Chapter 4

Towards Semantic Knowledge Maps Applications: Modelling the Ontological Nature of Data and Information Governance in a R&D Organization

Ivo Pierozzi Junior, Patrícia Rocha Bello Bertin,
Claudia De Laia Machado and
Alessandra Rodrigues da Silva

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/67978

Abstract

In organizational management, it is highly recommended that data and information be adequately prepared to match the knowledge needed to be used in decision-making processes. However, faced with the paradigm of complexity that currently dictates the dynamics of modern organizations, there is still a search for operational solutions that allow agility and flexibility to corporate information flows to meet that desired condition. In this context, the concept of data and information governance presents itself as a fundamental premise because it systematizes, reorganizes and reorients each element of the organizational system (people, processes, structures, etc.) without losing the notion of its contexts and causalities. For this, in the conceptual modelling of governance, the concept of systemism arises to support the balance between holistic and reductionist approaches, inherent in management processes, but often considered antagonistic or contradictory. The present chapter presents and discusses a data and information governance model for research and development (R&D) organizations. The model is based upon the concepts of data, information and knowledge life cycles and knowledge mapping, recovering and valuing the ontological nature of the elements of the system under analysis and constructing a pragmatic proposal for corporate application and operation.

Keywords: knowledge mapping, semantic maps, agricultural research, data management, information management, knowledge management

1. Introduction

In the face of complexity, an in-principle reductionist may be at the same time a pragmatic holist. [1]

One of the greatest challenges of management is to align and integrate the innumerable elements that constitute the intricate system of organizational processes and, from a broader perspective, being able to assess institutional development based on performance and effectiveness of specific actions on those same elements of the system. However, the boundaries between organizational processes are often ill defined, most of them being highly interdependent and transverse, which makes convergence of efforts and coherent results more difficult to achieve.

Management processes are highly dependent on a parallel and continuous process of decision-making. Every decision, in turn, depends on the availability of data and information, which is the basis of knowledge construction. In absolute terms, there is not 'the best' decision to be taken, as this would require access to the complete and utter repertoire of available data and information about a particular subject—a superhuman intellectual effort would be necessary for processing and synthesizing such amount of information and selecting the best of all possible decisions and their expected causalities. In front of these impossibilities, it is assumed that a general knowledge management program (implicitly included, data and information management processes) is needed to support the decision-making process, directing knowledge production or recombination for use or reuse in successive decision-making cycles.

Thus, the greatest dilemma of a decision-maker at any organizational level is perhaps to reconcile and resolve the balance between analysis and synthesis or, in other words, between reductionist and holistic viewpoints to better understand the system at hand. Even when this dilemma is sorted, there remains the difficulty of operationalizing, both logically and technologically, the relevant management processes.

This chapter presents an ontological approach to organizational systems, here explored in terms of a data and information governance (DIGov) framework for a research and development (R&D) organization. This governance framework encompasses a number of management processes relating to the triad 'data-information-knowledge'.

Section 2 of this chapter offers background information about the organization that served as the context of this study (Brazilian Agricultural Research Corporation—Embrapa), while Sections 3 and 4 present the theoretical foundations of the data and information governance (DIGov) framework and the methods that were used in this research, respectively. Embrapa's DIGov model is then explained in Section 5, with the two last sections of the chapter focusing on how the model can be used in practice, with support of knowledge-map-inspired conceptual structures (Section 6) and final considerations (Section 7).

http://dx.doi.org/10.5772/67978

2. Context

The design and implementation of efficient corporate management processes must be supported by a logical, conceptual or even ideological framework, which mediates the causal relations among premises, values, norms and rules, logical and physical structures, costs, personal skills as well as people behaviour and attitudes. All these elements relate to the notion of 'corporate governance' [2], which adheres to the organization's strategy, therefore differing fundamentally from the 'management' system, this one being ultimately focused on operational and tactical issues, such as the monitoring of activities and results achievement.

The Brazilian Agricultural Research Corporation (Embrapa), a large governmental organization, has been developing its own conception of corporate governance. As it is probably the case with many other R&D organizations, Embrapa now finds itself as an immense, diverse and complex organization with its multi-, inter- and transdisciplinary scientific scope; its heterogeneous thematic Research Units spread all over the Brazilian territory; its staff that includes not only scientific researchers but also a wide range of support personnel and experts from many knowledge backgrounds; and its broad stakeholder network, comprised of many sectors such as the State, governments, farmers, academics, students and society as a whole, which are ever more demanding and attentive to the implications of agricultural research, making their voice heard on the most controversial issues (an example is the heated debate and influence of the public opinion on research with Genetically Modified Organisms).

Embrapa is proud to present itself to society as a knowledge-producer organization—the word 'knowledge', in this context, meaning any solution or practical response to a specific demand or problem. It turns out that the knowledge that Embrapa produces is a direct consequence of the alignment of agricultural empirical data and resulting information, which are gathered, processed, shared, reused and disseminated in a dynamic, continuous, cyclical and fed-back process, aimed at consolidating a certain kind of knowledge that, in turn, inserts more cognitive value into decision-making. This is not to be seen as a linear process, since it often involves uncertain, unforeseen and even random developments. These general properties of complex systems, inherent to organizations such as Embrapa, have the potential to hinder or delay management decisions, since required data and information are not always timely and easily accessible.

This is not to say that Embrapa's data and information management processes are inappropriate or inefficient, nor that correlate governance is missing. It can be said, though, that the input and output flow of information and the chain of activities that make up both Embrapa's management and governance processes still leave room for improvement. Particularly, the development of coherent, commonly shared practices of data and information production, sharing, reuse and dissemination is highly desirable as a means to compensate for the traditional, hierarchical organization chart and its many decision-making structures, where information flows can be greatly impaired by power microstructures and bureaucracy.

Figure 1 illustrates how corporate management processes can benefit from properly managed data, information and knowledge, through improved decision-making.

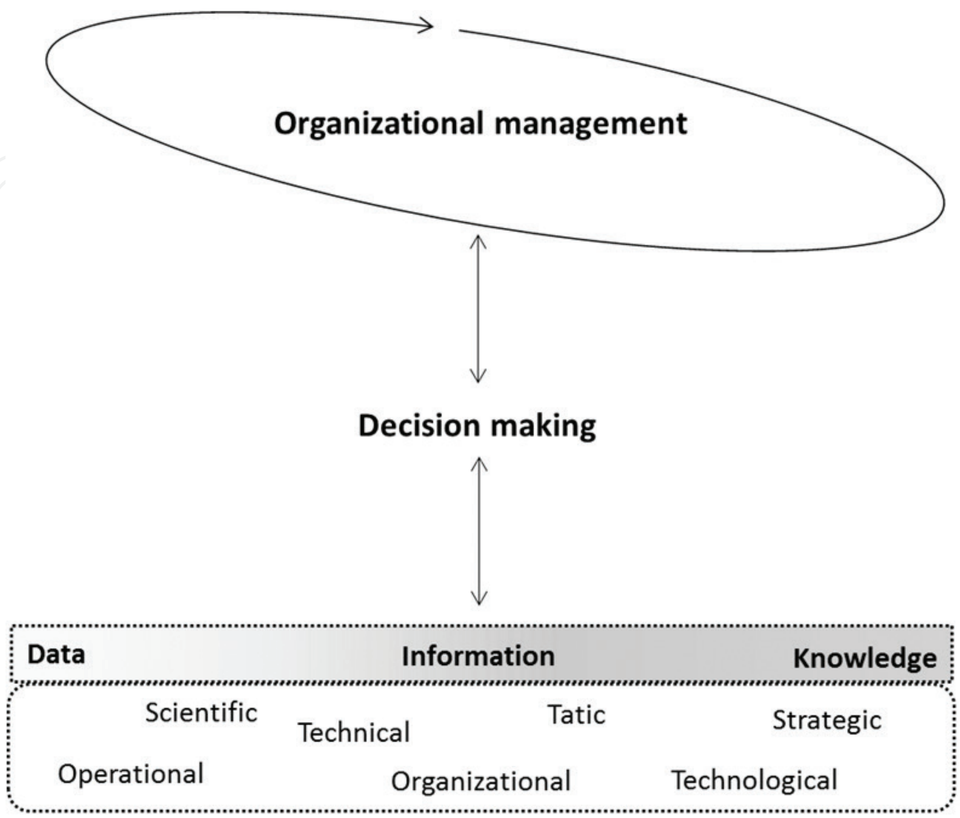

Figure 1. Attaching data, information and knowledge to organizational management through the decision-making process.

3. General approach to data gathering and analysis

Individually considered, the building blocks of Embrapa's Data and Information Governance (DIGov) model are not entirely new but relate to a range of previous conceptualizations and notions, which are detailed in Section 4. The DIGov model, however, is not only theoretically informed but also empirically grounded and based upon a deep understanding of Embrapa's information environment.

A data gathering plan was built with support of the well-known 5W2H management tool (What, Why, Where, When, Who, How and How much) and through questionnaires with both closed and open-ended questions, a large volume of data was gathered, categorized and reciprocally linked, pointing out actors; skills; logical, physical and computational structures; processes, workflows, rules and regulations; stakeholders and even potential or incipient, informal governance sub-systems.

It can be said, therefore, that this study applied both a deductive and an inductive approach [3], building upon prior knowledge and, at the same time, allowing new themes to emerge from data.

Knowledge mapping was identified as a useful tool for data analysis and representation, allowing for a complete system characterization that appreciated both the conceptual aspects of the entire complex system (level of ideas/macro-properties) and its instances (level of objects/entities). By doing so, particularities of each component could be explored without losing the sense of the whole.

Empirical data were thematically analysed [4] with support of the qualitative data analysis software Atlas.ti, and to produce a graphic representation of the interrelationships between data elements, these were translated into triple store format, A-R-B, where A and B represent the elements or concepts included in the system (entities, objects, facts, etc.), defined by terminological or textual labels; R means the relationship, defined by semantic labels and '-' can assume uni- or bi-directional paths. The triple stores were gathered and organized in a spreadsheet with three columns and innumerable rows where each row represented one triple.

For allowing the possibility of editing and visualizing a complex conceptual structure (holistic view) and then breaking it into snippets for a more detailed (reductionist) view, the software yEd (https://www.yworks.com/products/yed) was used. Besides generating high-quality diagrams and visualizations, yEd supports mathematical analysis of social relationships to provide insight into the structure of a social network. Metrics such as density and centralities (betweenness, closeness, degree, etc.), for instance, give a measure of the relevance of a particular element in the whole network.

4. Theoretical perspective

This section presents the theoretical foundations of the Data and Information Governance (DIGov) framework, exploring underpinning notions and conceptualizations to create a systemic approach to management problems at contemporary organizations, while acknowledging their ontological and complex nature. The notion of information as a complex, social phenomenon, the cognitive itinerary formed by Data-Information-Knowledge (DIK) and the conceptual alignment between the DIK life cycles form the basis of the theoretical framework, as explained in the following sections.

4.1. The corporate information environment as a complex system

The informational environment of an organization can be assumed as a complex system, that is, systems composed of innumerable autonomous but interactive elements, where the result (output) of the system is not simply the sum of the properties and particularities of its parts [5–7].

Under the systemic perspective [1, 8–10] and in the context of an organization, the corporate information environment is a social phenomenon [10, 11] and thus can be seen as a complex sub-system of another complex higher system containing, itself, other complex sub-systems within itself. As a social phenomenon, information (and, consequently, the data that originates it) assumes institutional properties and causality in the cognitive, communicative, documentary and normative or regulatory dimensions. So, one of the great challenges of complexity, when social sub-systems are involved, is [12]; since people do not act continuously and regularly in a rational way, nor are they compliant with norms and laws throughout the time, it is not uncommon for them to react in a way not intended or planned by managers and their management strategies or, at least, even if people do not disobey or do not manifest themselves contrary to superior guidelines, they may react inconsistently to corporate guidelines. These inconsistent reactions can meet random, uncertain or unpredictable situations when people are often pressured by the emergency or urgency of requested arrangements and by the huge volume of decision-making moments they experience in their daily work routines. Due to the complex interactions of these systems and the non-linear way in which their elements give rise to general behaviour patterns, complex systems can be very difficult to predict, control and manage [1, 8, 9].

It is in this context that data and information governance presents itself as a preventive, conciliatory solution, which is more concerned with guiding premises to foster good practices and less with guaranteeing results in an idealized world of strategic planning processes and deterministic projects.

4.2. An integrated look on data, information and knowledge

At R&D institutions, empirical research is an established practice. It means that one of their main concerns is to obtain and translate data into scientific knowledge, which can then be applied to solve real-world problems. For this itinerary to be complete, an improved understanding of the conceptual line between data, information and knowledge needs to be achieved.

The word 'information' is most commonly used to mean physical representations of knowledge: objects, data and documents that possess instructive character, a use that has been previously described as 'information-as-thing' [13]. Alternatively, the term is used in a wider sense, as in reference to the act of informing or becoming informed ('information-as-process'), or to what we know ('information-as-knowledge'), that is, whatever is perceived in 'information-as-process'. The interrelationship becomes then evident: it is difficult to define 'knowledge' without referring to 'information', as it is to describe 'information' without referring to 'data'. The following quotes illustrate this:

> *Knowledge is information evaluated and organized in the human mind so that it can be used purposefully. [14]*

> *[Knowledge is] information combined with experience, context, interpretation and reflection. [15]*

For taking many forms, both physical and digital, the term data can be difficult to define [16]. Among the most widely cited definitions is the following, from the National Academies of Science:

> *Data are facts, numbers, letters, and symbols that describe an object, idea, condition, situation, or other factors. [17]*

In practical terms, data are equivalent to a physically 'recorded symbol', which can be exemplified by printed characters; binary characters in magnetic, punched or optical form; spoken words; or images. Whatever the physical form may be, it becomes a recorded symbol when it is interpreted as representing something [14]. In the words of Feather and Sturges,

> *raw data is the building block, and knowledge is the construct; information is the cement. The effective management of information allows it to be stored by means that permit it to be systematically and efficiently retrieved in a format that will facilitate the tasks of the end-user. [14]*

Prior attempts to bring the concepts of data, information and knowledge closer together, in order to identify their boundaries and build a logical trajectory between them can be found in the literature of Cognitive, Information, Management and Computing Sciences. The notion of a Data-Information-Knowledge-Wisdom hierarchy [18] is opportunely recovered here, even though the philosophical discussions and implications arising from it are not addressed [19]. The hierarchical-pyramidal representation is one of the most influential perspectives, adopted in support of several lines of argumentation [19-22]. Other more cognitively elaborated appraisals represent the relation in a linear-progressive form [18]. Despite being conceptually instigating, such representations are criticized from a pragmatic point of view, as to their usefulness in supporting data, information and knowledge management in practice [19]. Conceptually speaking, the main criticism is the reasoning that, if wisdom is to be taken as an 'unquestionable and irrefutable truth', it might not be reached if the data supporting it turned out to be incorrect or untrue!

An alternative representation of the relationship between data, information and knowledge is offered in **Figure 2**. In this new conceptual set up, there is a deliberate preference, based on the logic of added cognitive value, for aligning data, information and knowledge in a circuit with a continuous feedback loop, rather than the conventional hierarchical-pyramidal or liner-progressive representations [23].

Despite their conceptual interrelatedness, however, 'data management', 'information management' and 'knowledge management' studies have specialized as different bodies of knowledge. Analogously, traditional management approaches in organizations also tend to treat 'data', 'information' and 'knowledge' separately, which might cause considerable confusion, given their high levels of complementarity and interdependence [24, 25].

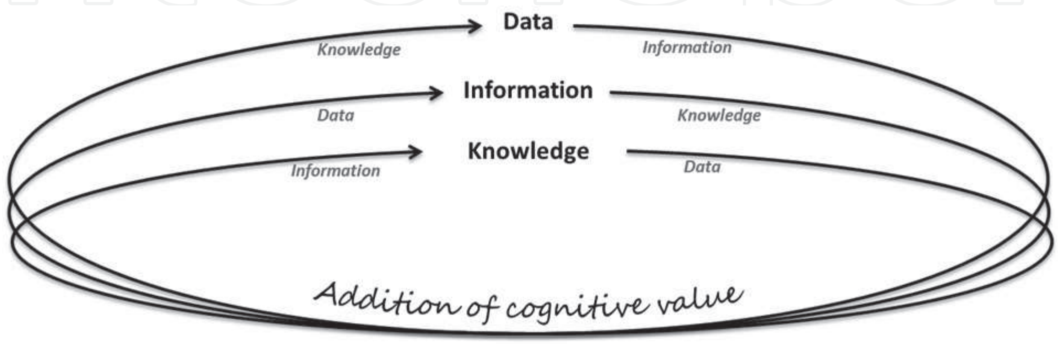

Figure 2. Data, information and knowledge in a cyclic, continuous feedback circuit. Source: Ref. [23].

To overcome this difficulty while building an integrative and systemic approach to data and information governance, these elements were arranged in a conceptual continuum—their interrelation has been acknowledged, so that their use can be maximized in the organization.

4.3. Aligning data, information and knowledge life cycles

In figurative sense, data, information and knowledge are taken here as a quantum triad: depending on the point of observation, what is seen is intellectual energy (thought, idea, cognition), or, alternatively, tangible matter (documents and media, printed or digital), which can then be collected, assembled, organized, analysed, shared, accessed and continuously reused independently of the human brain.

To better represent this perspective, the concepts of data, information and knowledge 'life cycle' emerged as a promising solution. Although offered in many different versions in the literature, these life cycles can be used, in a practical way, to support a cyclic, continuous and feedback view of the data, information and knowledge itinerary. **Figures 3–5** present examples of data, information and knowledge life cycles (DIK life cycles), as they are central to the understanding of the governance model developed in this chapter.

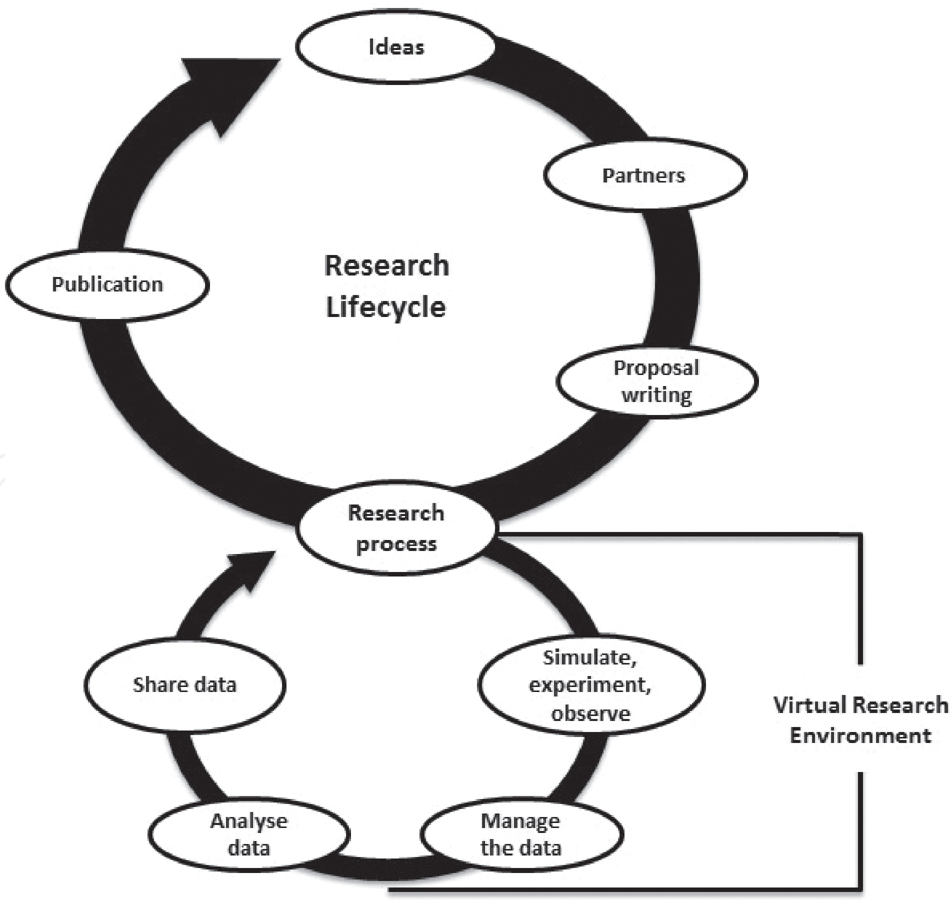

Figure 3. The data life cycle and its connection with the research process, as proposed by Tenopir et al. [26].

http://dx.doi.org/10.5772/67978

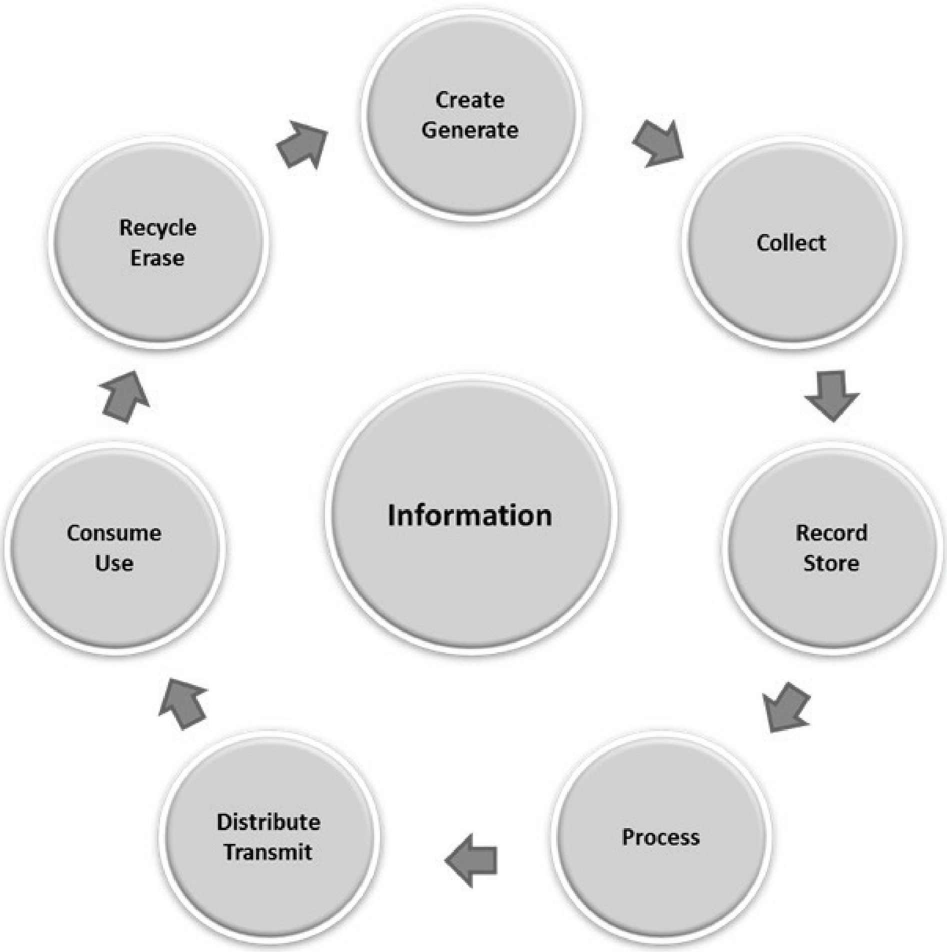

Figure 4. The information life cycle, according to Floridi [27].

On closer examination, the DIK life cycles seem comparable and have many similarities—which is not surprising, given the conceptual interrelatedness of its elements. **Figure 6** was obtained by expanding the conceptions of DIK life cycles that are available in the literature, in order to align and specify the fundamental processes (genesis, creation, production; organization; sharing; access and dissemination; appropriation, validation, evaluation). This is, therefore, a simplified view of the fundamental logic of alignment between these life cycles, here called the 'DIK Mandala', which served as a central construct to the concept of governance developed in this work.

This integrative look at the data, information and knowledge life cycles has practical significance, since it allows an appreciation of the processes involved in data, information and knowledge management, in terms of specific methodologies and technologies and parallel and progressive arrangements, as presented in **Figure 7**.

4.4. Conceptualizing data and information governance

The notion of data and information governance retrieves the understanding of governance already adopted in other initiatives of Embrapa, considering as part of a general corporate

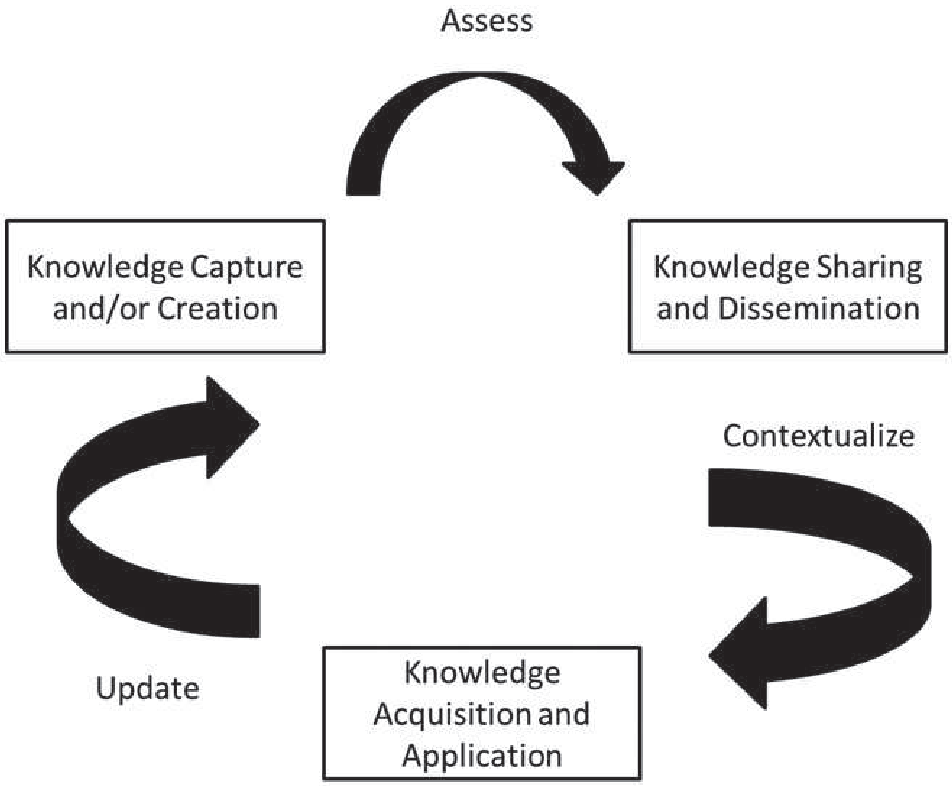

Figure 5. Knowledge management integrated cycle. Source: Dalkir [28].

governance proposal while aligns parallel to other subcategories of organizational governance, such as Information Technology governance, Human Resources governance and so forth. However, against the possible risks of creating an infinite number of instances of governance (reductionist paradigm of the organization), a systemic approach was adopted [8] to sustain the dynamics of information flows within and across organizational subunits, hierarchical levels and governance perspectives. For Embrapa, data and information governance for knowledge means: 'the determination, systematization and institutionalization of the principles, guidelines, structures, processes, culture, roles and responsibilities that drive, enable and transform data and information management in support of decision-making and the corporate governance system'.

The notion of 'governance' supported here, therefore, is not to be confused with that referring to information technology. On the contrary, it recognizes that problems related to information are often multifaceted—involving behavioural, cultural, regulatory, procedural and structural aspects that are not solved exclusively through the adoption of technological solutions.

As adopted here, data and information governance presents itself as an innovative approach that is situated in a higher and more strategic level than the operational, mechanistic and bureaucratic nature of management processes and technological tools. With constituent, morphological elements (principles, guidelines, structures, processes, cultures, roles and responsibilities) that relate to the corporate information environment, the

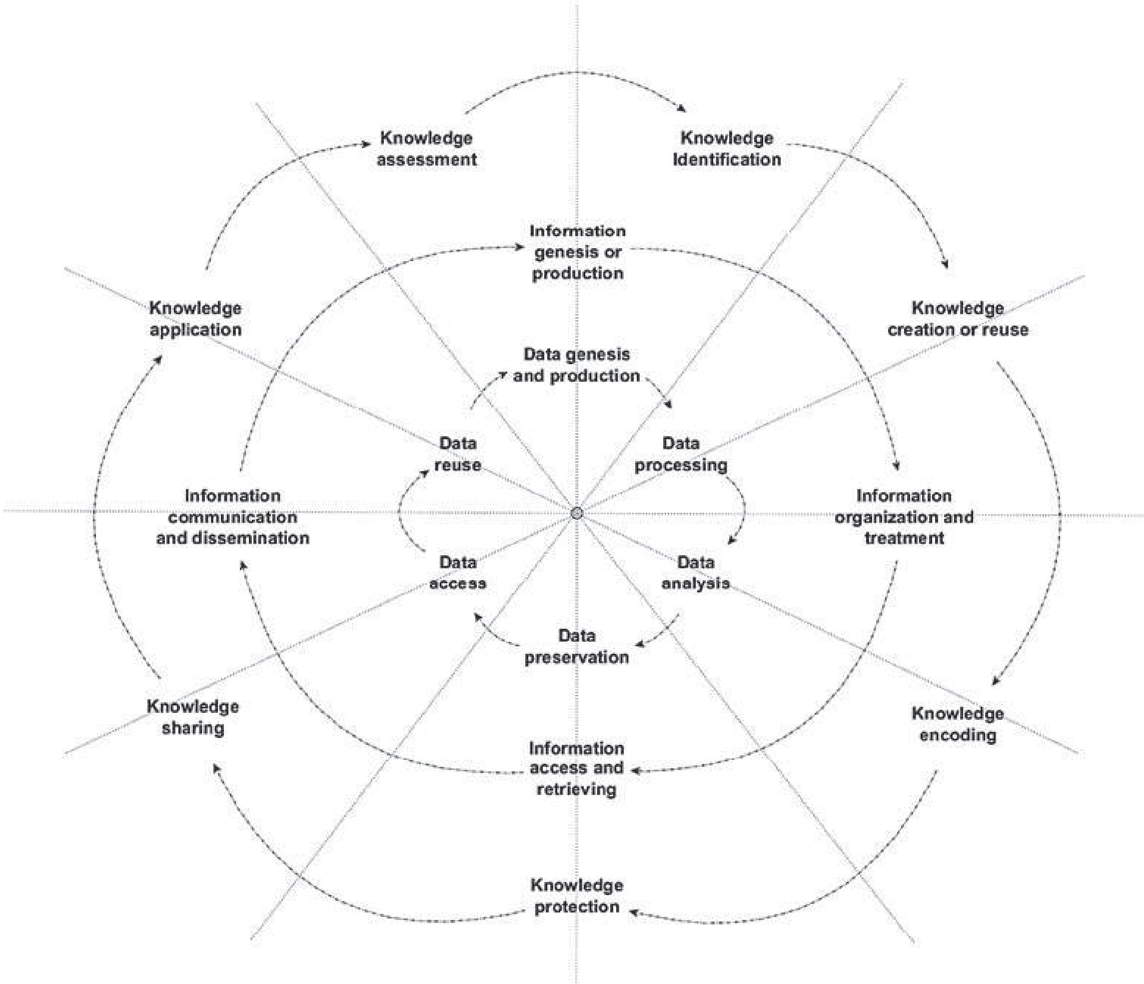

Figure 6. The data-information-knowledge Mandala.

data and information governance framework can be seen as a fluid, dynamic and aggregating material that permeates, integrates and interacts those elements with each other, assuming (also metaphorically) the 'physiology' of the system formed by the corporate management mechanisms, but with enough and sufficient room to deal with non-linear causalities.

This being said, the difficulty presented for the conception and development of a generic model of data and information governance was precisely to find adequate mechanisms and tools to turn these conceptual ideas into operational pragmatics. A natural way to pragmatize the DIK relation was sought, valuing its ontological nature and acknowledging, in this itinerary, the most logical solution to organize ideas around the corporate informational environment—which, in this particular case, is that of a R&D organization.

The following section presents Embrapa's DIGov model, which fills a gap in the literature of Information Science, for pursuing an ontological, systemic and conceptually integrative perspective on the reciprocal relationships of data, information and knowledge management, in the context of R&D organizations.

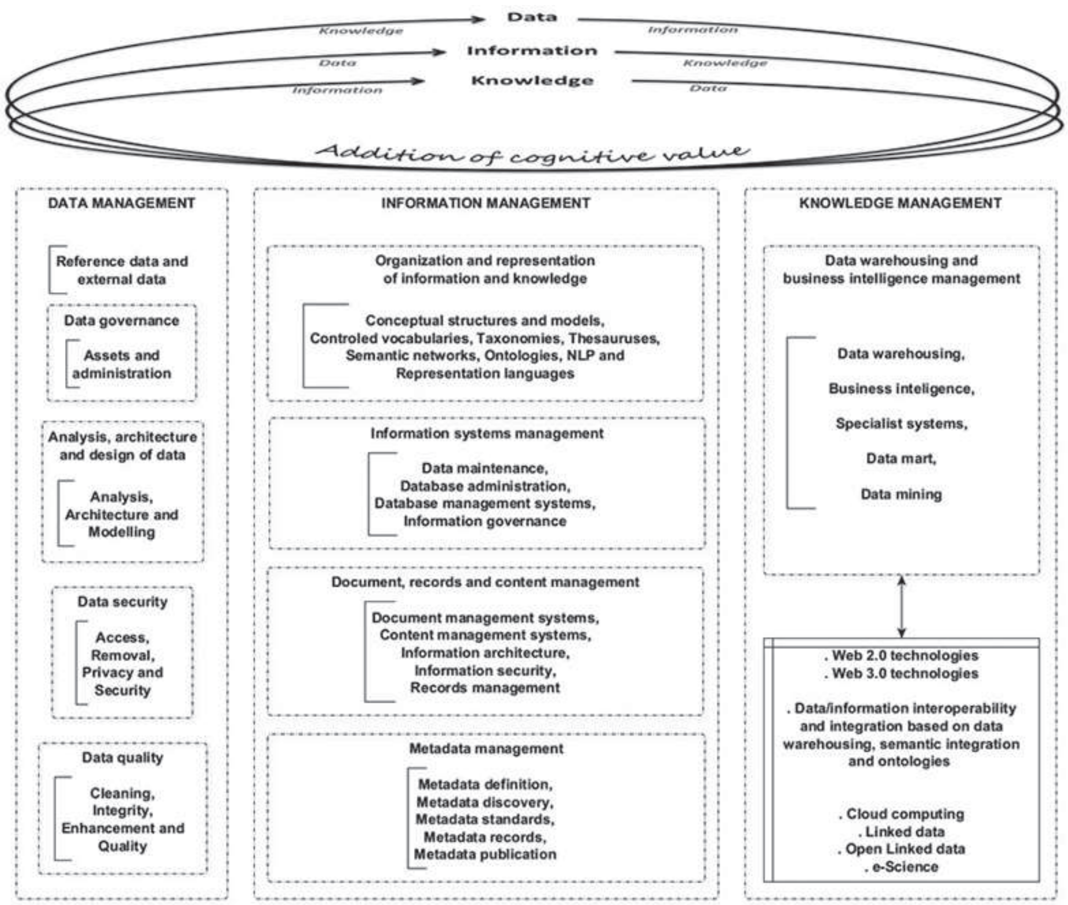

Figure 7. Alignment between data, information and knowledge management processes and activities, as a function of cognitive mechanisms and in relation to technological solutions. Source: Ref. [23].

5. Embrapa's data and information governance (DIGov) framework

The DIGov model is shown in **Figure 8**. Its main constituents and operation or physiology are described as following.

The DIGov model articulates with a series of theoretical, philosophical and conceptual notions, which are available in the literature: 'information ecology' [29], 'information policy' [30], 'informational audit' [31] and 'information culture' [32], among others. Some of these precedent approaches have arisen in parallel to the world views derived from or influenced by theories based on relationship or interdependence of system elements and global phenomena, which emerged by the end of World War II and were well consolidated in the 1990s. A creative and timely review and contextualization of the dynamics of human evolution of global knowledge over the past 70 years were presented by Brian Castellani and his work of successive editions of the 'Map of Complexity Sciences' (http://www.art-sciencefactory.com/complexity-map_feb09.html), where the main theories and scientific approaches are interrelated, as well as their authors. These theories, which support the current itineraries of knowledge that model our understanding of life and the Universe, derive from two main

http://dx.doi.org/10.5772/67978

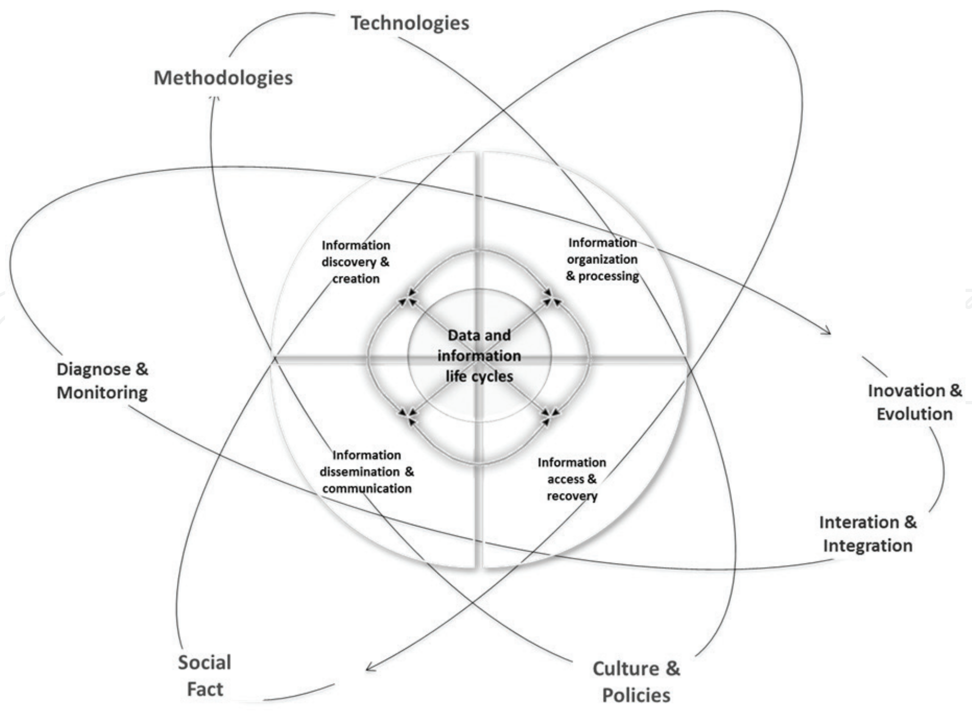

Figure 8. Embrapa's DIGov model.

intellectual strands—Systems Science and Cybernetics, both emerging in the second half of the twentieth century. The map registers the convergence of knowledge that today enables us a more systemic and organic view of the world, pointing out General Theory of Systems, Ecological Systems Theory and extending to discussions about complex adaptive systems until including, more recently, the social aspects of the knowledge development process, as in Social Systems Theory, Sociocybernetics and culminating with the e-Science and Data Science theorizations.

The influence and recovery of those chains of thought into business schools around the world have been previously analysed and contextualized [11], allowing new perspectives to emerge that make such conceptual approaches closer to the pragmatism of organizational management.

It is at this point that 'sistemism' [8] arises as one of the main theoretical-conceptual foundation of the DIGov model. It presents itself as a conciliatory solution between analysis and synthesis, the whole and the parts, holistic representations and reductionist ones, which are inherent to modelling processes, including those of organizational nature. As explained in Section 3, however, the DIGov model not only draws upon a vast theoretical background and the very definition concept adopted in this work, but it is also based upon an empirical analysis of data, information and knowledge management at a R&D organization.

The central axis of the DIGov model is formed by the data, information and knowledge life cycles—in a convergent and interrelated perspective, as explored in detail in the DIK Mandala (**Figure 6**). 'Discovery and Creation', 'Organization and Processing', 'Access and Recovery' and 'Dissemination and Communication' are, therefore, convergent dimensions of the data, information and knowledge life cycles, reflecting the managerial and operational activities that are involved in each phase of these cycles in the various organizational instances and levels. From the data collected at Embrapa, it was observed, however, that these steps, in practice, are not necessarily performed in a linear and sequential way: certain sets of data and information can be produced and immediately made available for 'access and recovery'; or, still, there are datasets and information that need to be previously organized and treated, for only then be made accessible. The several stages of the data, information and knowledge life cycles can be—and, in some cases, must be—navigated iteratively and non-sequentially, depending on the nature of the informational asset and corporate interest.

To ensure a continuous improvement of data and information management, however, it is necessary to observe trends and opportunities for improvement, so that the operation of the data, information and knowledge life cycles, which correspond to the DIGov model's 'micro-level', must be subject to the following processes, occurring in the 'meso-level': (a) continued monitoring and diagnostics of structures, skills and processes; (b) enrichment and strengthening of the interaction and integration among and between people and information and between those and available technologies; and (c) development and validation of innovations and developments in methodologies and technologies to support data, information and knowledge management.

Finally, complementing the systemic design of the DIGov model, the following components surround and guide the model in a 'macro-level': (i) culture and (ii) social fact, with regard to social aspects and people absorption and application of the model's morphology and physiology; (iii) policy, as regards the model's legal, regulatory and strategic framework; (iv) technologies; and (v) methodologies.

6. A knowledge mapping application for data and information governance

Knowledge mapping has been presented and discussed as a valuable tool for organizing, representing and retrieving knowledge, being particularly suitable for large amounts of information [33–38]. The main idea behind knowledge mapping is that facts, entities and objects of any kind can be identified, highlighted and interrelated to each other, in order to solve the deficiencies of documentary languages, which are conventionally based on categorization or classification systems. In DIGov model's scope, such deficiencies relate to: (a) the need of a conceptual framework for collective cognition and communication enhancement which respects the multidimensional and multifaceted nature of data, information and knowledge and (b) the need of mapping the numerous informational flows, emphasizing both its specific details as their general context in corporate information and knowledge management processes. Thus, knowledge mapping allows the reconciliation of reductionist and holistic paradigms, often conflicting in the choices of tools for organization and representation.

http://dx.doi.org/10.5772/67978

To illustrate the use of this concept and practical tool, a real-world example will be taken, among many possible ones that relate to data, information and knowledge in support of decision-making. This will be provided by the Open Access (OA) thematic, which is a recurrent, global demand to Embrapa. Acknowledged as a world reference in tropical agricultural knowledge, Embrapa is invited to present a clear, institutional position and to take practical steps towards OA to the knowledge it produces.

This was one of the main issues to emerge from the questionnaires applied in this research, which was therefore elected as one of the priority topics to be addressed by the company, so as to improve its information and knowledge management processes. In addition to other priority issues (e.g. controlled vocabularies, terminologies, semantic features and its applications; editorial process and policies; research data management; strategic information management, among others), a roadmap was designed to explore and achieve a corporative solution to the OA issue.

From the systemic point of view, which informed the DIGov model construction, a choice of tools was made to better represent this system as a complex one. Such tools are mainly editing software for conceptual structures that would allow the ontological design (concept mapping) of the represented system. In such type of representation, all of the components are assumed to be nodes and their mutual interrelationships are the vertices that connect the nodes to each other, creating a network. Embedded in these networks, it could be assumed that the corporative informational streams flow by the organization morphology (people, structures, processes) and feed its physiology, that is, the way in which an organization develops or operates its activities (*modus operandi*).

By projecting the DIGov model onto the empirical data and information that was gathered, the following framework was obtained, which forms the basis for the development and application of a knowledge map (**Figure 9**).

This higher level of abstraction and relational view of knowledge organization can still be useful in exploring the interrelationships of a system's components in the more operational level.

Figure 10(A) presents a multidimensional representation of Embrapa's data and information management panorama, mapping its main elements, as empirically observed. The large volume of data collected through the questionnaires was interconnected and their mutual interrelationships (causalities) were registered. **Figure 10(B)** highlights the contextualization of the OA issue in relation to the overall conceptual structure.

This exercise needs not to be complete and final. Being a complex system, its breadth, diversity and dynamism imposes the need of a model (and, consequently, of a representation tool) that meets the system's plasticity and allows its continued modification, due to uncertainties, randomness and unpredictabilities that characterizes the organizational time and space. However, features of this dynamics can be evidenced and captured to assist their operational management.

Figure 11 focuses on the concept of OA, isolating it from the whole it belongs to and highlighting the most relevant and immediate relationships that it establishes with other elements of the system.

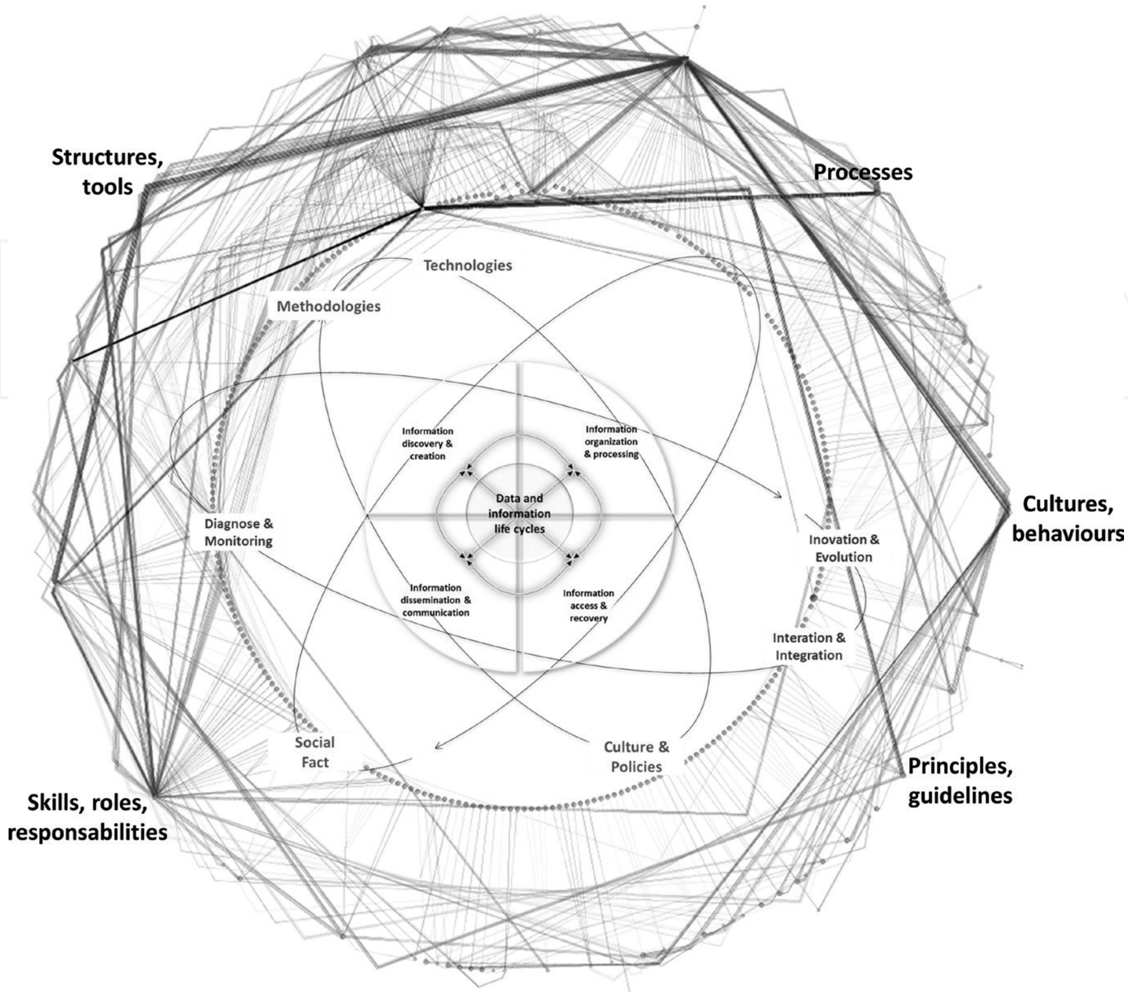

Figure 9. Contextualization of the DIGov model in a network of conceptual, procedural, structural and operational components of Embrapa's information environment.

The final stage of this exercise is to do with making concrete sense of this mapping, drawing it together with the relevant management processes. This is an intellectual, human oriented exercise, but which can be assisted by relevant tools. Doing so, the map presented in **Figure 11** can be graphically edited and semantically refined to incorporate a qualitative view of the positioning of the OA issue within Embrapa's informational environment, by identifying the elements that compose the DIGov model and designing causal relationships between them to recommend operational actions, whether to change culture and behaviour or to improve methodologies, structures or supporting technologies.

A semantic refinement of the relationships drawn in the model can and should be systematized and standardized corporately, strengthening a collectively agreed language process to ensure that the corporate learning processes become more aligned from cognitive, procedural, normative and communicational points of views. This would then result into the design that is shown in **Figure 12**—a pragmatic governance proposal to support decision-making regarding this particular issue in the corporate context.

http://dx.doi.org/10.5772/67978

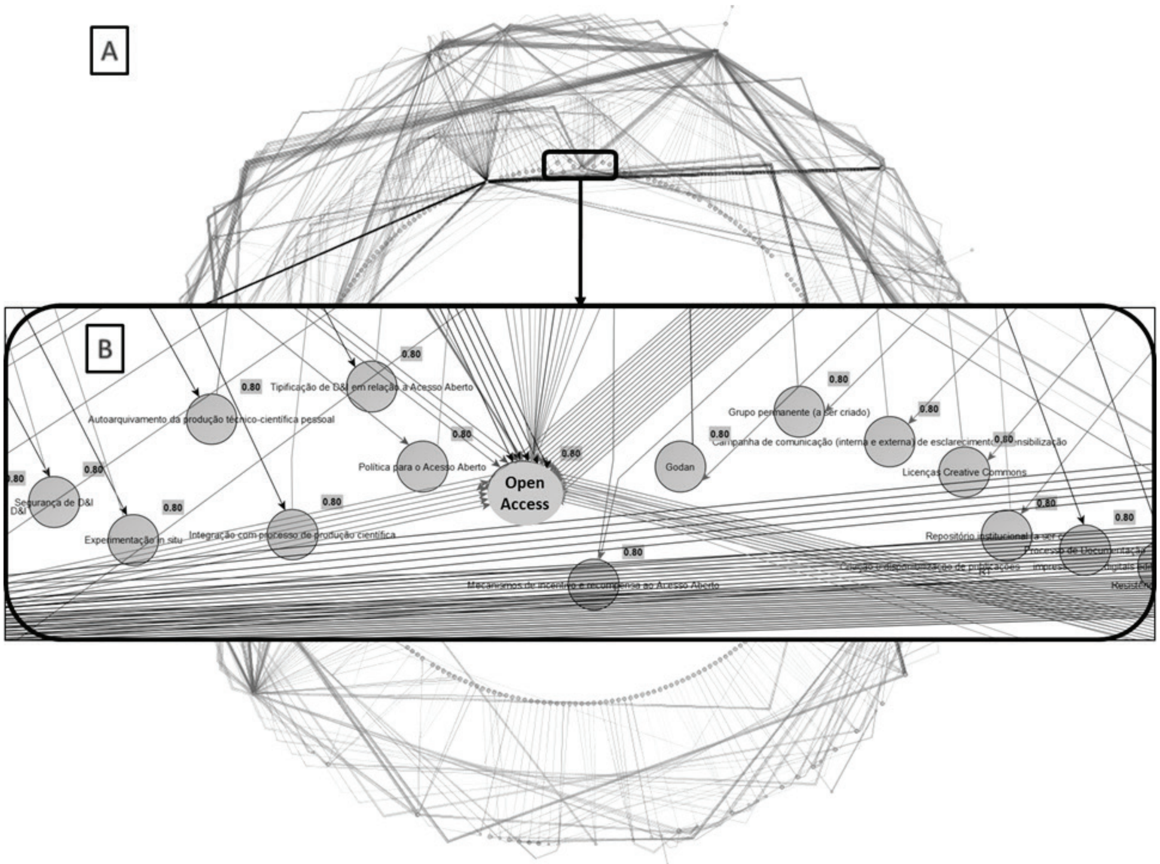

Figure 10. Conceptual representation of Embrapa's data and information management panorama (A), highlighting the positioning of the Open Access issue in relation to the other elements of the system (B).

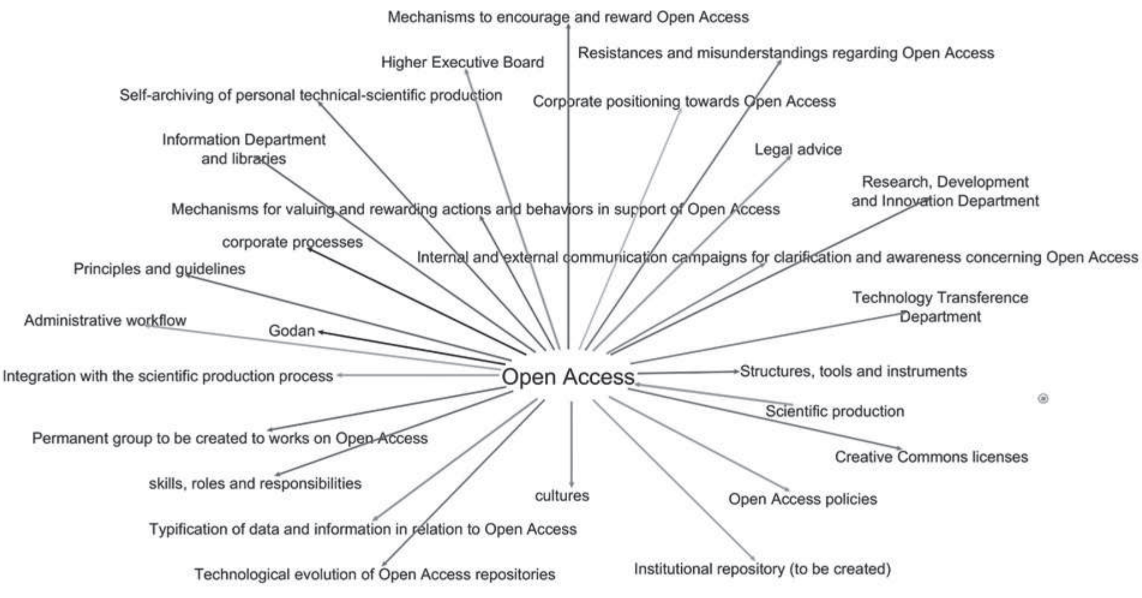

Figure 11. Apartness of the Open Access concept, identifying the most important relationships it establishes with other elements of Embrapa's informational environment.

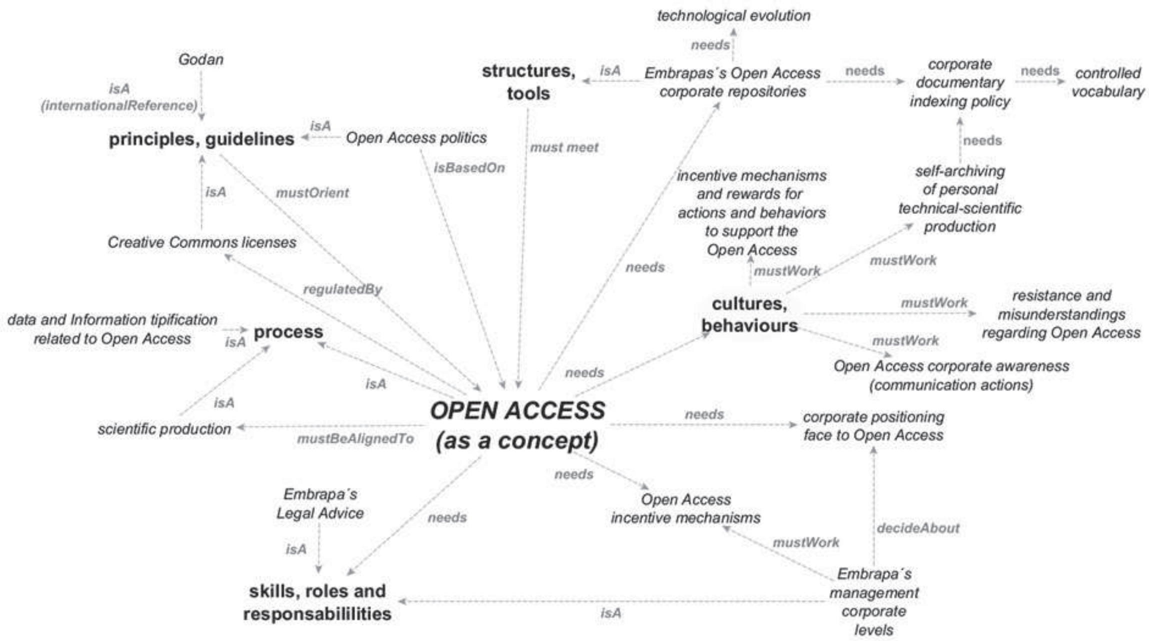

Figure 12. Snippet view (zoom in) of the ontological structure of Embrapa's DIGov model, focusing the Open Access issue within the corporate information environment.

7. Some considerations and future avenues of development

As already mentioned, corporate management processes are inherently aligned with the decision-making process. It was also mentioned that the best decision to be made is always dependent on the assessment of possible alternatives, which are formed by the gathering and analysis of data, information and knowledge relating to the matter of interest.

Given the complexity of R&D organizations' information environment, conventional models of data, information and knowledge organization and representation, which are conventionally of the categorization or classification types (metadata systems and taxonomies, for example), have proven to limit the effectiveness of information search and retrieval, disambiguation and making sense processes. In other words, relational models such as the thesauri, multi-faceted taxonomies, semantic networks and ontologies are more appropriate options for supporting conceptual designs that can better represent the ontological nature and the multidimensionality of causal relations in a complex system [11, 39].

Recent studies have reiterated [33] knowledge mapping as an important tool for knowledge management, since they increase the recognition, systematization, communication and sharing of common corporate practices. Despite being a promising area in the fields of conceptual modelling and tools' development, the operational adoption of knowledge maps still presents shortcomings. The present chapter contributes to filling this knowledge gap, while proposing a method for modelling complex, organizational processes. Furthermore, modelling data and

http://dx.doi.org/10.5772/67978

information governance in R&D organizations such as Embrapa, with support of knowledge mapping, has still the following advantages:

- Knowledge maps are very useful when the immensity and diversity of data and information involved in a given analysis need to be structured from vast content where the encoding of the tacit knowledge has already been processed, for example, textually.
- Knowledge maps do not need to be exhaustive. In fact, they are not committed to representing the real world. Complex systems are dynamic, mutants, reactive to externalities and therefore difficult to be modelled, and, therefore, they require tools which respect their natural plasticity and which, from an operational point of view, can be easily adjusted both in relation to its component elements and in relation to the nature of interrelationships that such elements establish among themselves.
- Knowledge maps are also useful tools for inferring on the social networks implicitly contained in organizational complexity. The identification of social networks come to meet the need to enhance the social fact in the context of organizational management, recognizing its importance and the way people influence and are influenced by corporate management actions. This utility of knowledge maps can be used in support of systemic approaches [8], which have as a first general methodological rule, the commandment to put the social fact in its broader context, that is, its own system [40].

Knowledge mapping (the process) or knowledge maps (the products) are useful for meaningfully representing data, information and knowledge, where and when large amounts of data, information and knowledge are involved. Embrapa's DIGov model, as presented in this work, suggests that this process and its products can be usefully employed as a conceptual and computational model for organizing and managing extensive corporate contents. But far beyond organizing data, information and knowledge repositories, knowledge mapping can be a useful tool and basic framework for navigational purposes through corporate informational flows supporting organizational and collective intelligence applications.

Future avenues of work for the further development, implementation and use of Embrapa's DIGov model would be:

- To evolve the governance proposal in parallel with efforts to incite the desired organizational culture change, towards one that would include the understanding, absorbing and embedding of complex and system thinking as new paradigms, in support of data and information management processes.
- To evolve conventional methods of data and information management, which are still predominantly based upon uni- or bi-dimensional models of knowledge organization systems (KOS) [41] like term lists, taxonomies, categorization or classification schemes, for example, to multidimensional relationships models like thesauri, semantic networks and ontologies, which are more suitable for operationally rearrange knowledge content of large volume and high potential for use and reuse in processes of collective intelligence and institutional development.

Author details

Ivo Pierozzi Junior[1]*, Patrícia Rocha Bello Bertin[2], Claudia De Laia Machado[3] and Alessandra Rodrigues da Silva[2]

*Address all correspondence to: ivo.pierozzi@embrapa.br

1 Embrapa Agricultural Informatics, Brasília, Brazil

2 Embrapa Technological Information, Brasília, Brazil

3 EmbrapaSoils, Brasília, Brazil

References

[1] Simon HA. The architecture of complexity: hierarchic systems. Proc Am Philos Soc. 1962;**106**(6):467-82. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+Architecture+of+Complexity+Hierarchic+systems#1

[2] Daily CM, Dalton DR, Cannella Jr AA. Corporate governance: decades of dialogue and data. Acad Manag Rev. 2003;**28**(3):371-82

[3] Bhattacherjee A. Social science research: principles, methods, and practices (2nd ed.). The global text project. Tampa: University of South Florida; 2012

[4] Boyatzis RE. Transforming qualitative information: thematic analysis and code development. London: Sage Publications; 1998

[5] Waldrop MM. Complexity: the emerging science at the edge of order and chaos. New York: Simon & Schuster; 1993

[6] Mitchell M. Complexity: a guided tour. New York: Oxford University Press; 2009

[7] Casti JL. Complexification: explaining a paradoxical world through the science of surprise. New York: Harper Collins; 1994. 320 p

[8] Bunge M. Systemism: the alternative to individualism and holism. J Socio Econ. 2000;**29**:147-57

[9] Morin E. From the concept of system to the paradigm of complexity. J Soc Evol Syst. 1992;**15**:371-85

[10] Furtado BA, Sakowski PAM, Tóvolli MH, editors. Modeling complex systems for public policies. Brasília: IPEA; 2015. 396 p

[11] Stacey RD. Strategic management and organisational dynamics: the challenge of complexity. 6th ed. Harlow: Pearson Education Limited; 2011

[12] Rand W. Complex systems: concepts, literature, possibilities and limitations. In: Furtado BA, Sakowski PAM, Tóvolli MH, editors. Modeling complex systems for public policies. Brasília: IPEA; 2015. pp. 37-54

[13] Buckland MK. Information as thing. J Am Soc Inf Sci. 1991;**42**(5):351-60

[14] Feather J, Sturges P, editors. International encyclopedia of information and library science (2nd ed.). London: Routledge; 2003

[15] Davenport TH, Prusak L. Working knowledge: how organizations manage what they know. Boston: Harvard Business School Press; 1998

[16] Borgman CL. The conundrum of sharing research data. J Am Soc Inf Sci Technol. 2012;**63**(6):1059-78

[17] National Research Council – US. A question of balance: private rights and the public interest in scientific and technical databases. Washington, DC: The National Academies Press; 1999. 158 p

[18] Bellinger G, Castro D, Mills A. Data, information, knowledge, and wisdom [Internet]. Mental Model Musings. 2004 [cited 2017 Jan 1]. http://www.systems-thinking.org/dikw/dikw.htm

[19] Frické M. The knowledge pyramid: a critique of the DIKW hierarchy. J Inf Sci. 2009;**35**(2):131-42

[20] Bernstein JH. The data-information-knowledge-wisdom hierarchy and its antithesis. In: Jacob EK, Kwasnik B, editors. Proceedings from North American symposium on knowledge organization [Internet]. 1989. pp. 68-75. Available from: http://hdl.handle.net/10150/105414

[21] Rowley J. Where is the wisdom that we have lost in knowledge? J Doc. 2006;**62**(2):251-70

[22] Rowley J. The wisdom hierarchy: representations of the DIKW hierarchy. J Inf Sci. 2007;**33**(2):163-80

[23] Pierozzi Junior I, Souza MIF, Torres TZ, Oliveira LHM de, Queiros LR. Information and knowledge management. In: Massruhá SMFS, Leite MA de A, Luchiari Junior A, Romani LAS, editor. Information and communication technologies and their relations with agriculture. Brasília: Embrapa; 2016. pp. 231-52

[24] Bertin PRB. Towards effective governance of information in a Brazilian agricultural research organisation. Loughborough, UK: Loughborough University; 2014

[25] Lueg C. Information, knowledge, and networked minds. J Knowl Manag. 2001;**5**(2):151-9

[26] Tenopir C, Allard S, Douglass K, Aydinoglu AU, Wu L, Read E, et al. Data sharing by scientists: Practices and perceptions. PLoS One. 2011;**6**(6):1-21

[27] Floridi L. Information: a very short introduction. New York: Oxford University Press; 2010

[28] Dalkir K. The knowledge management cycle. In: Knowledge management in theory and practice. Oxford: Elsevier; 2005. pp. 25-46

[29] Davenport TH. Information ecology: mastering the information and knowledge environment. New York: Oxford University Press; 1997

[30] Davenport T, Eccles RG, Prusak L. Information politics. Sloan Manage Rev. 1992;(Fall): 53-65

[31] Henczel S, editor. The information audit: a practical guide. München: Saur KG; 2001

[32] Choo CW, Bergeron P, Detlor B, Heaton L. Information culture and information use: an exploratory study of three organizations. J Am Soc Inf Sci Technol. 2008;**59**(5):792-804

[33] Balaid A, Rozan MZA, Hikmi SN, Memon J. Knowledge maps: a systematic literature review and directions for future research. Int J Inf Manage. 2016;**36**(3):451-75

[34] Cook DJ, Holder LB, editors. Mining graph data. Electrical engineering. New Jersey: Wiley-Interscience; 2007

[35] Wang Z, Zhang J, Feng J, Chen Z. Knowledge graph embedding by translating on hyperplanes. In: Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence. Québec City, Canada; 2014. pp. 1112-9

[36] Karpicke JD, Blunt JR. Retrieval practice produces more learning than elaborative studying with concept mapping. Science (80-). 2011;**331**(6018):772-5

[37] Pulido JRG, Flores SBF, Ramirez RCM, Diaz RA. Eliciting ontology components from semantic specific-domain maps: towards the next generation web. In: LA-WEB'09 Latin American. Washington, DC: IIEE; 2009. pp. 224-9

[38] Schwendimann BA. Making sense of knowledge integration maps. In: Ifenthaler D, Hanewald R, editors. Digital knowledge maps in education: technology-enhanced support for teachers and learners. New York: Springer; 2014. pp. 17-40

[39] Ohly HP. Mission, programs, and challenges of knowledge organization. Adv Knowl Organ. 2012;**13**:25-33

[40] Kern VM. O sistemismo de Bunge: Fundamentos, abordagem metodológica e aplicação a sistemas de informação. Anais. 2011;(i):2693-709

[41] Zeng, M. L. Knowledge organization systems (KOS). Knowl Org. 2008;**35**(2-3):160-82