Journal of Interdisciplinary Methodologies and Issues in Science

Mixing Biology and Computer Science Concepts to Design Resilient Data Lakes

Marzieh DERAKHSHANNIA[1], Anne LAURENT[2], Arnaud MARTIN[3]

[1]LIRMM, Univ Montpellier, CNRS, Montpellier, France; dm.derakhshannia@gmail.com

[2]LIRMM, Univ Montpellier, CNRS, Montpellier, France; anne.laurent@umontpellier.fr,

[3]CEFE, Univ Montpellier, CNRS, Montpellier, France; arnaud.martin@umontpellier.fr

DOI: 10.46298/jimis.11449
Submitted:10/02/2022 – Accepted: 26/06/2023
Issue: **11** – Année : **2023**
Volume: **Penser l'interdisciplinaire en pratique**
Editors: *Déborah Nourrit, Guillaume Alevêque, Anne Laurent, Thérèse Libourel*

Abstract

Data lakes appeared a few years ago, introduced in particular to meet the challenges of storing and exploiting IoT data. They were first considered as a new technical and commercial tool, sold by the main database software editors. More recently, they have become the subject of research, in particular to define what a data lake should be, what it should provide in terms of services, and how it should be built. In this work, we have tried to return to the origins of data lakes, starting from the name "lake". We present here how we worked, between biologists and computer scientists, to understand the links between natural and data lakes. In this article, we first explore the links between the disciplines of biology and computer science before declining these links for the particular theme of lakes. This could appear as a work of transferring knowledge from biology to computer science, and a "simple" application of the concepts. However, we had to interact and understand each other's concepts and issues to align a possible comparison between the disciplines, for example to determine at what scale to establish the biological comparison, from DNA to the more macro system of the animal and plant ecosystem present in a natural lake. For this reason, we are inspired by a hybrid method based on ecological and logistical network topology to propose the resilient structure for the data lake. Thus, we use the Ecological Network Analysis (ENA) as a bio-inspired method and Graph theory as a logistical-inspired framework to study the interdisciplinary resilience strategies for the data lake network.

Keywords

Biomimetics, Data Lake, Supply Chain, Natural system, Ecological Network analysis, Graph theory

I BIOLOGY AND COMPUTER SCIENCE: TWO FIELDS THAT SHARE SO MUCH-ORIGIN OF OUR WORK

Man has always been inspired by nature *lato sensu* for his inventions (Leonardo da Vinci, the measurement of things, international units). Biomimetics regularly appears as a field of all possibilities to draw ideas that allow us to develop new technologies. For example, the invention of “Velcro” imagined from seed hooks clinging to the fleece of sheep, to more recent examples such as water-repellent materials (Jeong et al., 2009). However, biomimetics is not so much an exchange between two disciplines, for example between biology and physics, but rather an imitation. What interests us is the exchanges between two or more disciplines and to see if they can really learn from each other. Our approach has more been initiated by the curiosity of the exchange more than by a purpose linked to yet another biomimetics experience.

In our opinion, biology and computer science are two disciplines whose latest developments are taking place quite synchronously. Indeed, if it is admitted that computer science is a young discipline, biology gives the impression of being very old. However, major advances in biology (for example, genetics, evolution and even ecology) are very recent. Thus, these two disciplines are currently in research of evolution of concepts, paradigms, etc. For us, this is the right time to develop exchanges between disciplines with a view to developing concepts. Natural and biological strategies are the primordial source of inspiration for innovative and effective approaches in other disciplines, particularly in computer science (Yang and He, 2020; Beyer and Schwefel, 2002). For instance, mechanisms of species evolution (such as combination, mutation, drift, and selection), competition, mutualism, parasitism, and predation are extensively used in the field of computer science, such as the development and design of complex algorithms, improvement of the AI framework for computer programming (coding), development of knowledge discovery approaches, and management of complex network systems with appropriate strategies. (Bush et al., 2001; Fister Jr et al., 2013; Chatterjee and Layton, 2020; Yu and Gen, 2010; Boucher, James, and Keeler, 1982.)

James Dixon is generally credited with coining the term “data lake” (Dixon, 2010). He describes a data mart (a subset of a data warehouse) as akin to a bottle of water… "cleansed, packaged and structured for easy consumption" while a data lake is more like a body of water in its natural state. Data flows from the streams (the source systems) to the lake. Users have access to the lake to examine, take samples or dive in. (Dixon, 2010), the inventor of the data lake concept, probably stopped at an isotropic system with the vision of the volume of water compared to the volume of data that is to be "contained" in a sort of reservoir for future use, without fine vision of the content of a natural lake.

In our work, we wanted to go further by trying to see if the metaphor could be extended beyond this simple initial comparison. Our starting hypothesis was that there must be links to be explored, that it was worth investigating how we could go beyond Dixon’s vision. To carry out this work, our collaborative work approach is based on a shared attitude that could be assimilated to that of the university teacher who reflects on the foundations of the subject thanks to his scientist background and intellectual posture (identifying the right questions, the research avenues to investigate…) This approach was made possible by the opportunity of the meeting and the prior knowledge of empathy, open-mindedness, curiosity, and a posture open to interdisciplinarity. We recognized each other in this “difference” to the usual posture of our colleagues. This mutual curiosity was not only the essential base but also the glue between our disciplines. It is yet an asymmetrical curiosity. From the IT side: we use a metaphor in my field, but I would like to know a little more about the concepts.

On the side of the biologist: you use concepts from my discipline, I would like to know why and how. And how can we go further to exploit this approach to go beyond the simplistic imitation. Indeed, the previous approach by Dixon is seen as annoying since it is based on a narrow vision of a natural lake seen as “only” a volume of water, erasing the aspects of ecology and the living. The methodology of the exchanges has been based on a form of “active listening” attitude so that everyone understands the questions of the other (Rogers, 1966; Trong, 2016). We are also relying on “analogy” that is often used in such interdisciplinary works (Barbot, Miclet, and Prade, 2019; Haaparanta, 1992).

Our work has thus been built between computer scientists working on data lakes questions and a biologist who was first asked, “how does a natural lake work?” When facing such a question, the researcher in biology obviously answers with the latest concepts from ecology and evolution, far from the representation of the natural lake in high school books. In our work, it was assumed that the answers of the biologist will bring answers and new questions (especially questions) to the computer scientist. We had in mind that our exchange could also allow the biologist to question her discipline. For a biologist, a lake seen as an ecosystem is a very complex system in which living organisms are likened to envelopes of data (genes). An ecosystem is also seen through its ability to evolve, to undergo disturbances while maintaining essential functions. Thus, the questions of stability, resilience and homeostasis of the system easily challenge the computer scientist confronted with questions of the same type.

II Contribution of biology to Computer Science: Ecological Metaphor

As explained above, a data lake can be seen as a flat, stable, and isotropic structure. We imagine that it is this vision that interested Dixon when calling “lake” the new data architecture he was proposing. The horizontal dimension suggests that the data is not arranged according to a structure designed by a demiurge. The stable character suggests that it can be moved around without difficulty. Finally, the isotropic character suggests that we can look in all directions without distorting the quality of the information perceived. These three properties in no way suggest the immanent character of the data lake. They are just water molecules, and they give us no information, no structure, no meaning.

In ecology, a lake is not necessarily a flat structure (there are waves caused by the winds), nor stable (on a geological scale, a lake is born and then dies) nor isotropic (if you look below the surface horizontally you will see further than if you look towards the bottom, simply because of the incidence of light provided by the sun). In our world, a lake is not just water. From the formation of a lake, even at very high altitude, it is immediately colonized by living beings (first, viruses and bacteria that move freely in the atmosphere and fall with the rain, then larger organisms carried by the feet of birds, etc.).

Introducing the living into the metaphor of the data lake can be useful in the search for new concepts for information sciences for three reasons. The first is that the living is endowed with autopoiesis, that is to say, it reproduces. The second is that the living thing causes the homeostasis of the system (itself ensures the stability of the system in which it evolves, modulo some limits such as hurricanes, volcanic eruptions, the fall of large asteroids, the end of life of the solar system). The third, derived from the two previous ones, is the immanent character of the living. It contains within itself its own characteristics, its own values, without the intervention of the demiurge (except perhaps originally, but no one was there to see it).

Let us use the metaphor of what is alive in the lake to advance in the concepts of information sciences. A lake is formed, for example, after a period of glaciation. Most mountain lakes are born this way. At the beginning it is almost virgin of life (in reality, viruses and bacteria are innumerable in ice cream). As it warms, the lake will fill with more and more complex life (after viruses and bacteria, plants, invertebrates, fish, etc.). Erosion will bring sediments to it and the lake may eventually fill with life and mud and in the end die by disappearing. This ecological process is called eutrophication. If we draw a parallel between the data and the living in the history of a lake, we can imagine that there is a filling phase, a stabilization phase (homeostasis, the living reproduces and dies at the same time, the number of fish, bacteria, remains almost identical over generations) and finally a phase of transformation (end of the life of the lake, overproduction of living things, chemical imbalances, loss of oxygen, death of the Lake).

The filling phase corresponds to the information accumulation phase, the stabilization phase to the data use phase. A lake in the stabilization phase is a structured ecosystem in which the living has filled all the possible space. Watching how it works can give us ideas about the organization of information in an autopoietic system and in homeostasis.

Biologists speak of the fundamental brick of life, which would correspond to information in data lakes. For most biologists the fundamental building block is the gene, and it is he who contains the information of the living (Trivers and Dawkins, 1976). Richard Dawkins explains that living things are made up of genes that reproduce themselves through envelopes, organisms, which are seen as simple avatars of genes. One may wonder why there are so many different forms of life. We share identical genes with many species (97% homology with great apes such as chimpanzees or gorillas) and certain fundamental genes such as the one that codes for hemoglobin are almost identical in very many species. However, ecosystems are very diverse, and they appear to us as relatively stable structures where information seems to be organized, distributed and redistributed permanently.

The living reproduces itself (it is considered that even before the appearance of the first cells, self-replicating molecules have developed). It reproduces with a prolixity well above the acceptance level of the system. There are therefore regulations which are carried out by the mechanism of natural selection (only the fittest survive). Natural selection is the constraint of living beings. During reproduction, sexuality allows the mixing of genes and introduces a factor of chance (in addition to other phenomena such as mutations, for example) (Monod, 1975). Thus, the two forces that frame living beings are chance and necessity (constraint).

Information is not produced by random processes. Random processes only produce complexity in the sense of Kolmogoroff (Kolmogorov and Sevastyanov, 1947). Necessity produces information. Let's take an example on the movement of animals. For example, it may be for elephants to cross a forest to search for a resource. This action will be repeated over generations. The first animal makes its way "randomly", the second also and so on. Very quickly, paths exist and will be taken by the following ones because it is less costly in terms of energy than creating a new one. Then there will be a selection of the most practical paths (to circumvent natural obstacles for example). At the end, there remains a reduced network of paths which forms an optimum for the shortest and least expensive path, this network results from the effect of necessity (going towards the least expensive). "The combined action of chance and necessity conditions not only the information as we observe it, but also its evolution (Dessalles, Gaucherel, and Gouyon, 2016). If we follow the paths created by the elephants, we can consider that at any time chance can initiate evolution in a new way, while necessity will force the new path to remain functional.

Based on the above statement, the resilience of nature (as an evolutionary factor of nature) could be seen as a consequence of chance and necessity combination (Pavé, 2007; Monod, 1975). Regarding the data lake as a natural lake, these two factors (chance and necessity) play a similar role in enhancing resilience properties in this mimetic system. To encounter disturbances, the random processes search for the initial and reasonable decisions and the necessity conditions finds the best solution to keep the data lake in a stable state even in a disturbing situation. Therefore, this study aims to discover the possible nature-inspired strategies that are applicable to any systemic structure to improve flexibility in complex systems. We present below the results of our work, with particular emphasis on the issue of resilience.

III Resilience, an interdisciplinary Concept in Complex Systems

Resilience is one of the significant properties of any systemic structure. The ability of a system to resist against the predicted or unpredicted disrupted situations and its capability to recover itself promptly in case of external disturbance are the important characteristics of a resilient system. In natural systems, resilience is considered as an ecological feature that reduces the negative impacts of environmental perturbations (like unusual environmental changes) on the structure and functions of ecosystems. Moreover, in logistical systems like supply chain, to be resilient is the optimal property that reduces the supply chain's vulnerability and improves its restoration capacity against the disruption scenarios. The data lake is a centralized data management system whose systematic structure is inspired of natural and logistical systems. The data lake, like any systematic composition, is influenced by the internal disruption (the structure and performance failure) and external perturbation (attempted fraud or global server rupture). Therefore, the concept of network resilience, which is an indispensable issue for any network structure, could be generalized to this data storage system. For this reason, we concoct two ecological and logistical methods to study and assess the resilient structure for the data lake. In this hybrid method, we indicate the Ecological Network Analysis (ENA) as a bio-inspired method and Graph theory as a logistical-inspired framework to investigate the data lake network resilience issues. This concatenation is a result of the mimetic properties of the data lake structure. Firstly, we refer to the network-based feature of the data lake to take advantage of the graph theory ontology secondly, we address the nature-based philosophy of the data lake (data as species) to leverage ENA metrics for a resilience assessment framework.

The concept of resilience refers to the capacities of any systematic structure (ecosystems, logistics systems, computer systems, biological systems, nervous system, etc.) to equip all the entities of the system for disturbance situations and to quickly restore them to their initial state (Rezapour, Farahani, and Pourakbar, 2017). Therefore, to make a network resilient, two important phases of system preparation are required, pre-disruption mitigation and post-disruption recovery. While it is difficult to define the universal quantitative or qualitative measures to verify the level of resilience (Chatterjee and Layton, 2020) indicates that certain characteristics such as survivability, recovery time and recoverability could be the effective criteria for evaluating the resilience capabilities of a system. According to these studies, survivability measures the ratio of maximum and minimum system functionality before and after disruption, time recovery means the return duration time from disturbed state to the normal state, and recoverability refers to the fraction of system function before and after the disruption event).

In the natural systems, the resilience capability is considered as one of the significant ecological properties to preserve homeostatic functions and to restore the steady and optimal conditions after the natural or human-origin disturbed events (Holling, 1973). Ecological resilience could be measured by some characteristics like inertia (degree of resistance), elasticity (recovery time), or malleability (the degree of recovery ability) (Westman, 1978). Based on systematic features, the actor's (species) competence of ecological systems and their interactions throughout the ecological network could influence natural resilience and adeptness. One of the popular ecological systems that are a source of inspiration for many organizational and engineering systems is the food web. The network structure of the food web is based on graph theory in which the nodes represent what-eats-what and arcs illustrate the food relationship flow. In order to analyze ecological systems like the food web and monitor the functionality of their principal members, many tools and strategies are proposed by ecologists. Among these tools, the Ecological Network Analysis (ENA) is more practical, whether for ecosystem study, or for bio-inspired systems research (Fath et al., 2007). ENA provides ecological metrics like ecological fitness function to assess natural characteristics and functions such as resilience, in terms of the interaction of network members in the context of the graph structure.

From the viewpoint of logistical systems like supply chain, resilience property refers to the managerial strategies to provide the facilities for chain participants to confront network disruptions and could return to a former position and state without minimum delay and restoration costs. Like any system, supply chain disruptions may be caused by local level like a firm's interruption or by network-level like unforeseen disasters. However, the global supply chain network structure plays an important role in managing the resilience strategies (Kim, Chen, and Linderman, 2015). For this reason, the supply chain network design and infrastructure evaluation are considered as the consequential decisions in the fields of resilient supply chain management. Based on the importance of supply chain network design on its resilience quality, many criteria and strategies are conducted to assess the level of resilience in supply chain. One of the usual methods for analyzing the structure of supply chains is graph theory. The supply chain structure like the food web follows the principles of graph theory in which each node represents the individual chain member, and the arcs indicate the logistical flow of material or information throughout the chain (Borgatti and Li, 2009). Considering the supply chain as a network of nodes and arcs helps decision makers to monitor all the probable local and global disruptions and propose the proper plans to design the resilient supply chain (Falasca, Zobel, and Cook, 2008). Based on this requirement, some network metrics are proposed to evaluate the level of resilience in logistical systems like, network density, connectivity, centralization, flexibility, and node criticality (Kleindorfer and Saad, 2005; Stauffer, 2003; Kim, Chen, and Linderman, 2015).

In terms of the data management systems like the data lake, resilience could be defined as the hardware and software potentiality that prevent the system from internal network perturbations such as machine failure or external disturbances like cybernetic attacks (Pilania and Chiueh, 2005). However, a data lake that is a centralized data storage system has a bio-inspired and logistics-inspired definition. The data lake, likewise, the traditional data management repositories, supports all kinds of data structures in their native formats. This bio-inspired feature of the data lake is borrowed from the natural lake where all kinds of species with different organs live in a commonplace. On the other hand, data in the data lake is considered a precious product for any organization that could add value in the long term. Therefore, the lifecycle of data in data lakes from generation to consumption by users reflects a logistical-inspired process in which different data preparation zones interact to make available the product or service (data) for final uses. Based on these mimic definitions, the data lake system obeys the network design rules like graph theory where each node refers to the special phase of data lifecycle (ingestion, storage, process, and access) and each arc refers to the forward and backward data flow in the data lake.

The issue of the resilient computer system is addressed in many types of research in computer science especially in the cybernetic field (Kharchenko et al., 2020). However, the structure of computer systems, the facilities, and the technologies which are used have a direct impact on network resilience. As we discussed, the data lakes as a data storage and management system, are launched by a network of hardware and software faculties who create the general infrastructure of the data lake. However, the resilient data lake and the metrics for evaluating the level of resilience, especially in terms of data lake architecture, have received less attention. For this reason, we intend to employ the mimic strategies and metrics to analyze the data lake vulnerabilities, recognize the probable disruption events, and assess the structure of the resilient data lake. In many studies, bio-inspired methods have been used to design and determine the resilient supply chain network. Based on the results of this research and the ecological and logistical architecture of the data lake, we can propose hybrid approaches which are derived from natural and logistical systems like graph theory and ENA, for assessing, designing, or redesigning the resilient data lake structure.

IV Methodology

The data lake is almost a recent generation of data management systems in the big data arena. Therefore, many problematic issues like data management and governance, optimal architecture, agile network structure, and improving quality services had been less addressed in the literature of this domain. The resilient data lake structure is one of the considerable challenges that must be taken into account as a risk management solution to confront uncertain situations and mitigate the vulnerabilities. Concerning the lack of a practical solution to make the data lake architecture resilient in terms of the infrastructure failure situations, we use a hybrid method of ecological and logistical approaches to address this issue. One the fundamental study in mimicking strategies in this field is the work of (Chatterjee and Layton, 2020) that shows the bio-inspired method could minimize the supply chain costs and efficiently improve chain network resilience. The logistic architecture of the data lake conducted us to use a similar bio-inspired method of supply chain optimization to design a resilient data lake. To introduce the driven method, the structure of the data lake must be defined in accordance with graph theory and an analogy of ecological system (like food web [chain] and logistics system [like supply chain]). The figure 1 shows graph schema of three networks, supply chain, data lake and food web in analogical manner. Based on the schema of figure 1, Table 1 shows the notations of graph theory for three systematic structures.

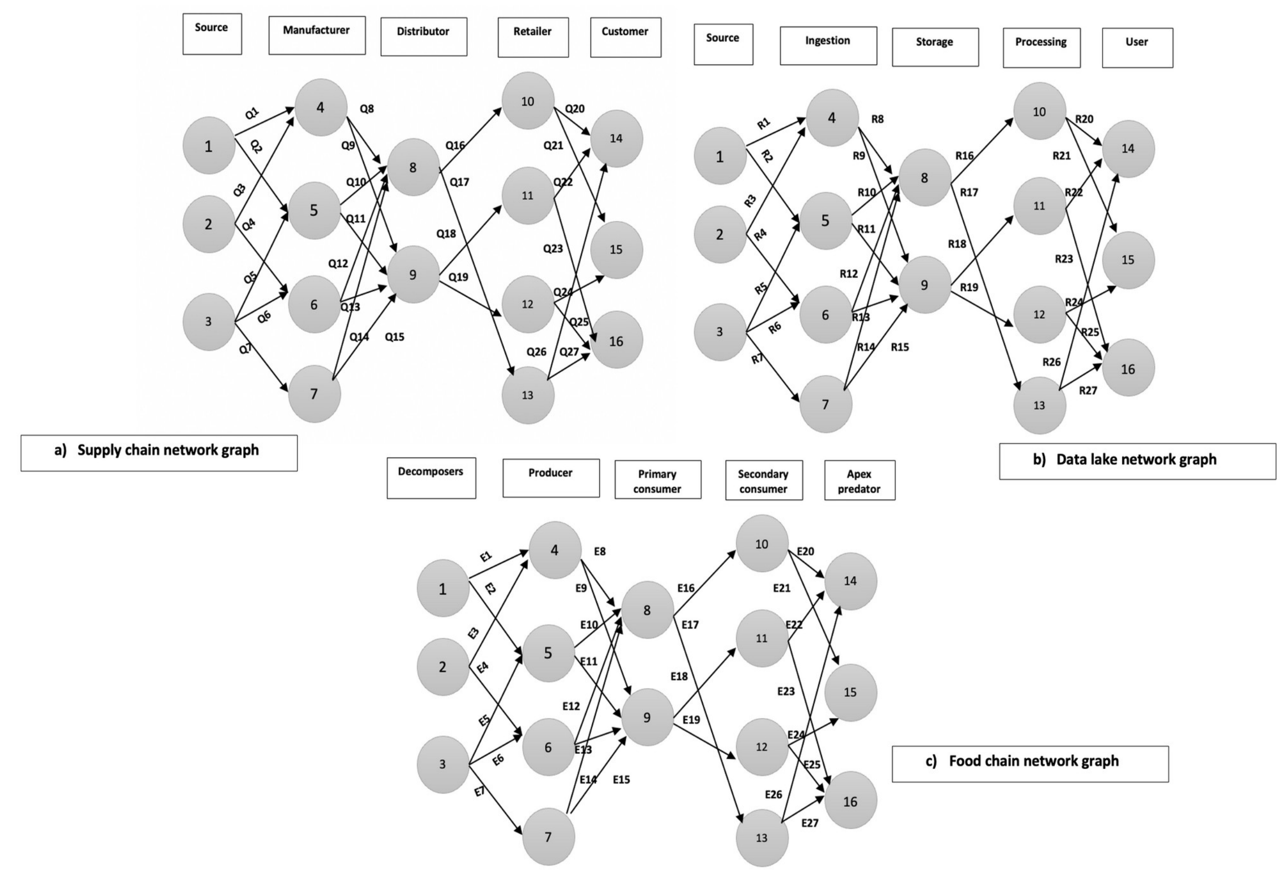

Figure 1: Three systemic network structures based on graph theory.

Based on graph theory, the disruption may occur on nodes, on arcs, or on the global network (Kim, Chen, and Linderman, 2015). However, all the graph compositions are interdependent and any disruptions on one element could engage other parts. In the same way, data lake graph disruptions are categorized in node perturbations (machine failure, high execution time, data loss, data swamp), in arcs interruptions (semantics errors, data traffic, data mapping complexity), or in-network disruptions (malfunctions of system hardware or software, network congestion and overload) (Revathi and Muneeswaran, 2010).

The ecological system as an intelligent self-repair system is a reliable source of inspiration for designing resilient systems. The ENA metrics provided by ecologists like, *flow matrix*, *total system throughput*, *average mutual information*, *Shannon index*, *ecological fitness functions*, and *conditional entropy* make available the mimic plan of resilience system analysis or design in any network structures (Chatterjee and Layton, 2020; Holling, 1973). This plan, that has been experimented with resilient supply chain design, could be an efficient method for the data management and storage systems in an environment like a data lake. The ENA metrics propose the framework based on graph theory to design the data lake structure with minimum disruptions in nodes, arcs, and global networks. This method determines the proper number of nodes and the efficient amount of data rate flow to reach the best amount of system fitness function that proves the appropriate state of resilience in the data lake.

Network system			
Graph notation	**Food chain**	**Supply chain**	**Data lake**
		Members	Pipeline layers
Node	Species	(Supplier, Manufacturer, Distributors, Retailers)	(Ingestion, Storage, Process, Access)
Arc	Food chain relationship	Product and information flow	Data flow
Start node (s)	Producer organisms	Raw materials	Data sources
End node (s)	Strong Predators	Customers	Users
Flow rate metric	Food energy (E)	Products quantity or shipment numbers (Q)	Data rate (mbps) (R)
Walk (A sequence of nodes & arcs)	Who eats Who	What creates what	What transforms what

Table 1: Graph theory notations for systematic structures.

As a result of this multidisciplinary comparison of different network structures, we can observe the effectiveness of interdisciplinary methods in solving complex problems. On the other hand, a wide range of unknown or undiscovered problems in each domain could be uncovered through discovery of cross-disciplinary solutions and methods. The research shows that the productivity of problem-solving strategies could be enhanced by the corporation of appropriate international, cross-cultural, and cross-disciplinary experiences. This effectiveness is gained from the competence of interdisciplinary methods to solve complex problems due to the diverse perspectives and alternative solutions (Vincenti, 2001). The study of mimetic strategies to design resilient data lakes, which has not received sufficient attention in the literature, is the culmination of this interdisciplinary brainstorming and multifaceted reflection. Our collaborative work was deemed quite successful thanks to the international experiences and cross-disciplinary knowledge of our colleagues.

V Conclusion

The work presented in this article required going beyond the simple application of concepts. However, it remains asymmetrical because it relies more on the transfer of concepts from biology to computer science and data science. The scientific contribution in the field of biology is less important, although there are other works in the literature showing a contribution of computer science to biology (Dessalles, Gaucherel, and Gouyon, 2016). However, examining the effectiveness of the proposed strategies on the technical architecture remains an important perspective to be discovered.

Our work required a lot of curiosity on the part of the researchers, and a great ability to question one's science on each side, to dig into the foundations of the concepts… The work led to reflections at a "philosophical" level, in an approach that could be described as "Socratic", constantly requiring "returning to the starting question" …

This work has led to interesting advances on the data lake side, making it possible to optimize their architecture. More broadly, the exchanges also made it possible to work on the question of the "death of data" which is a major issue for the sustainability of digital solutions with regard to the environmental impact (Derakhshannia et al., 2020). In nature, in fact, death is omnipresent and allows systems to remain resilient and to maintain themselves.

References

Barbot N., Miclet L., Prade H. (2019). Analogy between concepts. *Artificial Intelligence* 275, 487–539.

Borgatti S. P., Li X. (2009). On social network analysis in a supply chain context. *Journal of supply chain management* 45(2), 5–22.

Chatterjee A., Layton A. (2020). Mimicking nature for resilient resource and infrastructure network design. *Reliability Engineering & System Safety* 204, 107-142.

Derakhshannia M., Gervet C., Hajj-Hassan H., Laurent A., Martin A. (2020). Data lake governance: Towards a systemic and natural ecosystem analogy. *Future internet* 12(8), 126.

Dessalles J.-L., Gaucherel C., Gouyon P.-H. (2016). *Le fil de la vie: la face immatérielle du vivant*. Odile Jacob.

Dixon J. (2010). Pentaho, hadoop, and data lakes. https://jamesdixon.wordpress.com/2010/10/ 14/pentaho-hadoop-and-data-lakes/

Falasca M., Zobel C. W., Cook D. (2008). A decision support framework to assess supply chain resilience. In Proceedings of the 5th International ISCRAM Conference, pp. 596–605. Washington, DC.

Fath B. D., Scharler U. M., Ulanowicz R. E., Hannon B. (2007). Ecological network analysis: network construction. *Ecological modelling* 208(1), 49–55.

Holling C. S. (1973). Resilience and stability of ecological systems. *Annual review of ecology and systematics*, 1–23.

Jeong H. E., Lee J.-K., Kim H. N., Moon S. H., Suh K. Y. (2009). A nontransferring dry adhesive with hierarchical polymer nanohairs. *Proceedings of the National Academy of Sciences* 106(14), 5639–5644.

Kharchenko V., Dotsenko S., Ponochovnyi Y., Illiashenko O. (2020). Cybernetic approach to developing resilient systems: Concept, models and application. *Information & Security* 47(1), 77–90.

Kim Y., Chen Y.-S., Linderman K. (2015). Supply network disruption and resilience: A network structural perspective. *Journal of operations Management* 33, 43–59.

Kleindorfer P. R., Saad G. H. (2005). Managing disruption risks in supply chains. *Production and operations management* 14(1), 53–68.

Kolmogorov A. N., Sevastyanov B. (1947). The calculation of final probabilities for branching random processes. In Doklady, Volume 56, pp. 783–786.

Monod J. (1975). Chance and necessity: An essay on the natural philosophy of modern biology. *Religious Studies* 11(3), 366–368.

Pavé A. (2007). Necessity of chance: biological roulettes and biodiversity. Comptes rendus biologies 330(3), 189–198.

Chiueh T. C. & Pilania D. (2005). Design, implementation, and evaluation of a repairable database management system. In 21st International Conference on Data Engineering (ICDE'05) (pp. 1024-1035). IEEE.

Revathi T., Muneeswaran K. (2010). Congestion control using soft computing. In Soft Computing Applications for Database Technologies: Techniques and Issues, pp. 271–289. IGI Global.

Rezapour S., Farahani R. Z., Pourakbar M. (2017). Resilient supply chain network design under competition: A case study. European Journal of Operational Research 259(3), 1017–1035.

Rogers C. R. (1966). Le développement de la personne. Les Études philosophiques 21(4).

Stauffer D. (2003). Risk: The weak link in your supply chain. Harvard Management Update 8(3), 3–5.

Trivers R., Dawkins R. (1976). The selfish gene. New York: Oxford University Press.

Trong M. V. N. (2016). Impact d'une formation à la relation thérapeutique (FRT) sur les capacités empathiques en DFASM1 de médecine: analyse qualitative d'un groupe Balint de 7 séances. Thèse de doctorat.

Vincenti V. B. (2001). Exploration of the relationship between international experiences and the interdisciplinary work of university faculty. Journal of Studies in International Education 5(1), 42–63.

Westman W. E. (1978). Measuring the inertia and resilience of ecosystems. BioScience 28(11), 705–710.