008212470 (Library of Congress), URL: http://www.thinkmind.org/download.php?articleid=sysmea_v6_n12_2013_15 [accessed: 2014-11-23], URL: http://lccn.loc.gov/2008212470 [accessed: 2014-11-23].

[37] C.-P. Rückemann, "Sustainable Knowledge Resources Supporting Scientific Supercomputing for Archaeological and Geoscientific Information Systems," in Proceedings of The Third International Conference on Advanced Communications and Computation (INFOCOMP 2013), November 17–22, 2013, Lisbon, Portugal. XPS Press, 2013, pp. 55–60, ISSN: 2308-3484, ISBN: 978-1-61208-310-0, URL: http://www.thinkmind.org/download.php?articleid=infocomp_2012_3_10_10012 [accessed: 2014-11-30].

[38] C.-P. Rückemann, "High End Computing for Diffraction Amplitudes," in The Third Symposium on Advanced Computation and Information in Natural and Applied Sciences, Proceedings of The 11th International Conference of Numerical Analysis and Applied Mathematics (ICNAAM), September 21–27, 2013, Rhodes, Greece, Proceedings of the American Institute of Physics (AIP), AIP Conference Proceedings, vol. 1558. AIP Press, 2013, pp. 305–308, ISBN: 978-0-7354-1184-5, ISSN: 0094-243X, DOI: 10.1063/1.4825483.

[39] C.-P. Rückemann, "Advanced Scientific Computing and Multi-Disciplinary Documentation for Geosciences and Archaeology Information," in Proc. of The Int. Conf. on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2013), February 24 – March 1, 2013, Nice, France. XPS Press, 2013, pp. 81–88, ISSN: 2308-393X, ISBN: 978-1-61208-251-6, URL: http://www.thinkmind.org/download.php?articleid=geoprocessing_2013_4_10_30035 [accessed: 2014-11-30].

[40] U. Inden, D. T. Meridou, M.-E. C. Papadopoulou, A.-C. G. Anadiotis, and C.-P. Rückemann, "Complex Landscapes of Risk in Operations Systems Aspects of Processing and Modelling," in Proceedings of The Third International Conference on Advanced Communications and Computation (INFOCOMP 2013), November 17–22, 2013, Lisbon, Portugal. XPS Press, 2013, pp. 99–104, ISSN: 2308-3484, ISBN: 978-1-61208-310-0, URL: http://www.thinkmind.org/download.php?articleid=infocomp_2013_5_10_10114 [accessed: 2014-11-30].

[41] P. Leitão, U. Inden, and C.-P. Rückemann, "Parallelising Multi-agent Systems for High Performance Computing," in Proceedings of The Third International Conference on Advanced Communications and Computation (INFOCOMP 2013), November 17–22, 2013, Lisbon, Portugal. XPS Press, 2013, pp. 1–6, ISSN: 2308-3484, ISBN: 978-1-61208-310-0, URL: http://www.thinkmind.org/download.php?articleid=infocomp_2013_1_10_10055 [accessed: 2014-11-30].

[42] "GMT - Generic Mapping Tools," 2014, URL: http://imina.soest.hawaii.edu/gmt [accessed: 2014-11-30].

[43] International Commission on Stratigraphy, "International Chronostratigraphic Chart," 2014, URL: http://www.stratigraphy.org/ICSchart/ChronostratChart2013-01.pdf [accessed: 2014-11-30].

[44] C.-P. Rückemann, "Integrating Information Systems and Scientific Computing," International Journal on Advances in Systems and Measurements, vol. 5, no. 3&4, 2012, pp. 113–127, ISSN: 1942-261x, LCCN: 2008212470 (Library of Congress), URL: http://www.thinkmind.org/index.php?view=article&articleid=sysmea_v5_n34_2012_3/ [accessed: 2014-11-30].

[45] "LX SNDX, a Soundex Module Concept for Knowledge Resources," LX-Project Consortium Technical Report, 2014, URL: http://www.user.uni-hannover.de/cpr/x/rprojs/en/\lattenkreuZ{} [accessed: 2014-11-30].

[46] C.-P. Rückemann, "Computing Optimised Result Matrices for the Processing of Objects from Knowledge Resources," in Proceedings of The Fourth International Conference on Advanced Communications and Computation (INFOCOMP 2014), July 20–24, 2014, Paris, France. XPS Press, 2014, iSSN: 2308-3484.

[47] "Multilingual Universal Decimal Classification Summary," 2012, UDC Consortium, 2012, Web resource, v. 1.1. The Hague: UDC Consortium (UDCC Publication No. 088), URL: http://www.udcc.org/udcsummary/php/index.php [accessed: 2014-11-30].

[48] "Creative Commons Attribution Share Alike 3.0 license," 2012, URL: http://creativecommons.org/licenses/by-sa/3.0/ [accessed: 2014-11-30].

[49] M. My, Pompeii Reconstructed (DVD). ArcheoLibri, produced by MyMax, in: Bonaventura, Maria Antonietta Lozzi (2007): Pompeii Reconstructed, 2007.

[50] Guardasole, Vesuvio 1270 m. Guardasole SRL, Napoli, Via Argine, 313, Italia, 2013, Postcard, 40067, Description: 1944 eruptions and present day crater, Collection: LX, Provider: BGS, Entry date: 2013.

[51] M. A. L. Bonaventura, Pompeii Reconstructed. ArcheoLibri, 2007, ISBN: 978-88-95512-23-5.

[52] C. G. Newhall and S. Self, "The Volcanic Explosivity Index (VEI): An Estimate of Explosive Magnitude for Historical Volcanism," JGR, vol. 87, 1982, pp. 1231–1238.

[53] Achievements and Successful Solutions with Big Data and Computing Challenges: National and International Perspectives. International Expert Panel, July 23, 2014, The Fourth International Conference on Advanced Communications and Computation (INFOCOMP 2014), DataSys, July 20–24, 2014, Paris, France, 2014, URL: http://www.iaria.org/conferences2014/filesINFOCOMP14/INFOCOMP2014_EXPERT_PANEL.pdf.

[54] C.-P. Rückemann, "Big Data – Bigger Knowledge – Biggest Cognition: Optimising Organisation & Structure for Exa-Scale," International Expert Panel on Achievements and Successful Solutions with Big Data and Computing Challenges: National and International Perspectives, July 23, 2014, The Fourth International Conference on Advanced Communications and Computation (INFOCOMP 2014), DataSys, July 20–24, 2014, Paris, France, [Lecture], 2014, URL: http://www.iaria.org/conferences2014/filesINFOCOMP14/INFOCOMP2014_EXPERT_PANEL.pdf [accessed: 2014-11-23].

[55] "Google," 2014, URL: http://www.google.com [accessed: 2014-11-30].