Knowledge Processing for Geosciences, Volcanology, and Spatial Sciences Employing Universal Classification

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

***Abstract*—This paper presents the results from creating and using long-term knowledge resources for knowledge processing by employing a universal classification. The knowledge objects can further comprise specialised research data collections and refer to any kind of multi-disciplinary data. The resources can be sustainably used for documentation, universal classification, and structuring as well as with scientific supercomputing resources for advanced information systems, supporting discovery and decision making. The case study discusses selected examples from the geosciences context. Within the case studies the structure of the resources and the Universal Decimal Classification (UDC) has been used together with spatial information in order to handle and support workflows and visualisations for multi-disciplinary data and features. Implementations of various components have been created for discovery and development using 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 UDC. The paper discusses the new results and practical cases regarding a multi-disciplinary geosciences, volcanology, and spatial features scenario from long-term knowledge resources.**

***Keywords*–Geosciences; Knowledge Resources; Integrated Systems; Sustainability; Information Systems; Volcanology; Archaeology; Classification; UDC; High Performance Computing.**

I. Introduction

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, digital libraries and isolated content [1], [2]. 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 [3] with Integrated Information and Computing System (IICS) components and High End Computing (HEC) [4]. 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) [5]. 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 [6] concentrating on integrating knowledge from geosciences, volcanology, and spatial sciences disciplines.

This paper is organised as follows. Section II introduces the previous work and components used, Sections III and IV present the systematics and classification used for the processing, and discuss the results from the implementation case study, Sections V and VI evaluate main results, and summarise the lessons learned, conclusions and future work.

II. Components employed

The data used here is based on the content and context from the LX Foundation Scientific Resources [7], [8]. The LX structure and UDC [5], [9] 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 [10]. The classification 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 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 [11] in order to create knowledge object references and to modularise the required algorithms [12]. This provides a universal means for improving coverage, e.g., dark data, and quality within the workflow. 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 and intelligent components on HEC and HPC resources [13], [14]. With the IICS, the Generic Mapping Tools [15] have been used to visualise georeferenced data wherever a spatial representation is reasonable.

III. 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 [16] 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. 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. With the knowledge resources in this research handling 70000 classes, for 100000 objects and several million referenced data the algorithms are mostly non-linear. They allow interactive use, dynamical communication, computing, decision support, and pre- and postprocessing, e.g., visualisation.

The classification deployed for documentation [17] 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.

IV. Case study implementation and results

The following sections discusses 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 [6]. Mapping and timelining will be the natural result.

1) Space: Table I shows an excerpt of the resulting UDC classification practically used for spatial features and place implemented with knowledge resources and geo-coordinates.

Table I. Universal Decimal Classification of spatial features and place used with the knowledge resources (excerpt).

UDC Code	*Description*
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

2) Time: Table II 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 II. UNIVERSAL DECIMAL CLASSIFICATION OF SPACES OF TIME USED WITH THE KNOWLEDGE RESOURCES (EXCERPT).

UDC Code	*Description*
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

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) 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 1 shows the knowledge resources groups for volcano types, and Figure 2 provides the geological spaces of time references. For simplicity only the main groups are shown, subgroups like for Quarternary "Holocene" and "Pleistocene" create separate carousels (Figure 3). Most geological objects have references into some instance of these carousels. This enables to create numberless links to additional information.

Figure 1. Object Carousel for volcano and type references computed for terrestrial volcanism, providing volcano type references.

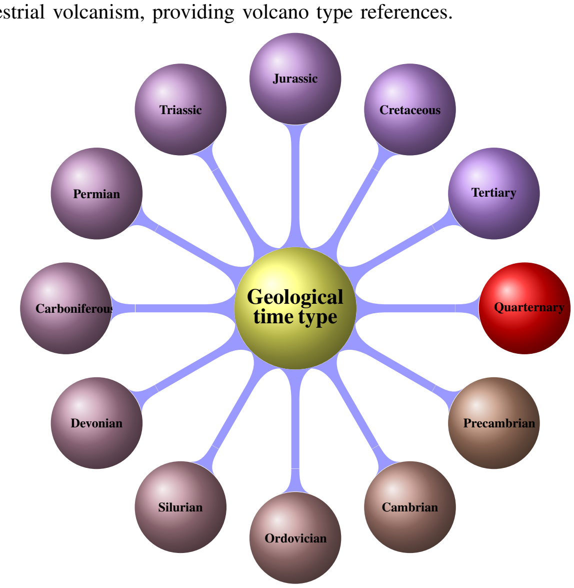

Figure 2. Object Carousel on geological spaces of time for computed references (terrestrial volcanoes, impact craters, and geological processes).

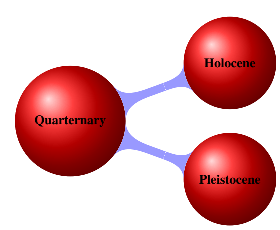

Figure 3. Object Carousel "Quarternary".

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

 ISBN: 978-1-61208-326-1

Object Carousels with the information referring to a requested object like “Vesuvius”. The subgroup Object Carousels, e.g., “Quarternary” (Figure 3), opens additional references to volcanological feature objects. The listing in Figure 4 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 4. 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.

4) Processing media citation references: Figure 5 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 5. 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 [18], [19], [20].

5) Classification development: All classifications are subject of a continuous development, review, and auditing process. Table III shows an example in different UDC editions.

TABLE III. 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. Figure 6 shows Object Carousels computed for a complete common system (top) as well as for an alternative system (below) used for some purposes [16] after the year 2000, missing “Tertiary”. The colours represent the term levels within the respective system.

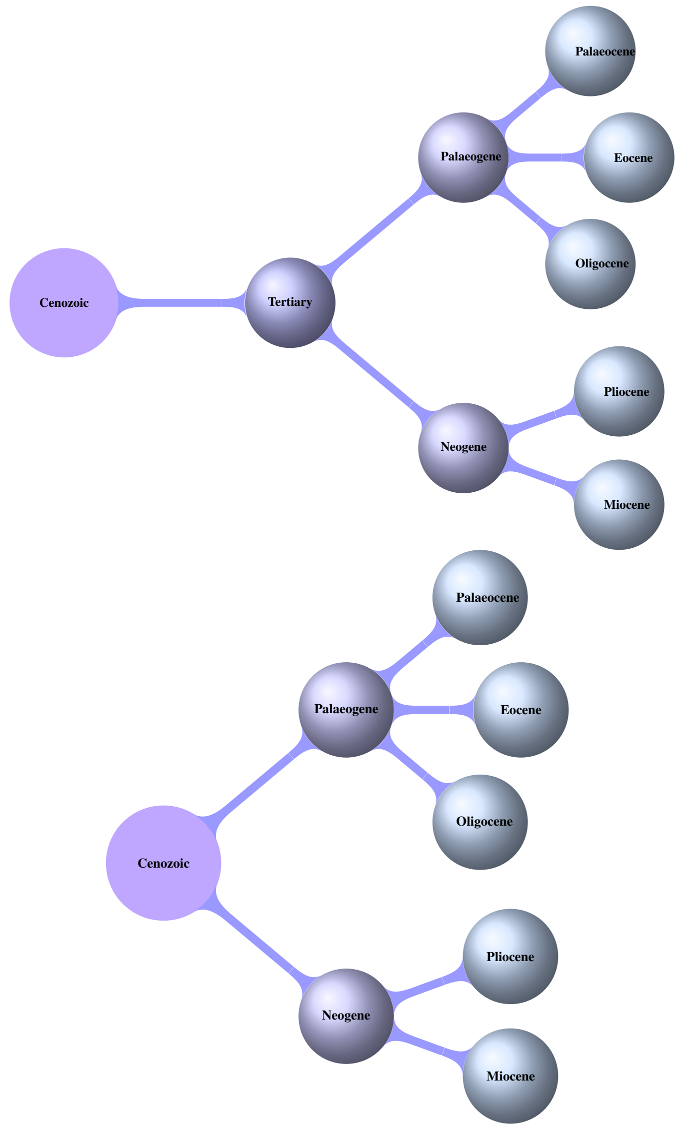

Figure 6. Object Carousel “Tertiary”: Common (top) and alternative (below).

Moved items have to be considered “persistent” within long-term knowledge resources appropriately with all consequences. It is possible to support any number of versions within the knowledge resources as long as each is handled consistently.

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

 ISBN: 978-1-61208-326-1

TABLE IV. 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 [7]. 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) [21]. The full result matrix for this request contains several hundreds of computed objects with tenthousands of references. A container represents a collection of equally structured groups of related objects on a certain topic. In addition, in depth completion within object containers has been enabled for the case of volcanological features. 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.

7) Knowledge generation, combination, and visualisation: The following visualisation (Figure 7) paradigmatically illustrates the results from the compute requests. An on-location attribute has been choosen 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.

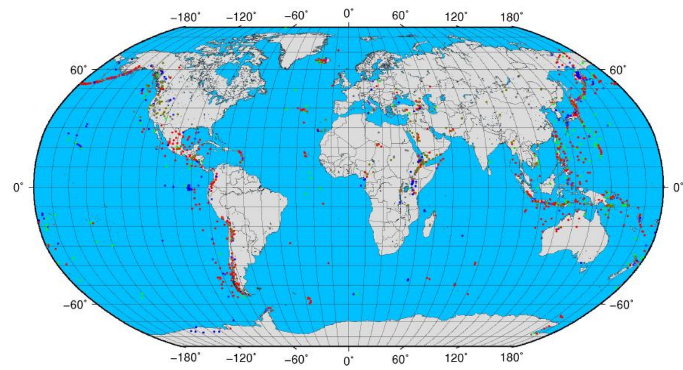

Figure 7. Volcanomap – computed worldwide spatial distribution of classified volcanological features from resulting object entries.

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.

In this example, colour groups have been computed via the result matrix (Table IV): A: green, B: red, C: blue, D: lighter

Copyright (c) IARIA, 2014. ISBN: 978-1-61208-326-1

blue, E: grey, F: dark green. The volcanic features are classified and several classification groups have been choosen for the result. The map shows the present situation according to the present state of the available volcanological data. It is possible to combine any information, e.g., computing a map animation varying in time, showing the development of volcanic features.

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

TABLE V. PROCESSING AND COMPUTATIONAL DEMANDS.

Item	*Value / Description*
UDC, number of classification items	70,000
Number of classification languages	50
Number of classification variations (50×70000)	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 [10].

V. DISCUSSION AND EVALUATION

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 [8]. 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. This will, e.g., 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.

VI. CONCLUSION AND FUTURE WORK

The knowledge processing employing UDC classification has shown to be a universal and most flexible solution for creating long-term multi-disciplinary 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 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.

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.

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, and at the Institute for Legal Informatics (IRI), Leibniz Universität Hannover, and WWU Münster, as well as to the participants of the EULISP Programme for prolific discussion of scientific, legal, and technical aspects over the last years.

REFERENCES

[1] "The Digital Archaeological Record (tDAR)," 2014, URL: http://www.tdar.org [accessed: 2014-01-12].

[2] "WDL, World Digital Library," 2014, URL: http://www.wdl.org [accessed: 2014-01-12].

[3] 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: 2014-01-12].

[4] 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.

[5] "Universal Decimal Classification Consortium (UDCC)," 2014, URL: http://www.udcc.org [accessed: 2014-01-12].

[6] "Geo Exploration and Information (GEXI)," 1996, 1999, 2010, 2014, URL: http://www.user.uni-hannover.de/cpr/x/rprojs/en/index.html#GEXI [accessed: 2014-01-12].

[7] "LX-Project," 2014, URL: http://www.user.uni-hannover.de/cpr/x/rprojs/en/#LX (Information) [accessed: 2014-01-12].

[8] C.-P. Rückemann, "Enabling Dynamical Use of Integrated Systems and Scientific Supercomputing Resources for Archaeological Information Systems," in *Proc. of The Int. Conf. on Advanced Communications and Computation (INFOCOMP 2012), October 21–26, 2012, Venice, Italy*. XPS, 2012, pp. 36–41, ISBN: 978-1-61208-226-4.

[9] "UDC Online," 2014, URL: http://www.udc-hub.com/ [accessed: 2014-01-12].

[10] 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, pp. 200–213, 2013, ISSN: 1942-261x, LCCN: 2008212470 (Library of Congress), URL: http://www.thinkmind.org/download.php?articleid=sysmea_v6_n12_2013_15 [accessed: 2014-01-12].

[11] 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.

[12] C.-P. Rückemann, "High End Computing for Diffraction Amplitudes," in *Proceedings ICNAAM 2013, Rhodes, Greece*, vol. 1558. AIP Press, 2013, pp. 305–308, ISBN: 978-0-7354-1184-5, ISSN: 0094-243X, DOI: 10.1063/1.4825483.

[13] 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.

[14] 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.

[15] "GMT - Generic Mapping Tools," 2014, URL: http://imina.soest.hawaii.edu/gmt [accessed: 2014-01-12].

[16] International Commission on Stratigraphy, "International Chronostratigraphic Chart," 2014, URL: http://www.stratigraphy.org/ICSchart/ChronostratChart2013-01.pdf [accessed: 2014-01-12].

[17] C.-P. Rückemann, "Integrating Information Systems and Scientific Computing," *International Journal on Advances in Systems and Measurements*, vol. 5, no. 3&4, pp. 113–127, 2012, ISSN: 1942-261x, LCCN: 2008212470 (Library of Congress), URL: http://www.thinkmind.org/index.php?view=article&articleid=sysmea_v5_n34_2012_3/ [accessed: 2014-01-12].

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

[19] 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.

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

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