Abstract—This paper presents the research and implementation of applying the new methodology Cogwheel Modules for creating new views and insights from knowledge integration. The target is advanced knowledge mining, e.g., complex discovery, and decision making. The paper provides both results of the present research of the methodology and an implementation, including a case study on different views and possible insight from an application in the spatial domain. The implementation includes modules required for a complete workflow as well as generators for creating results, specifying spatial data and content. The case study utilises topics, techniques, and data from geosciences, archaeology and multi-disciplinary context. The methodology is using integrated knowledge resources for complex knowledge mining by creating workflows applying specialised tools. The resulting methodology can be applied with any disciplines and with combinations of general, as well as specialised tools. The results of the knowledge mining can be used for gaining insight and creating automated learning processes, especially with long-term knowledge resources, which are continuously in development. The method can be used for practical mining procedures to gain insight as well as further develop available multi-disciplinary knowledge resources. The goal of this research is to create new views and insights from the available knowledge resources.

Keywords—Knowledge-based Views and Insights; Cogwheel Modules Methodology; Data-centric Knowledge Mining; Universal Decimal Classification; Advanced Computing.

I. INTRODUCTION

This research is focussed on creating new views and insights from content of knowledge resources. The methodology, which is deployed allows to compute “Cogwheel Modules” and peel information from knowledge resources. A workflow can use the process to iterate in an arbitrary number of turnarounds in order to create a possible knowledge integration.

The fundamentals of the new method of Cogwheel Modules were presented at the DigitalWorld and GEOProcessing 2017 conference in Nice, France [1].

The work is based on the integration of knowledge resources referring to universal classification and application components for solving complex tasks, e.g., for knowledge mining. Target of this research on the methodology of ‘Cogwheel Modules’ is to create different views based on integrating knowledge resources and specialised application components for a gain in knowledge, cognition, and insight.

Creating views means the creation of exhaustive context for knowledge objects and their entities. The primary context can be a knowledge context, which allows further analysis and processing. Based on the primary context, a secondary context can be created, e.g., a result matrix, a listing, or a visualisation.

The integration of knowledge discovery and decision making processes can result in extremely challenging tasks. The quality of results from knowledge mining is primarily connected with content and algorithms. The language or method used for expressing a ‘question’ and automating its translation in general is not of concern for this research.

Data resources, whatever their size is, do not automatically deliver high quality results. In most cases, content and algorithms are limiting possibilities to answer complex and staggered questions in reasonable ways. Contributions to these deficiencies result from data, algorithms, and their implementations. Therefore, high quality knowledge resources, including factual, conceptual, procedural, and metacognitive knowledge, description, and documentation are increasingly important. In consequence, advancing methodologies for knowledge mining is a focus of comparable importance.

Different knowledge references and data require different tools. Several disciplines contribute and specialised approaches and solutions have to be used on context for coping with any slightly complex question. Built on such in-deficit foundation, there is no direct and common practice on how to integrate specialised algorithms and applications with each other without a methodology. Appropriate methodologies will allow to integrate advanced knowledge resources and to modularise several tasks within a knowledge mining workflow. In addition, this research presents more close insights from a case study and the knowledge, especially conceptual knowledge required and provides additional new context examples, factual knowledge, and further case study results.

This paper is organised as follows. Section II introduces the methodology for creating views with advanced knowledge mining. Section III describes the Cogwheel Modules Methodology. Section IV presents an implementation and case study and how to create a primary context. Section V discusses an excerpt of secondary new resulting context, especially different visualisation views leading to new insights. The discussion includes references and associations with the workflow implementation resulting from the implementation and application of the methodology, based on previous work and re-usable components. Section VI summarises the lessons learned, conclusions, and future work.
II. MOTIVATION

The motivation for the research on a new methodology for creating new views from knowledge integration results from the unsatisfactory and non-knowledge centric instruments and state of integration available. For many knowledge mining challenges, e.g., seeking good answers to complex questions, there are no solutions available for integrating complex knowledge resources and arbitrary application components. A sample question is:

Which natural events associated with the creation of crater structures with a diameter larger than 100 m could have been directly notable by human population within the last thousands of years and are still observable on-land at the area of today’s continent of Europe and which knowledge is associated with such events?

The question is quite precise but present possibilities mostly cannot achieve appropriately precise results in order to answer such questions. If one is not satisfied with arbitrary lists of hundreds of snippets of information mostly not part of an answer instead of an on-topic result then we have to find better ways. A solution is to flexibly integrate high quality data with conceptual knowledge and suitable application components with appropriate features. Due to the complexity of integration, the state of the art resources together with supportive data and component resources will be presented and discussed when required in the following section.

III. COGWHEEL MODULES METHODOLOGY

With this research, a methodology is defined by a sequence of steps. The steps can be a set of procedures in order to create a result for a knowledge mining process, e.g., with a discovery process. The procedures can include data, knowledge, formal descriptions, and implementations, e.g., collecting data, retrieving information, and algorithmic specifications. The purpose can range from delivering to creating and answer to an open question, e.g., delivering knowledge for a learning or decision making process. The methodology uses a formal description of knowledge, data and information, as well as required research techniques. Content and context are represented by any knowledge objects and data available in time and space. Data may be structured and unstructured.

1a) Identification of a knowledge mining challenge.
1b) Phrasing of a problem or question.
1c) Identification of a solution or answering strategy.
1d) Context description and modeling.
1e) Mapping of sub-challenges to possible partial solutions.
1f) Interface creation for partial solutions.
2a) Creation and/or selection of Cogwheel Modules (modularisation into sub-challenges and partial solutions).
2b) Knowledge and information: Identification or creation and/or selection of nuclei and facets.
2c) Peeling of information-nuclei from existing evidence.
2d) Milling of nuclei.
2e) Information processing.
2f) Data selection including nuclei and facets.
2g) Information object turnaround.
3a) Workflow implementation (incl. Cogwheel Modules).
3b) Analysis of results.
3c) Learning process and persistent documentation.
3d) Improvement process.

We can identify three main groups within the methodology. 1a) to 1f) is a preparatory phase, 2a) to 2g) describes a gearbox of knowledge mining, and 3a) to 3d) is a consecutive phase.

The modules allow to assign specialised applications and specialised features to separate modules as will be shown in the following implementation. Options and features of specialised applications can be documented, including conceptual knowledge, with the learning process and to cope with recurring requirements. The methodology allows to create different approaches for a workflow.

IV. IMPLEMENTATION AND CASE STUDY

The methodology was applied to practical situations. The following case study presents a practical workflow implementation from 1 to 3 (challenge identifying question to workflow implementation) based on the above gearbox of knowledge, including the required Cogwheel Modules with their mapping to important components and steps, their implementation and results. The goal is to create primary context and –in a consecutive process– secondary context, which in the case means spatial visualisation.

The starting point is the above sample question. The required compositions of features and criteria can become quite complex and are commonly not implemented in any single application or component. Therefore, the integration of appropriate application components can be desirable or even required.

The plethora of information from the knowledge resources is narrowed by the conceptual knowledge, the references to classifications, e.g., to the mapping and data of:

• Craters (any, e.g., Earth and other planets),
  ◦ volcanic features including craters,
  ◦ impact craters including meteorites, . . .
• confirmed (and non-confirmed) structures/craters,
• structures observable on-land,
• age less than (about) 9999 years old,
• larger than 100 m diameter.

The respective workflow requires a number of special calculations as well as criteria Cogwheel Modules for knowledge resources and spatial components.

Applying a universal classification can be used to classify the appropriate objects, the associated application components, and the respective required options for a Cogwheel Module, e.g., for the calculations and filters.

In this case, the two groups of components involved with creating a solution are a) advanced knowledge resources and b) knowledge mining including conceptual knowledge references, spatial data and applications.
The definition of data-centricity used is: “The term data-centric refers to a focus, in which data is most relevant in context with a purpose. Data structuring, data shaping, and long-term aspects are important concerns. Data-centricity concentrates on data-based content and is beneficial for information and knowledge and for emphasizing their value. Technical implementations need to consider distributed data, non-distributed data, and data locality and enable advanced data handling and analysis. Implementations should support separating data from technical implementations as far as possible.” [2].

According to this, the implementation of the methodology is as far data-centric as possible and allows a systematic application.

The following sections describe the essentials of the preparatory phase up to the partial solutions and the Cogwheel Modules required, including the handling of the nuclei and information processing. The sub-challenges are presented with their mapping to applications. Relevant excerpts of data and information are discussed in anticipation of the final results. The concluding section shows the workflow implementation used for creating the final results.

A. Multi-disciplinary knowledge resources identification

The knowledge resources hold arbitrary multi-disciplinary knowledge (e.g., documentation of factual, conceptual, procedural, and metacognitive knowledge), in various structures as well as unstructured, objects, and references, including information on digital objects and realia objects, e.g., media objects and archived physical specimen. These resources provide the prerequisites in order to create efficient Cogwheel Modules and handle knowledge and information nuclei and facets for peeling and milling processes.

1) Factual knowledge: The knowledge resources also contain information on various types of crater features like volcanic craters and impact craters. Especially, the Earth’s impact crater container in the knowledge resources container holds data and references for all known impact craters on Earth.

The knowledge resources provide factual and conceptual data, e.g., crater types, crater/impact ages, and confirmed impact events.

The impact features container holds the Kaali impact, represented by its major impact crater. The minor craters of this impact event are referenced from this object and from sub-objects, all of which contain their factual and referenced data.

Figure 1 shows a spatial presentation overview of terrestrial (meteorite) impact features resulting from the impact features container. The spatial presentation is using a Robinson projection in order to cover arbitrary locations with a continuous visualisation in a common way.

The multi-disciplinary knowledge resources were used to create various computational views of impact craters on Earth [3] with any more details. The multi-disciplinary views, including conceptual knowledge represented by classifications, enable an association of various characteristics common with different information in collections [4].

Table I lists the factual container data used from the LX Foundation Scientific Resources [5] (not an acronym) referenced for the Kaali crater field object and relevant with the mining challenge.

<table>
<thead>
<tr>
<th>Crater Number</th>
<th>Coordinates (lat/lon)</th>
<th>Diameter (m)</th>
<th>Elevation (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>58.371270 22.664737</td>
<td>39</td>
<td>24.10</td>
</tr>
<tr>
<td>2</td>
<td>58.367407 22.672298</td>
<td>25</td>
<td>25.90</td>
</tr>
<tr>
<td>3</td>
<td>58.366556 22.677637</td>
<td>76</td>
<td>21.99</td>
</tr>
<tr>
<td>4</td>
<td>58.371982 22.675092</td>
<td>33</td>
<td>24.91</td>
</tr>
<tr>
<td>5</td>
<td>58.370815 22.675611</td>
<td>20</td>
<td>21.90</td>
</tr>
<tr>
<td>6</td>
<td>58.370861 22.663155</td>
<td>13</td>
<td>29.90</td>
</tr>
<tr>
<td>7</td>
<td>58.370306 22.687148</td>
<td>26</td>
<td>22.90</td>
</tr>
<tr>
<td>8</td>
<td>58.367460 22.675277</td>
<td>15</td>
<td>25.99</td>
</tr>
<tr>
<td>9</td>
<td>58.372715 22.669419</td>
<td>110</td>
<td>34.14</td>
</tr>
</tbody>
</table>

The crater field consists of 9 known craters. Crater number 9 is the major crater. Craters 1 to 8 form sub-container objects, which deliver the data.

In order to illustrate general facilities with modified Cogwheel Modules, information peeling and milling, even for case studies with different knowledge resources, we can take a look into the context and quality of the data involved in this case.

The factual knowledge criteria for impact crater classification on basis of a physical view (criteria 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,
- Physical attributes, e.g., porosity, of 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.

For creating Cogwheel Modules and enabling views, factual knowledge not only contains facts like measurements and documentation. Factual knowledge supports analysis and visualisation, e.g., comparing knowledge objects and creating a spatial distribution and visualisation.

2) Conceptual knowledge: Advanced knowledge from integration of universal classification and spatial information can provide new insights when applied with knowledge mining [6]. The use of the Universal Decimal Classification (UDC) is widely popular, e.g., in library context, geosciences [7], and mapping [8] as provided by the Natural Environment Research Council (NERC) [9] via the NERC Open Research Archive (NORA) [10].

The small 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) [11] released by the UDC Consortium under the Creative Commons Attribution Share Alike 3.0 license [12] (first release 2009, subsequent update 2012).

Data in the knowledge resources carries references to classifications. Examples are references to UDC for any discipline and object, e.g., natural sciences and history.

Here, besides the central UDC:539.63 (impact effects) and UDC:539.8 (other physico-mechanical effects), referred top level groups for geodesy, cartography, and geography are UDC:528 [13], UDC:910 [14], and UDC:912 [15]. Tables II and III show excerpts of the conceptual data (UDC) used for geodetic/cartographic and geographic classification.

### Table II. Classification with Knowledge Resources: Geodetic and Cartographic Conceptual Data (LX).

<table>
<thead>
<tr>
<th>UDC Code</th>
<th>Description (English, excerpt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDC:5</td>
<td>MATHEMATICS, NATURAL SCIENCES</td>
</tr>
<tr>
<td>UDC:52</td>
<td>Astronomy, Astrophysics. Space research. Geodesy</td>
</tr>
<tr>
<td>UDC:528.5</td>
<td>Geodetic instruments and equipment</td>
</tr>
<tr>
<td>UDC:528.7</td>
<td>Photogrammetry: aerial, terrestrial</td>
</tr>
<tr>
<td>UDC:528.8</td>
<td>Remote sensing</td>
</tr>
<tr>
<td>UDC:528.9</td>
<td>Cartography. Mapping (textual documents)</td>
</tr>
</tbody>
</table>

Composite classification based on these top level classification references can refer to special items, e.g., cartographic bibliographies, historical atlases, and globes. Summarised, the classification can be used as glueing component classifying the knowledge object space and the implementation space, e.g., respective resources, objects, application components, and features of application components. This also provides the base for the creation of conceptual knowledge objects.

For creating views, conceptual knowledge not only provides a universal system of knowledge space, it contains classification and allows context references. Conceptual knowledge can provide a range of precise as well as fuzzy context for knowledge objects. Especially, conceptual knowledge allows the creation of conceptual knowledge objects. For example, impact features and meteorites can be classified in the following groups.

Table IV shows conceptual data (UDC) used for the basic classification of impact events and meteorites.

### Table IV. Classification with Knowledge Resources: Impact Events Knowledge Resources Classification (LX).

<table>
<thead>
<tr>
<th>UDC Code</th>
<th>Description (English, excerpt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDC:500</td>
<td>Natural sciences</td>
</tr>
<tr>
<td>UDC:523</td>
<td>Solar system</td>
</tr>
<tr>
<td>UDC:523.68</td>
<td>Meteors. Meteoroids. Meteorites</td>
</tr>
<tr>
<td>UDC:530</td>
<td>Physics</td>
</tr>
<tr>
<td>UDC:539</td>
<td>Physical nature of matter</td>
</tr>
<tr>
<td>UDC:539.63</td>
<td>Impact effects</td>
</tr>
<tr>
<td>UDC:539.8</td>
<td>Other physico-mechanical effects</td>
</tr>
</tbody>
</table>

The excerpt also shows the context of meteorites and impact effects in UDC: 5.

An object carousel generated for impact craters, shows the different types present in the knowledge resources groups and their crater categories (Figure 2). For the task of creating a carousel all categories are selected (red colour). The resulting categories are micro crater, multi-ring crater, elongate crater, complex crater, and simple crater.

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. According conceptual knowledge “filters” have been applied to the other criteria like geological time types and sub-types.
Further, supportive components can be Google Earth or Google Maps presentation [25], Marble [26], and Open-StreetMap (OSM) [27], [28].

For creating views, supportive data and component resources can provide data and features, which allow to refer to different context and add different kind of interactivity.

C. Peeling and milling of context references for views

Advanced analysis of research data is becoming increasingly important. For example, services supporting researchers especially for categorising texts with a special context are in development for many years [29]. Nevertheless, these services do not provide features beyond term context and text analysis.

The knowledge resources can fully support context and provide references to multi-disciplinary knowledge, e.g., photo media objects related to an object (Figure 3).

Regarding the knowledge mining process all categories can be used. After finding results with a possibly high relevance the categories provide further information for context and analysis.

B. Supportive data and component resources

In this case, referring to spatial distribution and distance, supportive data and component resources are geoscientific data and mapping components.

Appropriate data was required for the criteria, which are related to topographic data. In the past, the georeferenced objects have been used with various data, e.g., with the Global Land One-kilometer Base Elevation Project (GLOBE) [16] and the 2-minute gridded global relief data (ETOPO2v2) [17].

For the required resolution of the results presented here, the knowledge resources had to be integrated with data based on the gridded ETOPO1 [18] 1 arc-minute global relief model data [19]. For special purposes data can be composed from various sources, e.g., adding Shuttle Radar Topography Mission (SRTM) data [20] from the Consultative Group on International Agricultural Research (CGIAR) [21].

The horizontal datum of ETOPO1 is World Geodetic System geographic, which was established in 1984 (WGS84) and later revised. The WGS84 specifications and references are provided by the National Geospatial-Intelligence Agency (NGA) [22] and as EPSG:4326 from the European Petroleum Survey Group Geodesy (EPSG) [23]. The vertical datum of ETOPO1 is “sea level”. The source elevation data were not converted by the authors of ETOPO1 to a common vertical datum because of the large cell size of 1 arc-minute.

The Generic Mapping Tools (GMT) [24] suite application components are used for handling the spatial data, applying the related criteria, and for the visualisation.
The referenced citation entries are the result of the information peeling process from the Kaali crater object and refer to bibliographic references for meteorite craters on the island of Saaremaa [30] as well as to meteorite craters in Estonia [31].

Other references point to information for meteorite-material-usage, e.g., in context with archaeological and historical or mythical context.

One example is King Arthur’s sword Excalibur (‘Ex-Kali-bur’) [32], which is directly associated with Kali and the mother goddess Kali and its metal material. An association exists via metal object classification and “sword” synonyms (Figure 6).

![Figure 6. Synonyms of ‘cutter-sword’ group from knowledge resources objects (LX, excerpt).](image)

The association links to King Tutankhamun’s ‘dagger’ in Egypt [33], which is made with meteorite iron from impact craters in the Libyan desert, as proved by available modern analysis.

This reference shows a remarkably comparable set of facts and references (king, sword, meteorite, iron, impact, ...) for which we still have the authentic realia object.

D. Workflow implementation and phases

For the case study, the required data and configuration is manually selected for the preparatory phase. The consequent modules act on basis of that data, especially conceptual knowledge and factual knowledge.

The central Cogwheel Module cogwheel_criteria in the knowledge mining gearbox utilises a sequence lximpactsselect_crae_criteria containing a number of components

1) lximpactsselect_crae_date
2) lximpactsselect_crae_confirmed
3) lximpactsselect_crae_age_historic
4) lximpactsselect_crae_diameter

for handling the criteria for the event date range, confirmed and not confirmed events, the date range, and the crater diameter. In this case the components can be considered as filter processes.

The spatial modules of the workflow (cogwheel_world, cogwheel_region) utilise the features latitude and longitude, wet/land criteria, criteria evaluation, spatial distance computation, map projection, and visualisation.

The respective components are provided by GMT suite applications, especially pscoast and gmtselect. The GMT applications have to care for longitude, latitude, elevation and contribute to the applying topographical data related criteria, for topography related decision making within the information object turnaround.

The later association of knowledge objects, referenced media objects, and citation objects is supported by conceptual knowledge and discovery processes. In the consecutive phase results are analysed and persistently documented in order to improve the knowledge resources and mining algorithms.

Please keep in mind that it is not the intention of the examples that others should repeat the case study and its modules but with realising the details required they can create modules for their own knowledge scenarios, based on the methodology using the named or their own, additional components.

V. Secondary context and resulting views

Earths’ impact crater objects from the classified LX factual knowledge resources are used as a factual and conceptual knowledge source for computing results, considering the respective context and selection criteria. Result can be a group of craters, fitting to all the criteria, after the mining algorithm is applied to the integrated knowledge resources and methods.

The following sections describe the creation of secondary views and possible new insights based on the Cogwheel Modules Methodology and provide a discussion of the above implementation case study with its resulting primary context.

A. Result of implemented workflow

Figure 7 shows the resulting output, including the necessary topography (longitude, latitude, elevation), data, and information used, after the result was visualised via GMT.

Criteria for decision making are the resulting target structures (meteorite craters) on land (topography and coverage), especially confirmed Earth crater groups (meteorite impact features, bullets, red, blue, and green colours), age and size of (on-land) structures, and a reasonable catchment area for Europe (blue).

A catchment center has been chosen, a circular area with a respective radius of 3000 km, automatically fitted with the map projection. The blue circle marks a reasonable area to cover the continent of Europe in this context. The blue and green bullets mark the craters inside that area. The data, items, and marks are automatically computed and visualised.

The final resulting object (bullet, green colour), which fits all criteria is the Kaaali crater field, Saaremaa, Estonia. This result is based on a large amount of knowledge resources and application resources in the preparatory phase, an advanced gearbox with compute intensive Cogwheel Modules, and a workflow implementation using a range of large supportive data and component resources, as described. Further analysis can, e.g., select a relevant area containing the resulting object in order to create additional context with the object itself.

Another strategy can be to find comparable objects and context, which were outside the range when having first phrased the question.

The region of positive final result of the applied knowledge mining is computed and presented via GMT, too. Figure 8 shows the region of the Kaali crater field on the island of Saaremaa, Estonia in its topographic context.

2017, © Copyright by authors, Published under agreement with IARIA - www.iaria.org
The bullet and the cross mark the center of the crater field (labeled Kaali Crater 9). The yellow ring marks an area of 25 km around the major crater.

**B. Resulting spatial description**

Arbitrary different representations can be computed and generated from the result matrices. It is possible even to generate many different types spatial descriptions.

The Keyhole Markup Language (KML) is an Extensible Markup Language (XML) based format for specifying spatial data and content. It is considered an official standard of the Open Geospatial Consortium (OGC).

The KML description can be used with many spatial components and purposes, e.g., with a Google Earth or Google Maps presentation [25], with a Marble representation [26], using OpenStreetMap (OSM) [27] and national instances [28] in order to create arbitrary context.

Figures 9, 10, and 11 show the complementary excerpts from KML data generated for the results of the discovery with this case study.

The excerpts contain the objects of the Kaali crater field, Saaremaa, Estonia. In detail, the first excerpt holds the top part of the generated KML.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2" xmlns:gx="http://www.google.com/kml/ext/2.2">
  <Document>
    <!-- (c) CPR, LX-Project, 1992 to 2016 -->
    <name>Kaali Meteor Crater Field</name>
  </Document>
</kml>
```

Figure 9. KML data top (excerpt) generated for results of the discovery from factual knowledge (LX): The Kaali crater field, Saaremaa, Estonia.

It contains the formal configuration, e.g., the XML version, the encoding, and the KML schemes to be used in addition with a general name for the generated spatial description. The middle part contains most of the factual data, which was compiled during the preparatory phase, the knowledge mining, and the consecutive phase.

It includes the major and minor crater groups with their coordinates and elevation and also includes balloon style label popup information.
The third excerpt holds the bottom part of the generated KML with the range markers. The ellipses mark the location of longer passages of data generated for the KML code, which repeat comparable entries and entities but which are not relevant for the demonstration here.

C. Resulting associated information: Spatial mapping

The resulting satellite view shows the area of the Kaali crater field, Saaremaa, Estonia (Figure 12). Besides the major crater, further features of the crater field are not immediately visible. The reason is that the features are small in relation and they can be hidden from the satellite view, e.g., under vegetation.

The integrated knowledge from different context can deliver relevant information. For example, topography, elevation data, vegetation coverage, water bodies, infrastructure information are important information, which can be used in context with the knowledge mining.

The final result from the knowledge mining with the classified LX factual knowledge can be projected onto online satellite data of the area of the Kaali crater field. The result from object and sub-objects is shown in Figure 13.

The interactive map shows the nine craters known for the crater field. The major crater is marked in red colour, the minor craters are marked in green colour.

The final result from the knowledge mining with the classified LX factual knowledge can be projected onto online vector and navigation data (Figure 14).
Figure 14. The resulting area of Kaali crater field, Saaremaa, Estonia, factual knowledge (craters 1 to 9) (LX) projected onto OSM data via Marble.

The integration shows craters 1 to 9 of the Kaali crater field area projected onto OSM data via Marble.

D. Resulting associated information: Media references

The integrated knowledge resources can contain references to any data, e.g., media objects. Media objects contain own references, e.g., classification, citations, documentation, and keywords and can therefore contribute in many ways to new insight – besides their intrinsic media content. The following photo data (Figure 15) from the media references for “Kaali crater” were delivered in association from the final result of the knowledge mining workflow.

Figure 15. Integrated media photo objects associated with the knowledge object “Kaali crater”, Saaremaa, Estonia, referring to [34] (LX resources).

The references of these media photo objects (Figure 3) are part of objects in the knowledge resources. Media results (1–5) [34] and specimen (6) photos from the Natural Sciences Specimen Archive are dated June 29, 2016.

The photos and physical samples have been taken in 2016 by the Knowledge in Motion (KiM) natural sciences and archaeology sections at the Kaali meteorite crater field on the island of Saaremaa, Estonia, during the Geo Exploration and Information (GEXI) [35] Baltic research and studies campaign.

In detail, the resulting photo objects of the examined site (from left to right, from top to bottom) show in this sort order:

1: Major crater, view in northern direction.
2: Major crater, view in north-eastern direction.
3: Major crater, view in western direction.
4: Path towards major crater, view from southern direction.
5: Vegetation, Lilium martagon, at top of crater rim (referred to Figure 4).
6: Specimen crater pond material (quartz, melane particles, lacustrine deposits, biogenic material).

The references included in the knowledge mining workflow (Figure 5) provide the complementary information that fine particles from the Kaali crater include impactor remains (esp. significant Ni-Wüstite, Ni-Maghemit, Ni-Goethite, Hematite, Magnetite, Taenite, Kamacite), spherules and splash-forms.

The analysis of the referenced media content, e.g., Lilium martagon, delivers the information that this flower is an indicator plant [36], indicating natural resources, e.g., showing mining resources. This will also show context with the references, both with impactor remains and with activities in prehistorical and historical times and associated remains and mythical context.

The media references are part of the context created for the views. These references can also be used when creating a secondary context, e.g., a spatial and dynamical visualisation, based on the results.

E. Consecutive criteria and range markers

The above resulting media references are directly referenced with the Kaali crater, especially with the major crater of the crater field. If we use further criteria, e.g., available with the spatial context and projection, we can associate additional context, e.g., Points Of Interest (POI), in the range around the Kaali crater. The generated KML can be used to express such ranges (Figures 16 and 17). The spatial algorithms and features available with the respective applications can be used to create complementary insight from individual context.

For example, further context for a primary view can be created by calculation of context considering a range. Considering range means calculating distance in a spatial representation.

The resulting perspective satellite view shows the area of the Kaali crater field, Saaremaa, Estonia, including circular range markers (Figure 16). The range markers (visible 1 km and 2 km diameter) mark an area around the major impact in the crater field. The same data is used with Marble (in this zoom the 1 km diameter range marker is visible in the window) and OSM context data (Figure 17). Not only that all the known crater structures appear to be located inside the close range around the major impact: That way the results can be analysed in arbitrary different context with the integrated knowledge, information, algorithms, and methods available from the chosen target.

Consecutive associations and further consecutive references can be computed, making use of the new context and features [37].
Figure 16. Secondary context with satellite data: Range markers at the area of the Kaali crater field, Saaremaa, Estonia (Google Earth), with range markers, major, and minor craters in a portable, interactive, and dynamical environment, presenting the resulting knowledge in context of a perspective satellite view.

Figure 17. Secondary context with street data and legends: Resulting area of the Kaali crater field, Saaremaa, Estonia (Marble, OSM data), with range markers, major, and minor craters in a portable, interactive, and dynamical environment, presenting the resulting knowledge in context of land use and transportation.
Google Earth provides satellite data and POI, Marble with OSM provides vector street and feature data and different POI data. The secondary context examples using satellite and OSM data showed the range circles and the generated selection legend with the circles from the generated KML data as well as the legends on available transportation infrastructure context and on area context, e.g., forest areas, farm areas, residential areas, and lake reservoirs.

The spatial context provides an arbitrary number of thematic knowledge, which can be integrated with the results in order to compute new views and insight. The applied components allow to combine an arbitrary number of methods. For example, perspective visualisations can give additional information and enable to create references in that “spatial-range” context of integrated knowledge KML based animations can allow flightover animations in the context of the associated thematic knowledge. These are just two of many examples for the spatial context only.

The region referenced fuzzy contextualisation for spatially expressed thematic context, e.g., weather and climate map data [38], can, e.g., also make the method beneficial for businesses like planning agencies and insurance companies.

These are the core items of the methodology implemented for this case study, from preparatory phase and knowledge mining to the consecutive phase with the analysis of results. There are no optional items, which should be described in detail, as the fine-tuning always depend on what the implementor wants to achieve for a certain task.

VI. Conclusion

Creating context and views for gaining insights from content of knowledge resources is a most challenging task. This research successfully deployed the Cogwheel Modules Methodology for advanced knowledge mining and generating data, especially for knowledge integration and the goal of creating new views, which lead to new insights and cognition.

The methodology case study implementation showed how primary context and secondary can be created and how the views can be visualised. The examples showed a series of newly generated spatial views, which open a wide new context of complementary dimensions, which can further be used to associate additional references. The case study also showed that this way the context can be extended very efficiently and that new insight can be the result, which was out of scope before.

The implementation proved that the methodology can be an excellent support for advanced knowledge mining. Generating views also means using general and specialised tools, which allow to add new knowledge from resources.

As soon as views proof to extend a context efficiently the process can be used for creating automated learning processes and saving the views with long-term knowledge resources for future use. Besides the practical benefits for knowledge mining the methodology also contributes to the further development of multi-disciplinary knowledge resources.

Creating spatial views is one of an arbitrary number of possible applications. The major phases of the methodology were applied for the implementation. Nevertheless, creating spatial Cogwheel Modules with spatial components and multi-disciplinary knowledge from knowledge resources demonstrated the methodology in a very instructive way.

The paper provides the results from the research and data-centric implementation of a case study of integrated knowledge and methods for answering knowledge mining challenges like complex questions and a number of instructive examples for creating primary and secondary context views.

The case study is focussing on knowledge mining challenges associated with geosciences and archaeology. Therefore, one category of the relevant generated context is spatial context, implemented in modules for spatial analysis and visualisation.

The base of the view creation is the identification and mapping of required resources – knowledge resources and partial solutions, mapping of complementary components in their context, and excerpts of associated knowledge used for information peeling generating a base for the information processing. The resources provide conceptual and factual knowledge in integration with appropriate context data and application components for computing and visualisation.

The mapped application components – tools and filters – were used complementary for handling the complex resources, systematically peeling of information nuclei and facets, milling, and consecutive information processing, including decision making integrating spatial and conceptual criteria. The results of the knowledge mining information object turnaround, can itself become part of the knowledge resources.

The methodology and the view creation can be applied to many application scenarios, especially where a solution can only be gained by integration of different data and approaches. Examples are multi-disciplinary knowledge mining scenarios integrating natural sciences and archaeology. Comparable reasonable results are not possible with any tested services, e.g., even large search engines accessing data in depth and width of the knowledge spectrum.

The various approaches also provide potential for optimisation for special priorities. In most cases, the optimisation can consider the individual challenges and the use of special algorithms and applications.

Future work concentrates on analysis of complementary context features, beyond spatial views, and further improving the long-term multi-disciplinary knowledge resources. On module side for knowledge mining the creation, utilisation, and documentation of advanced components with the Cogwheel Modules is in focus.

Acknowledgements

We are grateful to the Knowledge in Motion (KiM) long-term project, Unabhängiges Deutsches Institut für Multidisziplinäre Forschung (DIMF), for partially funding this implementation, view modules case study, research and studies campaign, and publication under grants D2015F1P04604, D2017F1P04708, and to its senior scientific members, especially to Dipl.-Biol. Birgit Gersbeck-Schierholz, Leibniz Universität Hannover, Germany, to Dr. Friedrich Hülsmann,
Gottfried Wilhelm Leibniz Bibliothek (GWLB) Hannover, Germany, and to Dipl.-Ing. Martin Hofmeister, Hannover, Germany, for fruitful discussion, inspiration, practical multidisciplinary case studies, and the analysis of advanced concepts. We are grateful to Dipl.-Ing. Hans-Günther Müller, Cray, Germany, for his work for providing flexible practical solutions to architectural challenges and excellent technical support. We are grateful to all national and international partners in the Geo Exploration and Information cooperations for their constructive and trans-disciplinary support. We are grateful to the postgraduate participants of the European Legal Informatics Study Programme (EULISP) for prolific discussion of scientific, legal, and technical aspects and for feedback, contributions, and reuse of results regarding High Performance Computing and Return of Investments projects and case studies over the last ten years. We thank the Science and High Performance Supercomputing Centre (SHPSC) for long-term support for the associated research cooperations.

REFERENCES


