Deliverable D3.3

Transforming data and information into knowledge within the MICA project

Project:	Mineral Intelligence Capacity Analysis
Acronym:	MICA
Grant Agreement:	689468
Funding Scheme:	Horizon 2020
Webpage:	www.mica-project.eu
Work Package:	Work Package 3
Work Package Leader:	Evi Petavratzi
Deliverable Title:	Transforming data and information into knowledge within the MICA project
Deliverable Number:	D3.3
Deliverable Leader:	NERC (BGS)
Involved beneficiaries:	
Dissemination level:	PU Public
Version:	Final
Status:	Submitted
Year:	2017
Authors:	Evi Petavratzi
Reviewed by:	Tom Bide
Approved by:	Erika Machacek, Kisser Thorsøe

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 689648.

Notice

TABLE OF CONTENTS

PURPOSE

Deliverable 3.3 explores how raw materials data and information investigated by the MICA project can be used to deliver knowledge and support mineral intelligence. This report provides a synthesis of work undertaken within WP2, WP3 and WP4, which assists in delivering knowledge on raw materials to various stakeholder groups. The key purpose of this report is to define the terms of data, information and knowledge within the raw materials context and describe the steps required to derive desirable knowledge and justify data/information needs.

Raw materials knowledge is delivered by the European Raw Materials Intelligence Capacity Platform (EU-RMICP), mainly through the form of the MICA flowSheets, factSheets and docSheets, but also by the project deliverables that provide knowledge independent of the platform addressing questions of relevance for mineral intelligence, and mineral policy. D3.3 presents the development of a knowledge management model, the Data-Information-Knowledge-Intelligence (DIKI) model for raw materials.

EXECUTIVE SUMMARY

This report is concerned with exploring how data and information are transformed into knowledge for raw materials. It delivers a conceptual framework, which explains the transformation process. It also produces clear definitions of the terms data-information-knowledge-intelligence, which constitute key components of the MICA project, as well as the EU Raw Materials Knowledge Management Base. Reference to background literature used as inspiration for the development of this model is made.

Several examples from the raw materials field and in particular the MICA project are given to demonstrate the applicability of the proposed model and explain the transformation process. Discussion of the term raw materials intelligence also takes place.

DELIVERABLE REPORT

1. Introduction

The ‘DIKI’ (Data-Information-Knowledge-Intelligence) model was presented in D.3.1 *Draft inventory of data on raw materials*, a previous WP3 deliverable, however without extensive explanation about its origin or how it could be used to explain the development process of raw materials knowledge.

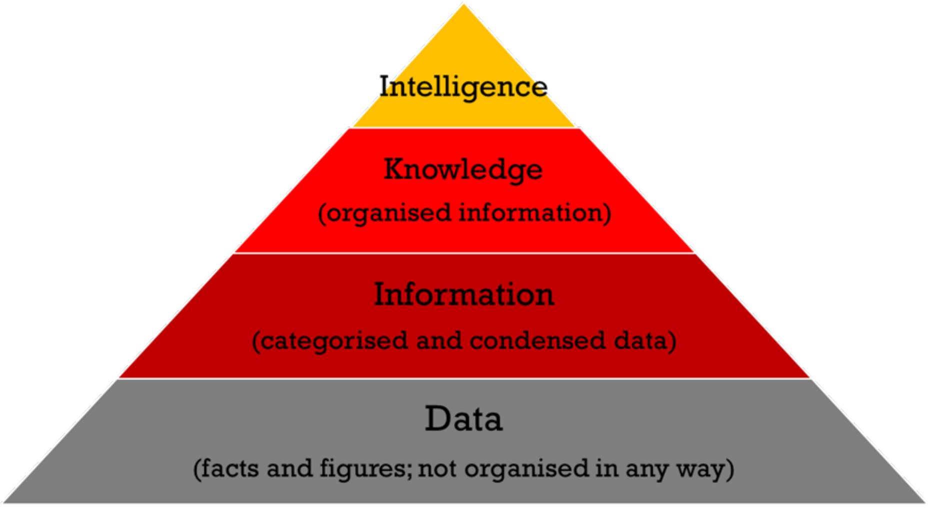

Figure 1 DIKI model (MICA D3.1).

The MICA project is at the centre of all the above, as it delivers a ‘system’ that allows a diverse range of stakeholders interested in raw materials to identify data, information, methods, tools and knowledge to support their needs, which can later on lead to raw material intelligence. The aforementioned terms, namely, data, information, knowledge, intelligence are currently misinterpreted or used interchangeably by stakeholders, which suggests that clarity of what such a ‘system’ actually delivers is diminished, or expectations from stakeholders are not met. The development of a tailored raw materials DIKI model is expected to provide the underlying explanation of the meaning of those terms and the role they play in EU-RMICP. It also investigates how essential components of what is conceived as knowledge (e.g. data and information) may be combined together to produce knowledge maps (flowSheets) that stakeholders can use to satisfy their needs.

2. Knowledge management – brief introduction

Knowledge is a critical resource and an asset for individuals, sectors, economies and companies (Foucault and Gordon, 1980). Knowledge is power and maintenance and enhancement of knowledge promotes innovation. Our societies are 'flooded' with knowledge at present and there has been a lot of work undertaken on the diffusion, sharing, generation and use of knowledge. Various systems, such as databases, portals, networks, libraries and so on are used to support the creation, use and sharing of knowledge. Technology progression has facilitated this move, for example computers, software tools and nowadays artificial intelligence are used to generate new knowledge or to share it across the globe. It has also assisted accommodating its dynamic nature, namely the continuous change in our understanding, as well as the change of the requirements for access to knowledge.

Knowledge management is the discipline concerned with understanding how knowledge is generated and conveyed to the right person to satisfy a need. The first mention of the term was in mid-1970s, including the first definitions and distinction between data, information and knowledge (Henry, 1974). Knowledge management made a presence as a distinguished discipline from 1990s onwards (Metaxiotis et al., 2005). Its main purpose has been to achieve improvements in business areas, such a quality management, and business process modelling, human resource management and information systems. Conceptual frameworks around knowledge management however are relevant to any other action, discipline, sector, project or organisation that generates, uses and disseminates knowledge.

3. Definitions of terms in the historic DIKW model

The 'Data-Information-Knowledge-Wisdom' model or as it is often referred to the 'DIKW pyramid', the 'DIKW hierarchy' or 'Knowledge hierarchy' (Figure 2) is widely used by the disciplines of knowledge management and information science. It is often used as a way to define the terms data, information, knowledge and wisdom and to explain the process of moving from data into wisdom (Rowley, 2007).

Several authors referred to the hierarchy and gave their own perspective and definitions and many more continue to do so (Henry, 1974, Zeleny, 1987, Ackoff, 1989, Davenport and Prusak, 1998, Rowley, 2007, Frické, 2009). A few authors (Hey, 2004, Sharma, 2008) suggest that the first mention of the hierarchy was in the poem "The Rock" (Eliot, 1934). Other citations by artists, well before knowledge management and information science showed an interest in this classification are available(Zappa, 1979).

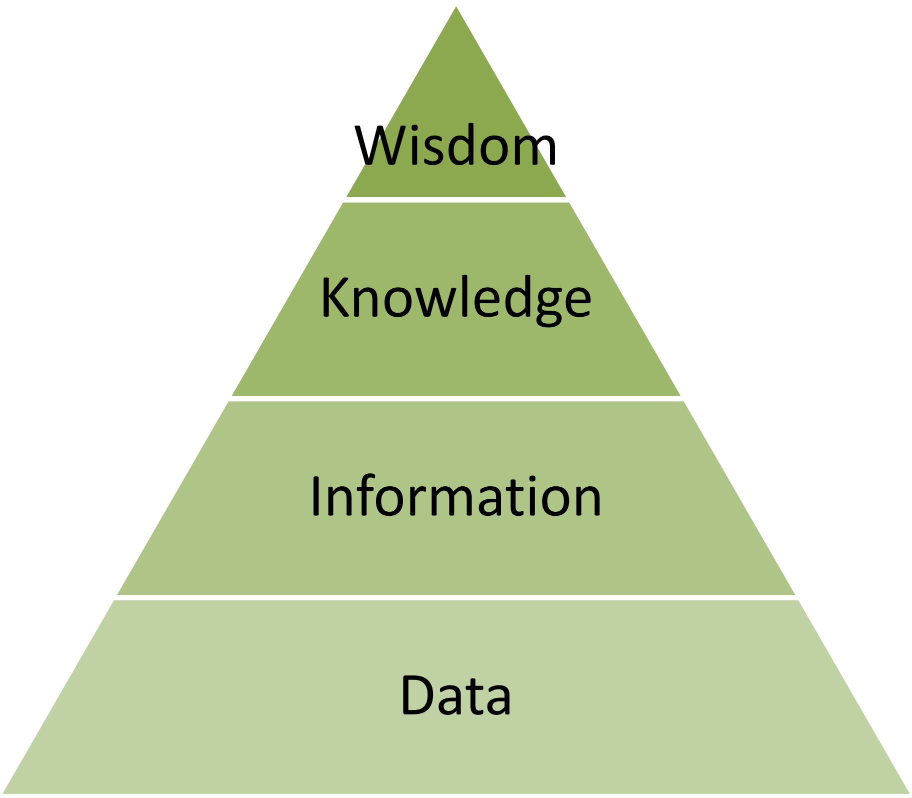

Figure 2 The DIKW model.

> *Where is life we have lost in living?*
> *Where is the wisdom we have lost in knowledge?*
> *Where is the knowledge we have lost in information?*
>
> T.S. Eliot (1934), "The Rock"

> *…Information is not knowledge*
> *Knowledge is not wisdom*
> *Wisdom is not truth*
> *Truth is not beauty….*
>
> Frank Zappa (1979)

The definitions of data, information and knowledge have been a subject of interest for many researchers and even though there have been several publications trying to distinguish between the terms, they still remain abstract and used in different context. For the purpose of this report the definitions given below represent the most widely accepted ones, which are in line with the concept of the MICA project.

Data

According to Davenport and Prusak (1998), 'data is a set of discrete, objective facts about events'. The English Oxford Dictionary states that 'data are facts and statistics collected together for reference or analysis'. The term originated from philosophy and it is defined as 'things known or assumed as facts making the basic of reasoning or calculation'. A variety of other definitions exist, depending on the context of its use.

Data are not structured, they do not convey any meaning and there are no built relationships between them. Nowadays, data are generated by many different processes and they are often referred to as 'raw data'. For example, statistics offices collect data during the different surveys they undertake.

Data are essential to all organisations and many industries are heavily dependent on them. They represent the foundation of any decision making process, but in reality data do not provide any judgment or interpretation or suggestions for action. Also, data is not known to be true (Frické, 2009). They may include mistakes and errors turning them into invalid and wrong data. However, the identification of incorrect data requires some degree of interpretation, which is outside the definition of data. The gathering of data is the essential process for making data available. Very often the principle followed is that of 'the more data the better', but this is not always valid. In many cases too much data can inhibit interpretation and after all there is no inherent meaning in them. Data are stored in some form of technology system, for instance a database or another data management system (Davenport and Prusak, 1998).

Information

Information is data put in context. This can only be achieved through interpretation and analysis of data to add meaning to them. Relationships between data and associations or patterns identified are all part of the interpretation and analysis process that result to the birth of information.

The role of information is to "inform", namely to provide insight into a subject of interest (Davenport and Prusak, 1998). This implies that whether the content provided has succeeded its purpose is defined by the person that receives it. If it does then it is considered as information or otherwise it may be judged as noise or an irrelevant 'message'. Information is commonly found in answers to 'who', 'when', 'where', 'what' and 'how many' questions (Ackoff, 1989).

Knowledge

Knowledge is neither information nor data. Knowledge signifies things known. The Oxford Dictionary defines knowledge as 'facts, information, and skills acquired through experience or education; the theoretical or practical understanding of a subject'. Several other definitions have been produced from various disciplines, without any consensus about which one constitutes a good one. For the purposes of this document the definition given by Davenport and Prusak is considered suitable: '*Knowledge is a fluid mix of framed experiences, values, contextual information and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of knowers. It often becomes embedded in documents or repositories, but also in organisational routines, processes, practices and norms'.*

The above definition and research of Davenport and Prusak distinguishes data and information from knowledge. They suggest that knowledge derives from information and information from data. Humans are responsible for transforming information into knowledge by comparing them, assessing their implications, exploring their relationships and identifying what the opinion of others are on specific information (Davenport and Prusak, 1998).

Two types of knowledge are often referenced in the literature, which are also apparent in the above given definition: explicit knowledge and tacit knowledge. Explicit knowledge is tangible and captured in documents, recordings or images. Tacit knowledge however is more difficult to define as it often sits with the knower and is not easily transferable. It is the know-how or expert insight. Explicit knowledge often represents the final outcome, but very rarely it provides insight about the actual research process undertaken to get to an outcome (Dalkir, 2005). Research in knowledge management suggests that only 15 to 20% of the valuable knowledge has been captured and is available through a tangible form. The remaining 80% is often in the form of tacit knowledge that remains with individuals. The need therefore for systems that mobilise tacit knowledge to create explicit knowledge is important (Dalkir, 2005).

Knowledge conveys expectations, instructions and rules and it often provides answers to 'how-to' questions (Ackoff, 1989, Boisot and Canals, 2004). Knowledge is closer to action than data and information and this is the main reason why it is seen as valuable. The process of how knowledge is generated is complex and dynamic. Experience, truth, judgment and guidance provision are key components of knowledge and are interlinked in complex ways (Davenport and Prusak, 1998).

Data and information are visual and quantifiable elements, whereas knowledge is something we know and it requires analytical thinking.

Wisdom

Wisdom is found at the top of the DIKW model. Wisdom is a much more complex term to define; it is multi-dimensional and requires diverse knowledge. Wisdom often represents the overall picture (Barlas et al., 2005). It combines knowledge to formulate actions and achieve goals. Wisdom therefore aims to increase effectiveness, which requires the ability to make judgments (Ackoff, 1989).

This deliverable report deals primarily with the data, information and knowledge stages of the DIKW model, as these are directly relevant to the MICA project and against which specific outputs are delivered.

4. The raw materials 'Data–Information–Knowledge–Intelligence' model

The DIKW model presented in the previous paragraph forms the basis for the proposed DIKI model for raw materials. In this section we look at the definitions given for the historic DIKW model and explore their content and meaning for the raw materials sector.

Data for raw materials

The definition of data in the DIKW model fits well with raw materials data too. Raw materials data are often found in two forms:

- Spatial data or geographic data, which define a location on the surface of the earth. Spatial data usually represent points, polygons, lines or pixels. They include a location, shape, size and orientation and they are therefore multi-dimensional. They are used to identify the geographic location of features and boundaries on Earth, such as man-made constructions, or natural ones (e.g. geological deposits or oceans). The synthesis and interpretation of spatial data results in the production of various maps.
- Non-spatial data relate to a specific, defined location. Statistical data are often found in this category, but text, images and multimedia data are also part of this group. One would argue that statistical data on raw materials should fall under the spatial data category rather than the non-spatial one, as they are influenced by location and they may even represent a spatial data interpretation, such as for example the statistical data on mineral resources and reserves. However, it is useful to distinguish statistical data from spatial data, as the latter are often used for the development of maps and geoportals. Statistical data on the other hand are often used to describe a process, for example aspects of the physical economy, or environmental impacts associated with the production of raw materials.

Data for raw materials are essential for many different organisations, such as for public authorities, the industry, NGOs, research and the general public. They are therefore collected from various sources across the globe and through a variety of routes, including surveys, field work and experimentation. They are commonly stored in databases, in electronic documents or hard-copy documents.

Information for raw materials

Information represents structured data. As defined in the historic DIKW model, the role of information is to inform someone about something and to satisfy an information need. The generation of data for raw materials requires the interpretation and analysis of raw data. A variety of different methodologies are used during this process, depending on the scope of the information need. For example, procedures specified for statistical data often include the stages of checking, contextualising, and categorising, standardising and harmonising data. Interesting is the case of spatial data for raw materials in Europe, for which their collection and transformation into information is currently defined by the INSPIRE Directive[1]. Geospatial data on many different domains, not just mineral resources are produced by all European countries. As with mineral resources, many countries share rivers, mountains, transport infrastructure and so on, therefore ensuring that information describing them are shared was seen essential for preventing and

[1] https://inspire.ec.europa.eu/

minimising impacts and promoting effective decision making at the European level. This has led to the development of common standards to cover different domains, including mineral resources, which allow the description and sharing of spatial data. The INSPIRE Directive has encouraged the development of an infrastructure for sharing such information between public authorities in Europe.

Information is generated from many other bodies, for example from research and industry, which although invaluable are often dispersed, static (e.g. relate to a single project), and in some cases non-accessible (e.g. confidential information).

Information is stored in databases, portals, or presented in documents (e.g. reports, scientific articles). Very often data are transformed into information through model development, which represent another output for information.

Knowledge on raw materials

Knowledge as defined by the DIKW model is contextualised information and it is produced by combining available information with expertise, insight and intuition. So if we were taking, as an example, statistical data on mineral production to identify how they are transformed into knowledge, then the following associations to the DIKW model could be made:

1. the data generated from a survey represent the first level of our model 'data for raw materials'
2. the production of a mineral statistics yearbook would represent the second tier of the model, namely the 'information for raw materials'
3. understanding whether investing in a new processing plant to increase the production of a single commodity would require access to mineral statistics to understand for instance global market trends, but also access to other information, for example cost and efficiency of a new processing plant, examples of similar case studies elsewhere, commodity prices, and expert knowledge and insight of the sector which cannot be delivered by readily available information.

The example given describes a knowledge need, for which information on mineral statistics are required. The availability of such information could satisfy multiple knowledge needs, other than the one described in the example. The example showcases that knowledge requires the use and judgment of several different information types combined with expert insight. Even though explicit knowledge is possible to capture and convey to the public, tacit knowledge is quite difficult to define and disseminate. However, ongoing research in the field of raw materials has assisted to transform parts of tacit knowledge into explicit knowledge. Also, open access to a wide range of resources has externalised tacit knowledge. An example of this transformation is seen for instance in the field of critical raw materials, where the active involvement of experts, as well as collaboration between countries (US, Europe, Japan) has enhanced our understanding of the various issues surrounding critical raw materials, which further on prompted strategic interventions (e.g. changes in Mining Law in some European countries, opening of new mines producing critical raw materials). Several other examples can be identified, for instance the work undertaken for compiling national Material Flow Accounts, or actions taken by international organisations, for example the UN on monitoring progress towards global sustainability goals on resource efficiency.

Intelligence on raw materials

The term intelligence has been chosen instead of wisdom to describe the top of the raw materials DIKI proposed model as it is often employed by the minerals sector. Intelligence is associated with an action taken based on the knowledge available. Raw materials intelligence is often seen in decisions taken, such as for instance new policy development, strategy development, investment decisions, and development of infrastructure and many more. Intelligence on raw materials often requires having an overall picture of a topic of interest or concern, hence knowledge, information and data of all the individual concepts that constitute that topic. The example (Figure 3) attempts to demonstrate the different knowledge requirements that may be needed to feed into the development of a National Raw Materials Strategy. The knowledge requirements list in this example is not exhaustive. Additional knowledge would be required from different disciplines to proceed with such action.

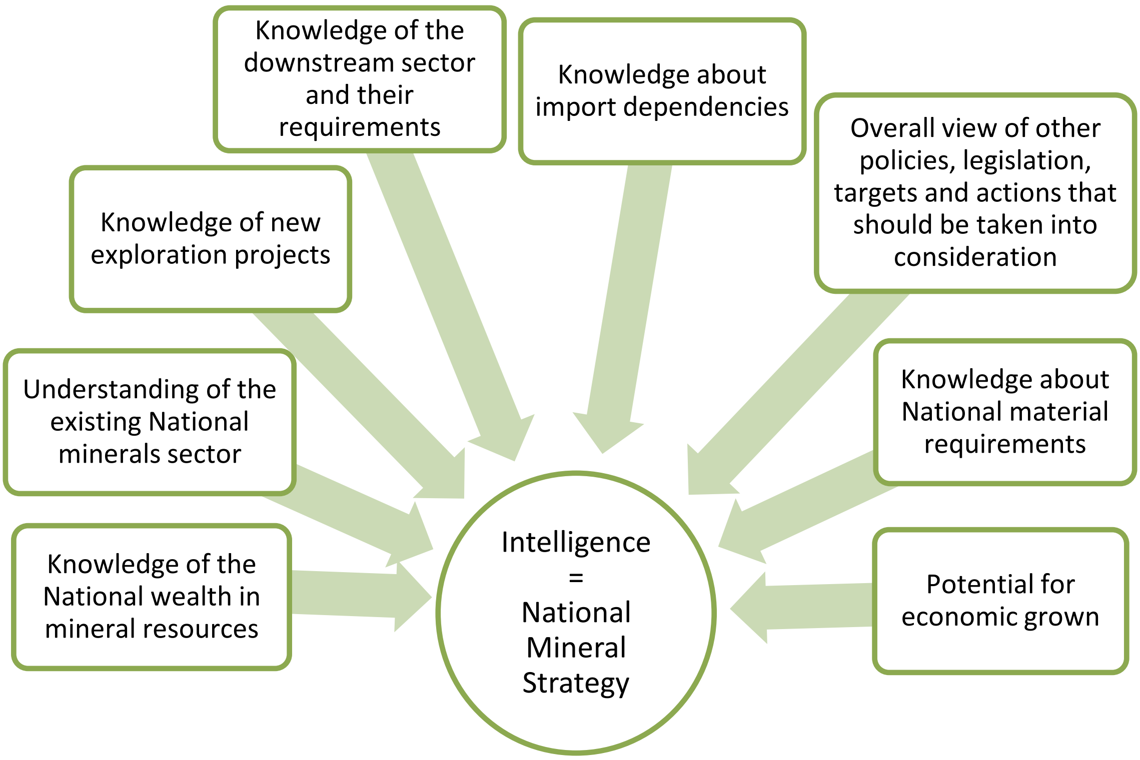

Figure 3 An example of a raw materials intelligence action demonstrating the different knowledge needs for developing a National mineral strategy (intelligence).

4.1 Discussion of the DIKI raw materials model

Based on the definitions given in the previous sections an improved DIKI model describing the data to intelligence transformation process for the raw materials sector has been developed (Figure 4). The model includes information about potential actors and needs associated with each tier of the pyramid. Actors represent different bodies that are responsible for delivering, data, and information. The 'needs' sections aim to elucidate why data, information, knowledge and intelligence are needed.

Figure 4 The proposed improved DIKI model for raw materials. Examples of Actors and Needs are included for each tier of the pyramid.

'Actors' and 'Needs' are seen as essential components of the DIKI model. Any resource, whether this is data or information or knowledge, is developed because there is a specific need for it. Without an identified need the pyramid or tiers in the pyramid are redundant. This can be the case, for example when data are generated by a process, often 'unconsciously', but are not transformed to information or knowledge. Such data do not serve any specific role and will not be part of a 'knowledge chain'. Being able to identify who the actors are is equally valuable. It provides insight about the capabilities, roles and importance of different organisations in the knowledge generation process and constitutes an essential component of a knowledge chain map.

The hierarchical representation employed by the raw materials DIKI model hides however intermediate steps that are taking place during the data to intelligence transformation process. These are:

- cyclical steps between stages, for instance knowledge may lead to requests for additional data, or there may be iterations between data and information before trends on an issue of concern is established;
- Knowledge input required for the generation of data, information or knowledge (Figure 5).

The 'hidden steps' in the proposed model are to some extent covered by the inclusion of 'actors' and 'needs' in the pyramid.

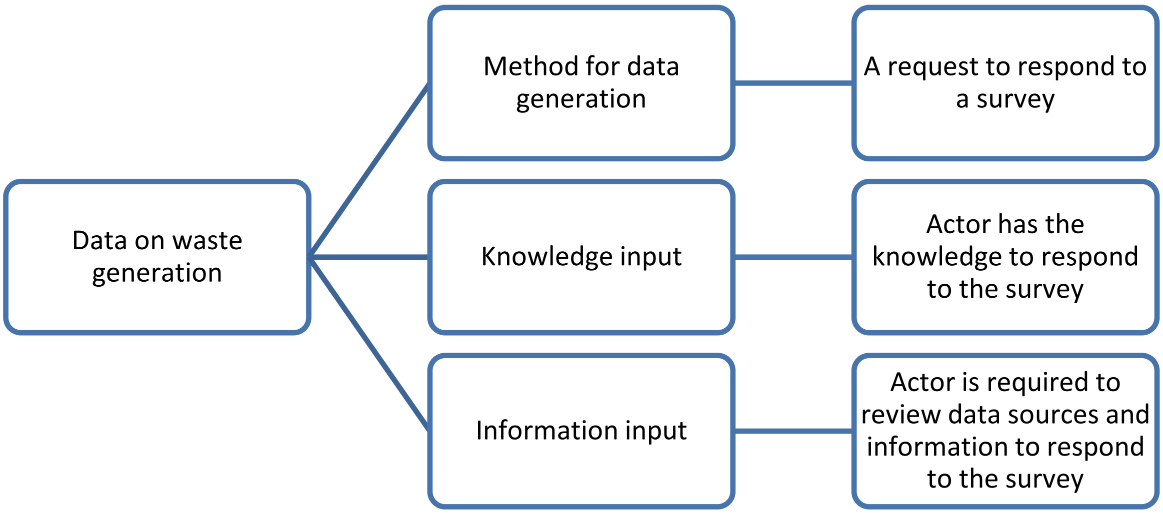

Figure 5 Example of 'hidden requirements' for information and knowledge associated with the generation of data.

Tracking the transformation process from the bottom to the top of the pyramid is not always straight forward due to the complexity of knowledge chains; however the proposed model provides a useful framework for:

- mapping knowledge chains and assessing their importance in decision making;
- identifying who the actors are and what role they play in a knowledge chain;
- identifying what the needs for resources are;
- increasing visibility and transparency across the knowledge chain and of the different steps required to get from basic data and information to knowledge and intelligence;

- connecting the different actors and informing them about the requirements of the tier above that could facilitate the development of fit-for-purpose resources;
- facilitating the generation of intelligence on raw materials by tracking who holds what knowledge in a knowledge chain.

5. MICA actions towards the population of the raw materials DIKI model

The raw materials DIKI model is in agreement with the actions of the MICA project, which aims to contribute to the EU Raw Materials Knowledge Management Base by delivering data, information and knowledge that satisfies the needs of a wide range of stakeholders. Data and information are delivered by WP3 (Data for raw materials intelligence) and WP4 (Methods and tools for raw materials intelligence) and knowledge is delivered by the EU-RMICP developed by WP6, and a range of project publications. 'Actors' and 'Needs' are identified by WP2 (Stakeholder needs for raw materials intelligence), all of this is used to inform policy and foresight, as developed in WP5, and project actions and outcomes are communicated by WP7 (among other through www.mica-project.eu).

The EU-RMICP is an important outcome of the MICA project and its components can be mapped against the DIKI model as shown below:

- Data and information are delivered by the MICA metadata inventory (Petavratzi and Brown, 2017, Petavratzi et al., 2016) and LinkSheets[2]
- Knowledge is delivered through the produced factSheets (Voet et al., 2016), docSheets, case studies, and flowSheets (Cassard et al., 2016).
- 'Actors' and 'Needs' have been investigated through extensive stakeholder engagement undertaken during the first year of the MICA project. Stakeholder requirements underpin the structure of the EU-RMICP. Stakeholders from multiple groups have been identified and classified into four different tiers depending on their power, legitimacy and interest for raw materials knowledge. An inventory of stakeholders has been delivered, which identifies their role and interests in the raw materials field (Erdmann et al., 2016; Erdmann et al., 2017).

Tacit knowledge in the MICA project

Tangible resources such as data, information and explicit knowledge are well represented in the EU-RMICP and the various MICA publications, but how does MICA deliver the 'hard to explain' tacit knowledge? There have been several approaches and attempts to capture tacit knowledge on raw materials throughout MICA, which has been conveyed in deliverable reports, as well as the EU-RMICP. In detail:

- Surveys: Prior to the kick-off meeting a survey questionnaire was distributed among the project partners to identify key topics of interest in the raw materials field for which related data should be identified and reported. The survey also identified data availability and uncertainty based on expert knowledge. The topics identified from this survey guided WP3 in the collation of related data on raw materials. Online surveys with public authorities and the industry, including a number of interviews were undertaken by WP2 during the stakeholder needs identification process. All the survey outcomes were used for the development of the MICA ontology.
- Workshops: Several workshops tried to capture tacit knowledge and blend it into the MICA project. Tacit knowledge is being delivered by the active involvement of experts throughout the lifespan of this project. For example, during the kick-off meeting a

[2] linkSheets present links to existing useful information, such as websites, online databases etc.

workshop took place that identified stakeholder needs using the expertise delivered by the MICA participants. An additional workshop took place later on with a wide range of external participants, which identified stakeholder needs beyond the consortium boundaries. Work Package 4 led a workshop that mapped methods to stakeholder needs. A combined workshop between WP3 and WP4 mapped data and methods and produced flowSheets, again using a pool of experts. Several tools were used to externalise tacit knowledge, including open discussions, group discussions, practical sessions and presentations. Most of the workshops captured the experience of experts, and allowed them also to reflect on what was achieved during the day, for instance during discussions and wrap up session. The findings of all these workshops have fed into the development of the Dynamic Decision Graph (Cassard et al., 2016) underpinning the EU-RMICP, and are also presented in various project deliverables.

- Sheets production: A wide range of factSheets and docSheets are produced in the MICA project and will become available through the EU-RMICP. factSheets represent descriptions of key methods used by experts for producing models and interpreting data. docSheets are short descriptive summaries of information on a range of topics related to raw materials. Both types of sheets are produced by experts in a field and have been written in a language that is easily understood by a wider audience. Tacit knowledge is communicated and shared in these documents.
- Metadata: Metadata included in the MICA metadata inventory are also prepared by experts and their knowledge is required in producing several of the fields, such as the abstract, completing information on data uncertainty, as well as linking the metadata to the EU-RMICP.
- Case studies: Several have been developed to explain how methods and tools can be used and what outcomes can be delivered. The case studies are produced by practitioners and researchers and they are linked to the stakeholder needs identified by MICA.
- Advisory Board: Input from the MICA Advisory Board has been useful in framing the project and identifying areas for improvement. Members of the Advisory Board represent different stakeholder groups, such as industry, research institutes and public authorities, all within the field of raw materials and they have actively participated in the workshops and meetings undertaken so far. The role of the Advisory Board members, given their influential roles in these various sectors is to reflect on the deliverables of MICA and assist in developing the project towards the direction with maximal benefit for the targeted users, and ensuring the appropriate dissemination of its findings, including beyond the project timeframe. Tacit knowledge is conveyed through exchanges with the project partners.
- MICA publications: Knowledge is not only conveyed by MICA through the EU-RMICP, but also through the various deliverable reports that are produced by the project partners. Tacit knowledge in these reports is delivered from the consortium experts, reflecting on the exchange of information and experiences gained during interactions with experts on raw materials.
- Linkages in the EU-RMICP: Metadata, factSheets, docSheets, flowSheet, linkSheets all link to the MICA ontology to enable users to access the right resources when selecting topics of

interest from the thematic domains of the MICA ontology[3]. The development of linkages is critical. The right Domains and Concepts from the MICA ontology should be selected to link the various 'Sheets' as well as the metadata. The linkages can be quite complex, for example a variety of Domains and Concepts may need to be selected for a single document or metadata record. The development of linkages requires the use of 'intuitive' expertise in the various fields, namely the use of tacit knowledge.

5.1 The development of MICA flowSheets

The term 'flowSheet' in the MICA project describes a knowledge chain relevant to a specific topic of interest or question on raw materials. The identification of stakeholder needs has developed a range of questions and topics for which knowledge is required. flowSheets are developed using the questions/topics as a starting point. A multidisciplinary group of experts is then involved in the development of a 'recipe', 'thinking process' or 'knowledge chain' for users to be followed to identify data, methods and to gain knowledge to assist them in formulating an answer. The involvement of a group of experts enables the investigation of a question/topic from different viewpoints and facilitates the dissemination of knowledge, both explicit and tacit. The end result is a flowSheet describing the 'thinking' steps that someone ought to follow accompanied by tangible resources such as data, information and methods, as well as qualitative information, discussions, expert knowledge and gaps.

FlowSheets are a direct reflection of the DIKI model, as they provide data, information and knowledge to users, namely the key ingredients and recipe to allow them to take informed actions (intelligence). MICA flowSheets were developed during two combined WP3 and WP4 workshops in Paris (15/06/2017) and in Leiden (10-11/10/2017), which involved an interdisciplinary group of MICA experts and external stakeholders. An example of a flowSheet produced during this workshop is provided in the following section.

MICA flowSheet Example

The following stakeholder question was explored by a group of experts:

What is the future demand and supply of zinc in 2050 at the global level?

The flowSheet development consists of the following steps:

Step 1: Interpretation of the question and translation into a precise one with identified boundaries

The question given to the expert group was considered imprecise as it does not specify who asks the question and why (actors and needs in the DIKI model). As explained in previous sections of this report, different actors have different knowledge needs, but unless these two elements

[3] There are 7 thematic domains in the MICA ontology and each domain includes multiple levels of concepts (sub-domains). The 7 domains are: 'Primary Mineral Resources', 'Secondary Mineral Resources', 'Industrial Processing and Transformation', 'Raw Materials economics' (including CRMs), 'Raw materials Policy & Legal Framework', 'Sustainability of Raw Materials' and 'International Reporting'.

become clear from the start, the knowledge delivered may not be fit-for-use. Taking this in mind, several alternative questions that reflect different points of view may be hidden behind the given question (Table 1).

Table 1 Examples of underlying questions that may refer to the original question depending on 'who' is asking and 'why'.

Who asks the question? (Actor)	Point of view (Need)	Alternative question
Steel producer	Supply chain disruption	Are zinc supplies sufficient to meet the future demand?
Investor	Economic impact	What is the zinc price likely to be in 2050?
Exploration company	Geological availability	Should we explore for more zinc deposits?
Materials researchers	Technology impact	If zinc is going to be in short supply, should substitutes be considered and developed?

The alternative identified questions are recorded and used to guide the subsequent steps of the flowSheet development process. Looking back at the original question boundary conditions are set in order to define it better (Table 2). Several of these boundary conditions are not defined by the question and assumptions need to be made, for example on the flows and stocks to be considered in the analysis.

Table 2 Boundary conditions.

Commodity	Zinc
Value chain	All stages (whole life cycle – exploration to end-of-life)
Impacts	Economic, but related social and environmental impacts to be taken into account, if possible
Spatial (activity)	Global level
Spatial (impact)	Global level, but depending on who is asking the question, it could also be national, continental or local level
Temporal (activity)	Future (2050)
Temporal (impact)	Future (2050)
Flows	All (trade, production, consumption, waste, emissions)
Stocks	All (lithosphere and anthroposphere)

<u>Step 2: Identification of the first sub-topic or question to explore further – identify related data and methods</u>

The first sub-topic to be examined is predicting the demand for zinc in 2050. Demand predictions are quite complicated. There is no single indicator to be used to define demand overall and any predictions made, require numerous influential parameters to be taken into consideration. Some key ones are outlined below:

- Supply changes, e.g. production of zinc
- Population changes, e.g. population growth
- Economic changes, e.g. changes in the gross domestic product (GDP)

- Technology changes, e.g. zinc applications and markets
- Country changes, e.g. changes in the developing/emerging countries (such as increases in comparative wealth might mean more demand for specific mineral elements, such as zinc)

Relevant datasets are then identified to provide tangible information for the aforementioned parameters (Table 3), which will allow the production of trend lines of past demand and the gathering of information on driving forces for demand. Demand scenarios are developed to create visions of possible futures based on assessments using the identified datasets as their foundation. Scenario development represents therefore the proposed method to be used for predicting future zinc demand.

Table 3 Key parameters of influence in predicting zinc demand and related datasets.

Parameter of influence – demand	Dataset
Supply change	Zinc production data (time series)
Populations changes	Population statistics (time series and forecasts)
Technology changes	Zinc applications and market shares; also future forecasts
Economic changes	GDP (time series)
Country changes	Country development indicators; strategic documents with visions to the future

<u>Step3: Second sub-topic or question and related data and methods</u>

The third step in the process includes the incorporation of supply predictions for zinc to 2050.

Key considerations to be taken for predicting future supply include:

- Mine production changes
- Market changes, e.g. price volatility
- Capacity change, e.g. upscaling existing operations or developing new ones
- Geological availability, e.g. exploration activity and mineral resources

Datasets that should be used to quantify the identified considerations are presented in Table 4.

Table 4 Key parameters of influence in predicting zinc supply and related datasets.

Parameter of influence –supply	Dataset
Mine production change	Zinc production data (time series and forecasts)
Market changes	Price volatility and price forecasts
Capacity change	Existing capacity and future opportunities based on existing capacity
Geological availability	Zinc deposit models; mineral resources; undiscovered resources; exploration activity

Supply scenarios should then be developed using the information and data outlined above, which will lead to an economic model, which in turn will enable an estimate to be produced.

Step 4: Conclusive steps and data and methods

The next step of the process requires examining the interactions and relationships between supply and demand in 2050. As substantial work has already been undertaken to understand future demand and supply, then the balance between supply and demand in the future is assessed using Material Flow Analysis and developing a dynamic model that can trace changes in the zinc cycle from past to future and across the supply chain. Information on stocks and flows of zinc will be needed to develop such a model. Imbalances between supply and demand could easily be identified using such a model that can inform decision making, for example in increasing capacity, identifying new deposits, supporting certain technology interventions and so on.

Step 5: Capturing expert knowledge

Questions such as the one explored in this example are not easy to answer. Even though some key parameters of interest have been identified that can influence supply and demand, in reality the interconnections between these parameters and other factors, for example, geopolitics, governance between buyers and suppliers (value chain dynamics), environmental impacts or social impacts, are very complicated and influence how the zinc market will develop in the future. The visibility that any expert has in predicting the future is limited and any predictions made are based on specific scenarios. Several assumptions are required during scenario building, which include different degrees of uncertainty.

6. Conclusion

The report explained how data and information are transformed into knowledge using the DIKI model for raw materials. The model underpins the objectives of MICA and it is therefore used to explain how the various elements of MICA and the EU-RMICP contribute towards knowledge generation and sharing.

The model provides clear definitions of the terms data, information, knowledge and intelligence for raw materials that are often used interchangeably, even though they are different. It comprises a framework for classifying 'resources' based on their content, what they deliver, by whom and for what purpose. The incorporation of 'Actors and Needs' in the DIKI model has enhanced its applicability and relevance to raw materials.

7. References

Ackoff, R. L. (1989). "From data to wisdom." Journal of Applied Systems Analysis **16**: 3-6.
Barlas, I., A. Ginart and J. L. Dorrity (2005). Self-evolution in knowledge bases. IEEE Autotestcon, 2005.

Boisot, M. and A. Canals (2004). "Data, information and knowledge: have we got it right?" Journal of Evolutionary Economics **14**(1): 43-67.

Dalkir, K. (2005). Knowledge management in theory and practice Oxford, UK, Elsevier

Davenport, T. and L. Prusak (1998). Working knowledge. How Organisations Manage What They Know. Boston, Harvard Business School Press.

Eliot, T. S. (1934). The Rock, Faber & Faber.

Foucault, M. and C. Gordon (1980). Power/knowledge : selected interviews and other writings, 1972-1977. New York, Pantheon Books.

Frické, M. (2009). "The knowledge pyramid: a critique of the DIKW hierarchy." Journal of Information Science **35**(2): 131-142.

Henry, N. L. (1974). "Knowledge Management: A New Concern for Public Administration." Public Administration Review **34**(3): 189-196.

Hey, J. (2004). "The Data, Information, Knowledge, Wisdom Chain: The Metaphorical link ", from www.dataschemata.com/uploads/7/4/8/7/7487334/dikwchain.pdf.

Metaxiotis, K., K. Ergazakis and J. Psarras (2005). "Exploring the world of knowledge management: agreements and disagreements in the academic/practitioner community." Journal of Knowledge Management **9**(2): 6-18.

Rowley, J. (2007). "The wisdom hierarchy: representations of the DIKW hierarchy." Journal of Information Science **33**(2): 163-180.

Sharma, N. (2008). "The Origin of Data Information Knowledge Wisdom (DIKW) Hierarch." from www.researchgate.net/publication/292335202_The_Origin_of_Data_Information_Knowledge_Wisdom_DIKW_Hierarchy.

Zappa, F. (1979). Packard Goose. Joe's Garage: Act II & III Tower Recrods

Zeleny, M. (1987). "Management Support Systems: Towards Integrated Knowledge Management." Human Systems Management **7**(1): 59-70.