Article

Elastic Stack and GRAPHYP Knowledge Graph of Web Usage: A Win–Win Workflow for Semantic Interoperability in Decision Making

Otmane Azeroual [1,2,*], Renaud Fabre [3], Uta Störl [2] and Ruidong Qi [4]

1 German Centre for Higher Education Research and Science Studies (DZHW), 10117 Berlin, Germany
2 Chair of Databases and Information Systems, University of Hagen, 58097 Hagen, Germany; uta.stoerl@fernuni-hagen.de
3 Dionysian Economics Lab (LED), University Paris 8, 93200 Saint-Denis, France; renaud.fabre01@gmail.com
4 College of Computer Science, Inner Mongolia University, Hohhot 010031, China; csqrdong@mail.imu.edu.cn
* Correspondence: azeroual@dzhw.eu; Tel.: +49-30-2064177-38

Abstract: The use of Elastic Stack (ELK) solutions and Knowledge Graphs (KGs) has attracted a lot of attention lately, with promises of vastly improving business performance based on new business insights and better decisions. This allows organizations not only to reap the ultimate benefits of data governance but also to consider the widest possible range of relevant information when deciding their next steps. In this paper, we examine how data management and data visualization are used in organizations that use ELK solutions to collect integrated data from different sources in one place and visualize and analyze them in near-real time. We also present some interpretable Knowledge Graphs, GRAPHYP, which are innovative by processing an analytical information geometry and can be used together with an ELK to improve data quality and visualize the data to make informed decisions in organizations. Good decisions are the backbone of successful organizations. Ultimately, this research is about integrating a combined solution between ELK and SKG GRAPHYP and showing users the advantages in this area.

Keywords: data management; databases and information systems; IoT; knowledge graph (KG); scientific knowledge graph (SKG); GRAPHYP; elastic stack (ELK); visualization; data quality

1. Introduction

Semantic interoperability: Bridging data discovery and data re-use in strategic decision making.

Semantic interoperability requires integrated solutions for all steps where data need to be processed. In this context, innovative use of the web draws attention to levels where strategic decisions may require timely and relevant responses. However, since web usage data contain all the knowledge from ancestors, only comprehensive, efficient analytical modeling could delineate new knowledge from experienced decisions, data reuse, and data discovery, thereby directing the pace of browsing activity toward meaningful assessments.

As we would like to emphasize below, a systemic valorization of web usage in the detection of innovative features is an important factor for efficient complementarity of Elastic Stack (ELK) and the GRAPHYP Knowledge Graph (GR) in efficiently processing semantic interoperability Big Data contexts. The respective features of ELK and GR could be appreciated as follows.

A recent review of various databases suggests that ELK meets the key requirements where resiliency needs to be demonstrated and tested, as ELK is an optimal solution for data management in any context while "allowing users to easily monitor their blockchain systems" [1]. However, in current strategic decision making, ELK needs to be connected to

a system that prepares data for contextualized sensory tasks: Knowledge Graphs (KGs) seem to be the best solution.

A large consensus in the literature states that "the accurate derivation and justification of KG plays an important role in controlling the operation and decision-making of energy plants" [2]. Nevertheless, research is still open to new solutions on the expressiveness and efficiency of KG; the important question of knowledge justification in this context remains a critical and challenging issue [3]. This edition recently benefited from the new results of KG GRAPHYP. This novel framework—a modeling web usage model for creating web graphs (https://webgraph.di.unimi.it/ (accessed on 15 May 2023)) from log files—develops a novel additional browsing function as a "scout" for discoveries, highlighting the challenges of decision making, with an analytical typology of recordings and a compass for explorable paths to discovery [4].

ELK and GR could thus appear as systems that benefit from synergies and possible functional complementarities in meeting strategic requirements in data discovery and reuse. In the following, we deal with identifying an optimal articulation of ELK and GR at the levels of three research questions; in the meantime, we outline the conditions of other real decision applications.

Surprisingly, attention to the science appears to be scant: an up-to-date overview of current and future research on semantic interoperability in meaning-making activities hardly accounts for new features of the emerging workflow of current scientific knowledge assembly. Meanwhile, research in the areas of informed decision making and meaning-finding technologies seems to point to the same behaviors [5].

For our side, we underline the new venues of the late birth diagnosis of "multimodal knowledge acquisition" [6]. At the same time, research on digital twins shows the need for further monitoring and control requirements for their real industrial applications from a communication and computing perspective [7].

While research diagnostics are still not very oriented, the area of Big Data has also grown more and more in recent years [8]. Therefore, the question has evolved as to how to collect and store these large volumes and derive valuable insights from these unstructured data. In the following, our answer depends on three research questions:

- RQ1: Is ELK a suitable solution for decision makers when analyzing and visualizing data?

The ELK can provide an open-source answer to the question of gaining insights from big datasets [9]. The stack is intended to enable the analysis and visualization of data and consists of three programs that build on one another: Logstash, Elasticsearch, and Kibana. Logstash represents a pipeline that can read data, prepare them, and then forward them to Elasticsearch. Elasticsearch enables distributed storage, full-text search, and analysis of the data. Elasticsearch capabilities can be leveraged in Kibana to create visualizations and metrics.

The first goal of this paper is to investigate whether the ELK is a suitable method to analyze and visualize data and help end users in decision making. To investigate this, a test setup is created with a requirements catalog that uses all components of the ELK and evaluates the suitability of the ELK.

- RQ2: Does GR bring data analytics that could benefit from ELK and optimize services for strategic decision making?

Our research question will explore the efficient and resilient new frontiers involved in optimizing the expressiveness of KGs and their ability to offload semantic interoperability in digital twin environments. We address the challenge of semantic interoperability by experiencing applications of the new generation of interpretable graphs (IGs) [10] and examining how knowledge acquisition is changing in interactions between ELK and the type of KG in which GRAPHYP innovates.

GRAPHYP innovates by deciphering relevant interactions in the information structure of the geometry of graphs and visualizing their categories of possible interpretations at the expected scale.

IGs with appropriate algorithms will lead to a conceptual framework for managing these two mutually influencing analytical requirements. The first requirement is a toolkit: a systemic geometric mapping of the user's heuristics in any cognitive community with digital twins obtained and analyzed from their connection logs on selected keywords. The second requirement is a step-by-step schema of an end-user as the rater transitions between the contextual challenges of their cognitive discoveries [11]. The end goal is that the end user can select the "best possible" search pathway among the recognized and geometrically designed "possible" paths resulting from the selection. These key features describe the value that the bi-partite symmetric scientific Knowledge Graphs (SKGs) could add significantly to a fully efficient ELK.

- RQ3: How could future benefits of a merger of ELK and GR be presented?

In this last research question, we want to show the integration of ELK and GR and discuss the benefits that can result for the user.

Our paper contains six sections. Section 1, the introduction, provides an overview of the topic. Section 2 depicts our research methodology and approach. Section 3 describes the materials on the ELK concept and its optimal capacities for data discovery and reuse, and presents a catalog of requirements when implementing ELK. Section 4 presents the GRAPHYP results as an explainable Knowledge Graph and shows how synergies and interactions of GRAPHYP and ELK could create a unique collaboration and also their potential for informed strategic decisions. After that, the further developments of web use are analyzed in coordinated approaches of ELK and GR. Section 5 discusses the results of the work, and in Section 6, the results are summarized.

2. Research Methodology and Approach

Our research methodology is user-centric and we analyze systems based on the outcomes and benefits of ELK and GR for decision makers. Understanding users' dataset behavior is essential to providing effective data discovery services [12]. From the search for predecessors, the analysis of log files has developed a variety of methods that efficiently determine contextualized answers to user needs. A helpful approach is to capture, at best, the "relatively relevant" characteristics of user needs in each identified context [13]. With GRAPHYP KG, we propose a comprehensive modeling of the detection and responses of a user-centric approach: the personalization of services supposedly depends on an effective modeling process [14].

When data are stored in Elasticsearch, be it product descriptions, user data, documents, or logs, they often contain properties or references that can describe relationships between objects, entities, people, or machines, for example. Visualization is the best way to analyze such relationships.

KG will play an increasingly important role in the future, since it is able to operate analytical connections between URLs and thus between linked hypertexts. It is about us collecting new types of information from different sources. The focus is on network effects, for example, to personalize products in this direction. GRAPHYP offers a unique solution for linking and displaying in comprehensive modeling, a first analysis system of past web use, and KG modeling of usability test logs at any scale.

The methodological approach of this paper is based on a multi-layered literature review with the aim of finding answers to our three research questions. We focus on questions pertaining to theoretical and applied characteristics of web usage as a source of innovative strategic decision making in both ELK and GR contexts. Relevant publications were researched and prepared. There is a gap in the literature dedicated to clarifying the relationship between ELK and GRAPHYP, their integration, and their benefits. Therefore, we want to develop an approach or a design solution for researchers in this field.

3. Materials

In this section, we briefly describe and discuss our materials and methods.

3.1. Elastic Stack (ELK) Concept

The task of the ELK is the collection, storage, and analysis or visualization of data. Each of these tasks is accomplished by a different component, namely Elasticsearch, Logstash, and Kibana (hence the name ELK stack). Figure 1 shows the structure of the stack and the interactions between the components.
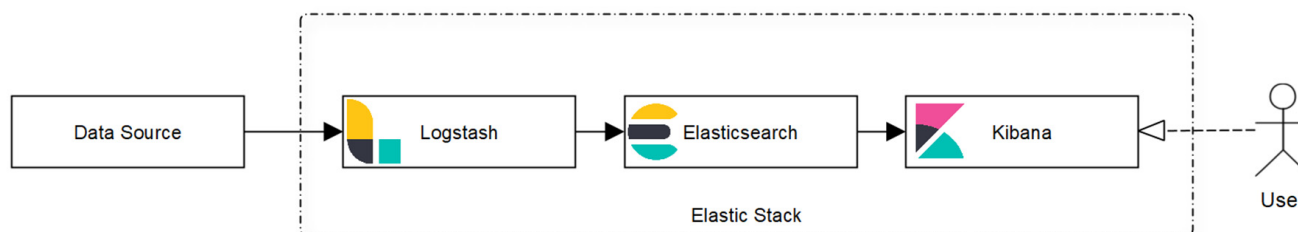

Figure 1. Elastic stack (ELK).

(i)   Logstash

Developed by Elastic companies such as Elasticsearch and Kibana, Logstash is an open-source data ingestion and processing engine that reads, transforms, manipulates, and forwards data streams and events from various sources [15]. It is written in Ruby and can be launched in a Java virtual machine environment. This is a pipeline that processes the incoming data stream with various filter plugins and forwards it to various systems. Logstash can be extended with plug-ins, so it supports many source and target systems. By default, Logstash creates tag-based indexes in the logstash-YYYY.MM.dd format. The index pattern can also be determined. Logstash requires a configuration file [16].

Logstash configuration files are in JSON format and are usually located in the */etc/logstash/conf.d* directory. A configuration file is structured as follows (see Listing 1).

Listing 1. Structure of a Logstash configuration file.

```
Input {
[...]
}
#Filters_are_optional
Filter {
[...]
}
Output {
[...]
}
```

The input and output parts are mandatory, while the filter section is only optional. Input is responsible for reading data from various sources (e.g., database, Apache server, file, Sysin, Beats, HTTP, etc.) at certain intervals. These data can then be parsed, structured, transformed, expanded, or reduced by information in the filter section, allowing the data to be analyzed quickly and have consistent formats and timestamps. Finally, one or more target destinations for the data must be configured in the output block in order to forward the data to them. Logstash offers many outputs where the data can be stored. The data can be easily dumped with Stdout and also stored in a database such as Elasticsearch [17]. An index pattern to be generated in Elasticsearch can be defined in Output and the path of the template can be entered. By default, the data are stored in the tag-based index "logstash-YYYY-MM.dd". Each of these 3 components requires different functions (plugins), which can be selected and configured as needed. Several functions can be used one after the other, which are applied in the order in which they are named.

(ii)   Elasticsearch

Elasticsearch is a distributed database with full-text search and analytics, based on the Apache Lucene Java library. The core of Elasticsearch is Apache Lucene, but it is enhanced with clustering, monitoring, aggregation, etc., hiding the complexities that come with search and analytics and providing rest interfaces for data exchange. Besides the Rest API, Elasticsearch also offers official clients for Java, JavaScript, Python, and .NET. The program searches and indexes documents of various formats and saves the search results in a NoSQL format (JSON), which can be easily grouped and evaluated with aggregation [18]. Wikipedia, Guardian, Stack Overflow, GitHub, etc., use Elasticsearch to meet their requirements such as high availability, scalability, and fast delivery of results. As seen in Figure 1, Elasticsearch has the role of Kibana storing and creating responses to queries.

Elasticsearch is designed to scale easily and without much effort. Although vertical scaling is possible, i.e., using computers with more power, horizontal scaling, in which the number of computers is increased, is more efficient and reliable because, for example, data loss in the event of a media failure can be prevented by having backup copies on one of the other computers [19]. Each Elasticsearch system consists of a network of one or more computers, also called clusters. Each of these computers represents a node, which divides the data and computing load among themselves, with one node in the computer network always having to be the main node. Main nodes are responsible for managing changes in the computer network such as adding an index or removing a node [20]. Users can make requests to any node in the network, since each node knows the location of a stored document. The query can thus be forwarded to the responsible body in order to provide the user with the answer. Nodes, in turn, consist of shards that contain the documents in indexed form and are stored and distributed across the nodes. Each shard represents an independent Lucene instance and thus a full-fledged search engine. There are 2 different types of shards. One is the primary shard, which contains a limited number (up to 128) of documents depending on the hardware [21]. Each document is only stored in a single primary shard. The second type, a replica shard, represents backup copies of primary shards in order to provide an alternative to a primary shard in the event of media errors or in the event of a large number of search or read requests [22]. The number of primary shards can be configured before creating an index, but not after. In the case of replica shards, on the other hand, the number per primary shard can be changed even after an index has been created. If there is more than one node in the computer network, the primary and replica shards are divided if possible in such a way that all documents are still available if one node fails. New documents are stored first on the primary and then in parallel on all associated replica shards. This mechanism is demonstrated in the following example.

In the example, additional nodes and shards are added to a computer network. At the start, the bond has 2 nodes and 3 primary shards, each with a replica shard. A failure of a node could thus be coped with even now, since each of them owns all the documents (see Figure 2).

Figure 2. Example with 2 nodes, 3 primary shards each with a replica shard.

If a node is added, the primary shards are automatically distributed in such a way that reliability and performance increase, since fewer shards now have to share a node (see Figure 3).
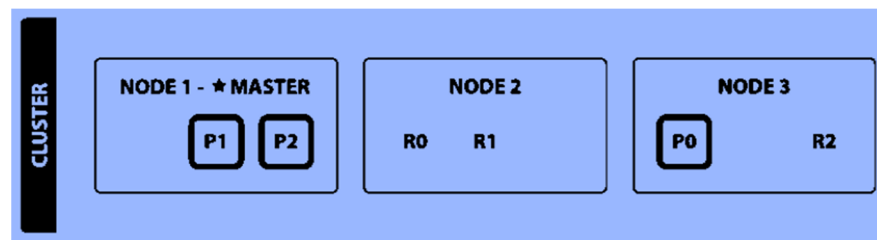
Figure 3. Example with 3 nodes, 3 primary shards with one replica shard each.

If the number of replica shards per primary shard is increased to two, the computer network looks as shown in Figure 4.
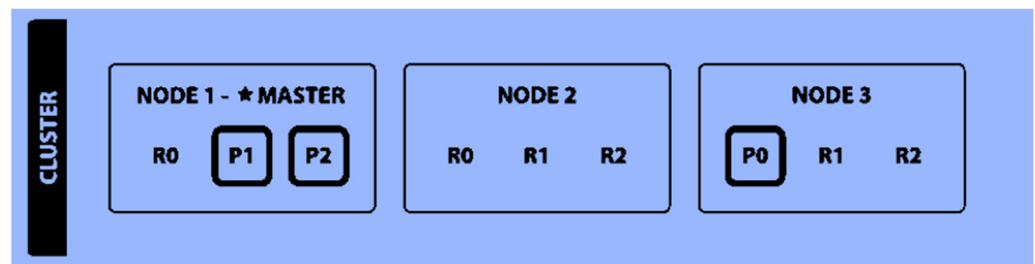

Figure 4. Example with 3 nodes, 3 primary shards with two replica shards each.

With this configuration, the failure of 2 nodes would still be manageable, since the data from all shards are stored on each node. If one of the failed nodes was the master node, the rest must immediately agree on a new master node.

There are basically two different ways of making a search query in Elasticsearch. On the one hand, there is the Lite search, in which the entire search query with its parameters consists of a query string [22]. On the other hand, there is the search using the Query Domain Specific Language (DSL for short), in which the query consists of a JSON object [23]. The Lite search is more compact and faster with a few parameters, but it is very confusing for more complex searches and therefore difficult to debug. Unless otherwise defined, only the top 10 hits are returned by default for performance reasons. Each search query is structured as follows and can be executed using the cURL command line program, for example [22].

Each data field has a mapping that defines each type of field so Elasticsearch knows how to handle the data. Every document in Elasticsearch is a JSON object and individual data fields have types even if they are not explicitly given [24]. The data can also be stored schema-free in Elasticsearch. When documents of unknown types are indexed, a new document type is created based on the content. New attributes can also be dynamically added to existing types. Elasticsearch uses dynamic mapping by default. A mapping to a type is defined to improve the search. The individual fields can be defined as date, number, string full text, or exact value. Elasticsearch defines many data types such as string, byte, short, integer, long, Boolean, date, object, array, etc. [25]. If no explicit mapping is entered for a document, Elasticsearch automatically creates a mapping as Boolean, long, double, and date depending on the value or string.

One of the most important modules of Elasticsearch is aggregation, which makes it possible to calculate, query, sort, and combine the datasets. Aggregation as a verb means joining parts together to formulate a whole [26]. With the aggregation in Elasticsearch, the datasets can be analyzed and evaluated. Aggregation is very powerful for reports and dashboards.

(iii) Kibana

Kibana is an open-source analytics tool designed to visualize data from Elasticsearch (see Figure 5).
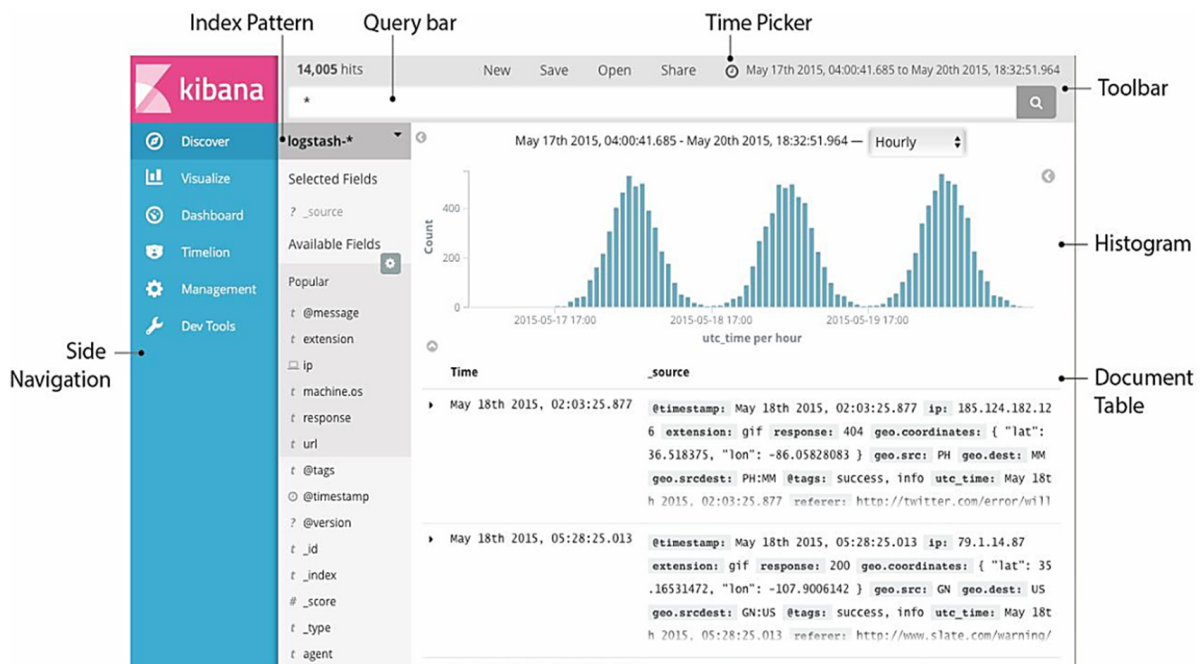
Figure 5. Overview of the structure of Kibana.

Kibana offers almost real-time analysis and flexible display options for the information stored in Elasticsearch [27]. Kibana accesses Elasticsearch data via the REST API and offers the user the possibility to filter the results as desired [28]. Kibana is implemented to work optimally with Elasticsearch. Many different views of the information can be generated through various display and aggregation options. Displays are arranged and organized on so-called dashboards. The display and views can be defined as required using the various configuration options. Kibana is very powerful and easy to use. Kibana offers many visualization tools to obtain an overview of computers, applications, and users. It offers many visualizations such as histogram, line chart, pie chart, metric chart, etc. The front end accepts the data from Elasticsearch and offers the user the possibility to filter the results as desired. The result is a dynamic, interactive, and attractive representation of the data in real time. The data stored in Elasticsearch's document-based structure can be explored and compiled into dashboards in visualizations such as pie and bar charts [26].

3.2. ELK Implementation Requirements

This section covers the requirements for using ELK. The requirements are intended to ensure that, after implementation, it can be evaluated whether the ELK is a suitable method for analyzing and visualizing data. Two different types of requirements can be placed on an experimental setup. The functional requirements deal with what functionalities a system should have. The non-functional requirements define the required quality properties for the overall system. These include, for example, performance, scalability, or reliability.

(i)   Functional requirements (FR)

    a.   FR 1: The retrieved data should be filtered and structured if necessary.

    b.   FR 2: Any missing information should be added.

    c.   FR 3: The data should be assigned to an index that contains, for example, the time of indexing.

    d.   FR 4: The data should be forwarded to an Elasticsearch system.

    e.   FR 5: Elasticsearch shall store the data received from Logstash.

    f.   FR 6: Elasticsearch should recognize and delete duplicates.

    g.   FR 7: Search or aggregation requests made by Kibana should be processed by Elasticsearch.

h. FR 8: Indexes stored in Elasticsearch should be searchable and retrievable in Kibana.
i. FR 9: Visualization and statistical values should be able to be created from these indices.
j. FR 10: Visualizations and statistical values should be accessible in an individually adaptable dashboard.
k. FR 11: The dashboard should be able to be updated automatically if data have been changed or added.
l. FR 12: It should be possible to organize and display data by time in order to analyze changes in the data over time.

(ii) Non-functional requirements (NFR)

a. NFR 1: ELK should be able to scale vertically and horizontally.
b. NFR 2: It should be possible to compensate for the failure of one or more computers in the system.
c. NFR 3: Actions such as searches or aggregations should deliver their results after a maximum of 2 s.
d. NFR 4: The dashboard should be accessible under the specific port.

Based on the requirements and the applications to be used, the experimental setup can be outlined as follows. The experimental setup should consist of three components, namely Logstash, Elasticsearch, and Kibana. Users then access Kibana, where the dashboard can be viewed with metrics and graphs, but also the data can be browsed and additional graphs can be added to the dashboard. Logstash works with configuration files, which consist of 3 components, namely input, filter, and output. Input is responsible for fetching data from various sources. These data can then be filtered and structured in the filter section to then be routed to their destination—in this case, Elasticsearch—in the output section. After the data have arrived in Elasticsearch, they are stored and indexed, and are now searchable or aggregable. Kibana uses these functions and provides users with the results in the form of datasets or graphics.

When implementing the ELK, the following points must be observed and configured accordingly:

- Logstash

  ○ How can the desired data be retrieved and what intervals make sense for this?
  ○ What information do the retrieved data contain?
  ○ What is the structure of the data?
  ○ Which index should the data be assigned to?

- Elasticsearch

  ○ How many computers are available?
  ○ How many primary and replica shards per index make sense?

- Kibana

  ○ How can the data be visualized well?
  ○ How can conclusions be drawn from the data? How can the dashboard be designed clearly?

In summary, functional and non-functional requirements continue to require the use of specific software from ELK. If ELK does not meet these requirements 100%, there are various tools that can be used as an alternative to ELK (such as Algolia, Splunk Enterprise, etc.).

4. Results

In this section, we show how synergies and interactions of GRAPHYP and ELK could create a unique collaboration and also their potential for informed strategic decisions.

ELK and GRAPHYP: A Win–Win Fit in Innovative Strategic Assessment

(1)　A case for win–win strategic assessment: Why GRAPHYP?

Using the GRAPHYP bipartite hypergraph, we design web usage captures from log files based on an explainable graph geometry that relates web graphs (https://webgraph.di.unimi.it/ (accessed on 15 May 2023)) to meaningful activities in an analytical typology of captures. Browsing could thus reach new frontiers, being equipped with a compass on explorable pathways to discovery: the users are experiencing multiverse reasoning on "possibilistic choices" between paths of linked URL, confronting at scale spatial and temporal grids structured by sub-graphs "geobots" as scouts to discovery. GRAPHYP thus proposes a novel approach to the breadth and depth of meaning, enabling "analytical" browsing and an original tool to support human artificial intelligence (AI).

The dynamics of today's networked learning inform each other while expanding the searchable space, the size and shape of hidden knowledge, and the demand for contextualized answers. Networked learning, therefore, requires advances in services that optimize AI support for human decisions recorded from logs and identify new "trade zones" for shared knowledge and potential directions for its interactive use. It has been tested that GRAPHYP belongs to the current generation of GNN representations as a hypergraph network with nonlinear activation functions applied to both hypernodes and hyperedges.

The approach proposes both representations of knowledge and its use, ranging from entries in users' log files to items. As a result, we design an original grid of diagrams of geometrically built robots ("geobots") that are represented by the mean of contextualized, searchable space that categorizes a typology of possible and recorded choices.

Geometry has recently become one of the main tools of data exploration, starting in molecular representations [29], with the goal of finding new applications for uninterpretable representations of information symmetry and dissymmetry. This approach of explainable geometries applied to graph interpretation is currently finding new extensions in KG analysis [30]. GRAPHYP is innovative because, like some other recent new approaches (see below), it allows the insertion of user-choice data into a predefined geometric architecture of subgraphs and positions information oriented and explained through the use of symmetric hypergraph geometry functions. It thus forms a lattice of "geometric robots" (geobots) capturing the cognitive communities of subgraphs. Such an interpretable forms approach does not seem so far removed from that of DeepMind's "relational networks", which propose "structured thinking even without structured inputs and outputs" and have the ambition to develop "structured learning" [31] ("Future work should apply RNs to a variety of problems that can benefit from structure learning and exploitation, such as rich scene understanding in RL agents, modeling social networks, and abstract problem solving. Future work could also improve the efficiency of RN computations. Though our results show that no knowledge about the particular relations among objects are necessary, RNs can exploit such knowledge if available or useful").

We then constructed graph similarity measures of knowledge use that allow for a specific grid of GRAPHYP, based on an "equivariant messaging network" [32]. Using the geometry of GRAPHYP, the activation function of neurons allows for the representation of all possible circuits of user choices according to the intensity and novelty of the preferences expressed by one keyword. Applications based on this contextualized log analysis of texts and images have two functions: first, they provide efficient and complete information about the user's practices (classical tasks, obtaining appropriate conceptual labels, uncovering trade zones between elements, pattern recognition, community recognition . . .); second, they enable original contextualized mapping and navigation between the quantitative and qualitative representations of mutual user/item information.

(2)　A case for win–win strategic assessment between ELK and GR: How?

ELK and GRAPHYP (GR) are systems that fully share the problems of data discovery and data reuse. We advocate that they together provide a unique opportunity to represent complementary between discovery, where ELK acts as a generator, and reuse, where GR

appears to currently be the most innovative discriminator to deliver personalized strategic decisions, delineated from GRAPHYP's analytics of web usage.

We assume that within the interaction framework of ELK and GR, those systems supply solutions for strategic decision making, combining, at the best possible level, flexibility and adaptation in the main strategic personalized issues of strategic decision making in any business.

Efficient search technologies are the background of strategic decision making: they optimize the adaptation of any kind of solution by extracting data features from an accurate analysis of web usage in the domain; the main condition of success is the optimal usability of testing logs. In turn, log modeling provides alternatives for building further strategic issues adapted to user demands. Moreover, efficient search technologies are decisive first-rank conditions to personalize messages in performative strategic initiatives and mining tasks that will shape up new answers on processes, contents, and linkages.

(3)　A common Challenge for ELK and GR future interactions

The current shortcomings in search technologies are discussed below.

Surprisingly, search technologies currently have common shortcomings: the adaptive efficiency of search systems advances without a general conceptual framework that could ensure fail-safe, accurate operational answers in every domain. Furthermore, the range of these is still wide open with no clear assessment of efficiency. At the same time, however, as search data analysis cannot be replaced, it remains essential to build up the background of strategic decision making by collecting, analyzing, and reusing the opinions of past users, mainly from log analysis. Near-real-time adaptative technologies have developed to offer the requested services of assistance to strategic decision making in any business context, with high refinement of explainable AI.

Explainability is the heart of innovative decision making. It combines new methodologies such as ELK with the digital activation of the familiar principle of informed decision making: "explainable AI" of KG such as GRAPHYP mitigates the uncertainty of "black box" approaches and helps interpret machine learning models with a large range of expected benefits, with the breakthrough of a human-in-the-loop approach.

The taxonomy of interpretability techniques lets us observe, in the array of methods, purposes, and models, the specificities and value added of integrated interactions between a generator of data such as ELK and a discriminator such as KG GRAPHYP [33]. We could qualify this approach as web-friendly, neutral to users, and human-in-the-loop. It thus radically differs from the "classical" demonstrative approaches (black box explanations or white box creations, sensitivity of predictions . . .) and supports any kind of proposed goal in any context.

This section answers the following questions: What and how is the main field of interactions between ELK and GR? What is the extent of a win–win partnership? What are the main components of GRAPHYP, and how do they potentially interact with ELK?

We first give a brief overview of the implementation requirements of KG GRAPHYPs; second, we highlight how and why interactions with ELK should be developed.

The Scientific Knowledge Graph (SKG) GRAPHYP was created with the support of the Scientific Knowledge Collection, which is currently making significant efforts to innovate the architecture of SKGs and to optimize and innovate the performance areas of the KG as a whole [34,35]. In addition, attention is rapidly evolving in many new directions that go beyond formal methods of knowledge representation and argumentation techniques (https://kwarc.info/ (accessed on 5 April 2023)). The requirements include both architecture and analytics.

- Architecture: A searchable space with "possible" choices

With the GRAPHYP community, we first coined the need for a "searchable space" in any situation where the availability of data allows differentiating approaches in analytics of existing practices and especially in decisive choices and assessments, in order to represent the selection of "possible choices" according to objective measurement parameters (see

further metrics); conversely, we consider that a structural analysis of search experience on the same data corpus is a mirror for the interpretation of the above-mentioned practices, as it reveals choices of predecessors on the same topic. Both operations, building a searchable space and representing search experience, being two sides of the same coin, allow bi-partite knowledge of typical representations of the architecture of information that exists or could be found on a given topic, combining, if necessary, probabilistic and possibilistic approaches.

KG allows realizing such a program, which we chose to represent with two interconnected subsystems: one for the representation of data, and the other for interpretation. One subsystem represents the "possible" options of the architecture of the analyzed dataset (see the above-quoted paper), while the second subsystem produces pathways to select the "best possible" option according to pre-existing prerequisites of the user on the topic, in a human-in-the-loop approach with multiple hops. That powerful reasoning workflow transposes, as a first application to the KG, the "multiverse" Leibnizian program for scientific reasoning [36]; we thus denominate SKG GRAPHYP as a novel "multiverse graph" ("A multiverse graph for scientific reasoning among adversarial tracks: mapping assessor's shifts from documentary pathways of explainable search experience") that applies the highly performative principles of multiverse analysis [37].

Figure 6 represents the articulation of tasks in the exploitation of data tracks in GR1.
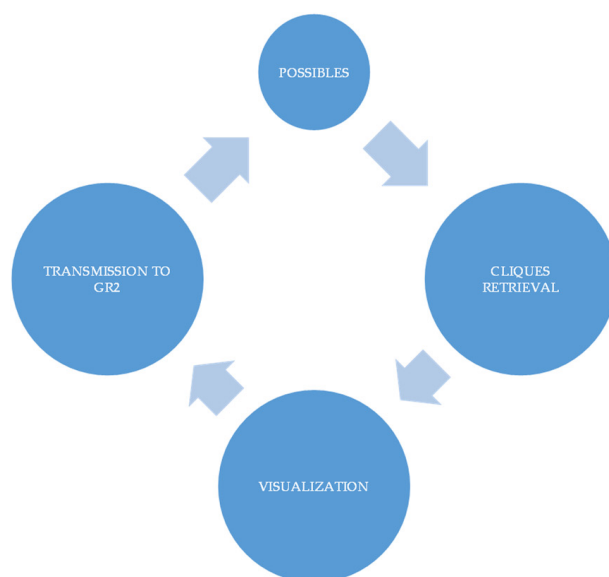

Figure 6. Steps in GENERATOR GR1: Extraction of possible data tracks.

- Reasoning with GRAPHYP: Pathways to the "best possible" option

GRAPHYP, a knowledge hypergraph, is born of the idea to represent any datum as being located in its modeled environment, within its manifold of structured possible adversarial uses. We assume that there is typicality of a set of data captured at the relevant scale on a given theme or keyword, as a "documentary track" of any search experience; we assume by default either that documentary tracks are good proxies of the assessments that they are grounding, or that documentary tracks structured to be compared allow finding alternative ways to the same assessment or alternative assessment to be found by explainable assessor's shifts.

We first designed and experimented with this conceptual framework using queries from research data logs in [35] that deliver a full description and first testing of the modeling [29]. Figure 7 reproduces the main representation of this conceptual framework of information representation in GRAPHYP, as a graph modeling from logs, on which we have performed extensive tests reported in the cited paper [35]. Behaviors of N users (parameter 1: Intensity) of K scientific materials (Parameter 2: Variety) are altogether measured,

combined by the geometry of the graph with a third parameter (Parameter 3: Attention) that we measure by differences between the observed value and the mean value of the α/β ratio (N/K) on the whole sample.
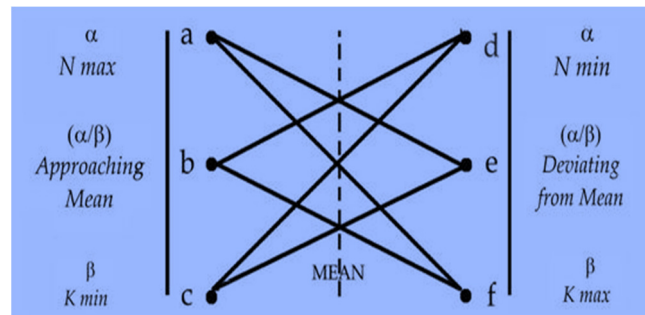

Figure 7. Conceptual framework of GRAPHYP. Mapping search experience in GRAPHYP: a–f, as the six typical modellable cliques combining triplets of nodes, included in a cognitive community [35].

This tested measure of "possible" results is the first of two subsystems, the second being, with same structure and data, the assessors' shifts that could allow a user to circulate towards the "best possible" choice, according to their own personal set of hypotheses. One could recognize in this kind of "inverted reasoning" reflections of the above-mentioned Leibnizian reasoning on the "best possible" choice. We thus call this KG a "Multiverse Graph" as the first application to graphs of the popular accurate multiverse analysis in machine learning [38]: an explainable geometric Knowledge Graph's expressive power gives a decisive additional extension to the already significant applications of multiverse analysis (see discussion, research on graphs and decision-making technologies). Figure 8 shows the corresponding steps of the workflow.
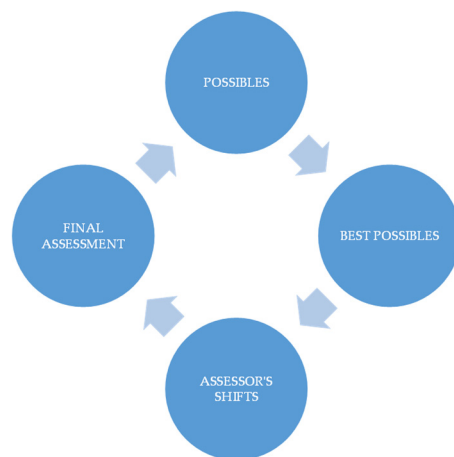

Figure 8. Steps in DISCRIMINATOR GR2 between assessor's shift: Choice of the best possible data track.

(i) GRAPHYP fits with ELK

The conditions for ELK compatibility are met. As seen above, as an analytics module with strategic capabilities, GRAPHYP is a KG that offers improvements in data quality and visualization to make informed decisions in organizations. These tasks could work with ELK operational requirements. ELK is highly configurable, allowing metrics to be tailored to needs in a superb global warehouse that acts as a scalable, searchable database. Files read by Beats applications are reached and collected by Logtash, which can also reach various multimodal sources. Logtash can also filter data from multiple systems and collect it in one location.

Kibana provides an easy-to-use interface for collecting, analyzing, and converting data into charts that allow patterns in the data to be visualized and interpreted. With third-party

applications such as Apache, Kafka can read multiple data sources simultaneously as a streaming platform for real-time distribution. With the ELK–Hadoop Connector, Hadoop operates as a massive batch data storage system that could interact as an in-system, bi-directional data discovery and visualization platform. New ideas could come also from the community of assembly information representation [39].

(ii)   Cooperation: Mutual benefits

ELK is known as an open-source solution for companies that handle Big Data scenarios quickly and conveniently [40]. There is less observation that ELK could accommodate scientific knowledge tasks such as EZ Paarse-EZ Measure tasks, where Kibana offers many users in libraries and laboratories real-time assessments of scientific documentation (https://www.inist.fr/services/acceder/ezmesure/ (accessed on 9 April 2023)) analysis at a very large scale with several visualized functions for the sophisticated management of scientific documents. This example also shows that in this context, a wide range of services, both analytical and administrative, are switched to use logins that enable the ELK.

(iii)   COUNTER: A potential testing case of Elastic Stack and GRAPHYP interactions in strategic management

An example at the global level is the COUNTER system (https://www.projectcounter.org/code-of-practice/counter-release-5-faqs/ (accessed on 22 May 2023)), which is the first wide-ranging use case of web usage within the ELK framework and overall a very operational case of strategic management in a decentralized context, and at least a potential case test for GRAPHYP functionalities. COUNTER manages the resources of academic libraries worldwide. The input of Project COUNTER consists of log files of decisions in academic libraries, just like in EZ Paarse, and from web usage data, a large amount of strategic data are processed, which can be consulted at the link mentioned earlier, all of which could be coupled with GRAPHYP's triplet analytical parameter. We have started to prepare tests for this case, which are well organized and allow observing a wide range of strategic situations that could be translated into a corresponding situation of sophisticated information management.

(iv)   A general workflow of cooperation between ELK and GR

In addition, many tracks are opened to joint approaches in the current leading topics of research and service development at the crossroads of ELK and GRAPHYP technologies, such as the context-aware representation of digital twins' data [41]. SKG GRAPHYP could offer numerous sandboxes and blackboards in addition to existing uses of ELK in order to explore the numerous interfaces between documentary tracks and ways to discovery by analyzing the possible alternatives in scientific documentation [42].

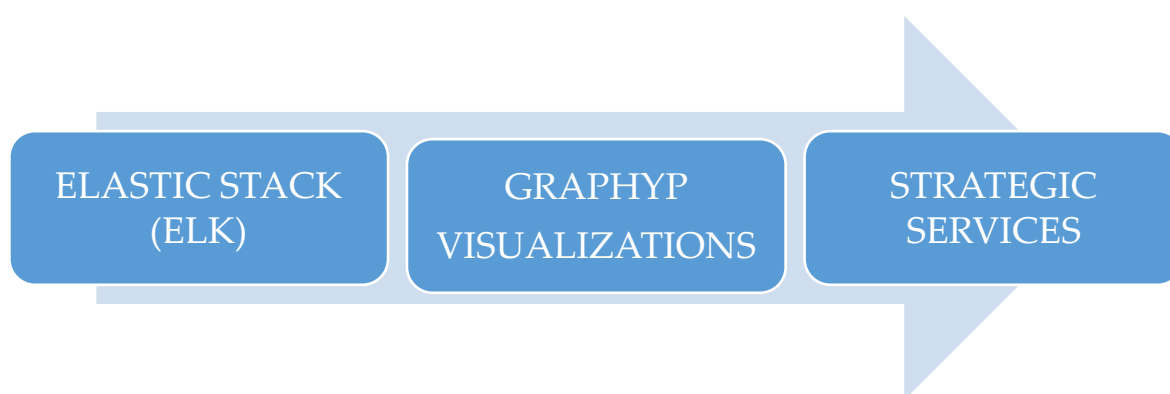Figure 9 shows a general workflow of possible future cooperation.

Figure 9. Workflow of ELK and GRAPHYP.

This example suggests that connections and experiments could be made with representations that SKG GRAPHYP could offer in addition to existing uses of ELK in order to

explore the numerous interfaces between documentary tracks and ways to discovery by analyzing the possible alternatives in scientific documentation [38]. The field of geometric interpretable graphs to which GRAPHYP belongs could also be a way to find synergies between analysis and interpretation, capturing the benefits of the new, expressive geometry of KG [43].

A useful example of future cooperation of ELK with the GRAPHYP Knowledge Graph could be derived from the background of the above-mentioned EZ Paarse-EZ Measure application of ELK.

This application would install cooperation between ELK and GRAPHYP in line with the workflow shown in Figure 8. The EZ Paarse-EZ Measure is a massive, highly skilled service to provide documentary strategies to suit the needs of a large national community of researchers, with multiple homemade original applications, including, as an example, an innovative simple localization service for the real-time analysis and mapping of downloads (https://bibliomap.inist.fr/ (accessed on 9 April 2023)).

Daily figures of management for 1 February 2023 are as follows: individual sources of data: 337 platforms of 134 entities of research and higher education in France; volume of data: 3953 days of logs. Additionally, as could be seen from the above-mentioned sites, it is a strong community with agile practices and community cooperation on analytic tasks. Service development for users has been an outstanding success in public as well as private deliveries (see blogs of the sites and related services) and allows analyzing logs of highly differentiated sources: logs come from 10 254 983 distinct titles of scientific journals and e-books. GRAPHYP could take charge of the above-described complementary services from downloads of logs with the following:

- Typologies of visualized downloads according to GR1 possible choices as compared to a user's search experiences (see above GR1 requirements in GRAPHYP);
- Representation of an assessor's shifts with GR2 methodology in an additional service of assistance to visualize the comparative search experience of the users;
- Assistance in the management of research at various scales with the visualization of documentary strategies that could compare the results of search experiences, as represented in GR1 and GR2, with scenarios of evolutions of documentation according to funding or research priorities.

5. Discussion

The following general observations pertain to our three research questions.

Organizations are faced with the challenge of converting historically grown processes, IT systems, and data structures into the logic of modern artificial intelligence technologies. In addition to the formatting of data and the resolution of redundancies, the data quality is often in the foreground, because not all systems have the same data status of the same quality. However, the advantages of artificial intelligence in business processes often only arise when it is used in real time.

The increasing demand for KG and ELK is driven by three trends. Organizations continue to invest in automation as part of the digital transformation and increasingly rely on artificial intelligence and machine learning. In order to flexibly meet the dynamic business requirements, it is also necessary to provide highly scalable applications that grow with current developments. Finally, the increasing complexity of data requires powerful solutions to easily and quickly navigate, analyze, and understand highly interconnected data.

In this paper, we examined how systems that process data in different steps of strategic decision making could improve their semantic interoperability. We assumed that semantic interoperability requires a common methodology of text analysis and benchmarking, as well as actions based on web intelligence. We found that there are a variety of different methods and techniques, and that search log analysis is mentioned as one of them [44]. We chose the option to explore a "win–win" solution of semantic interoperability that we propose to explore the overall functional complementarity of ELK and KG GRAPHYP. The discussion here follows the points of our three research questions.

RQ1 prompted us to explore the benefits and limitations of ELK as a suitable solution for decision makers in analyzing and visualizing data. We unequivocally concluded that among the wide range of database management system solutions, ELK deserves positive recognition for its internally clear and efficient division of tasks and functionalities, bundling benefits of scalability and flexibility in all directions where semantic interoperability is required.

RQ2 was about KG GRAPHYP's ability to provide data analytics that could benefit from ELK and enhance strategic decision-making services.

KG and the ELK solutions allow disparate data sources to be quickly connected by simply storing them. As for the applicability of the Elastic solution, Logstash can handle the querying and structuring of many data types. Elasticsearch enables the storage of large amounts of data and can therefore guarantee failures. Many general visualizations are available for the visualizations. With the support of a new generation of explainable KGs such as GRAPHYP, we have observed that organizations can find connections between disparate information and relationships between ideas and concepts that may not be related or explicitly stated. When organizations have information that needs to flow back and forth between multiple systems while preserving the relationship between those data entities, KGs serve as a good way to represent and normalize data. Finding and editing charts is a quick and easy way to visualize how information relates to each other.

We have also observed that KG research has made significant advances in the geometry of expressively explainable KGs by applying new methods of causal inference to transfer learning or multiverse analysis. Progress is made on new frontiers for KG management such as the optimization of processing heterogeneous data at any scale [45]. Workflow execution is also a hot topic for both, and ELK is studied in environments of exascale computing and for training data in automated analysis [46]. The logics and visualization of KG are areas of discovery that could appear as new frontiers [46].

It turned out that today, KG GRAPHYP can be seen as a new extension for ELK to successfully meet the difficult challenge of inspecting and managing data and its contents with the accuracy that the analysis of log files provides for strategic decision making. KG helps to discover, understand, and research connections in data. With KG, users can find answers to complex questions faster and more accurately. Conceivable areas of application for this include behavior analysis, fraud detection, cyber security, personalization, recommendations, etc. KG automatically identifies the most important relationships and can use the relevance ranking to extract the most important information from a large amount of data, and GRAPHYP operates with an explainable geometry. As shown in Figure 10, KG integrates with Elasticsearch and Kibana as an extension and scales easily, providing near real-time data for analysis. Another big advantage of the KG is the high level of comprehensibility of the information generated. Not only data scientists and data engineers, but also people from other disciplines can read and interpret a graph.

Using KG with ELK has a number of advantages. This includes the possibility of merging different KGs at the same time and seeing data from different perspectives in the overall view without losing the crucial focus, because KGs are able to aggregate new types of information from different sources. For example, interests, needs, and products can be personalized even better in practice.

In our RQ3, we wanted to explore how the future benefits of a merger of ELK and GR could be presented. In this last research question, we tested the integration of ELK and GR and discussed the benefits that could result for the user.

From all those standpoints, plus another one, we conclude that GRAPHYP is a win–win partner of ELK. As we highlighted in Section 4, GRAPHYP's value added to another KG is not only to propose an additional qualitative information process in ELK but also to provide a decisive strategic service to users: GRAPHYP is the first available KG that proposes a comprehensive representation of all possible assessor shifts in an interpretable modeling of possible answers to a query, with a human-in-the-loop system. To the best of our knowledge, this last detection function is currently represented exclusively in GRAPHYP

modeling: as we show in Section 4, it detects, discriminates, and analyzes the whole world of possible answers, because of its unique, well-tested, explainable subgraph geometry.
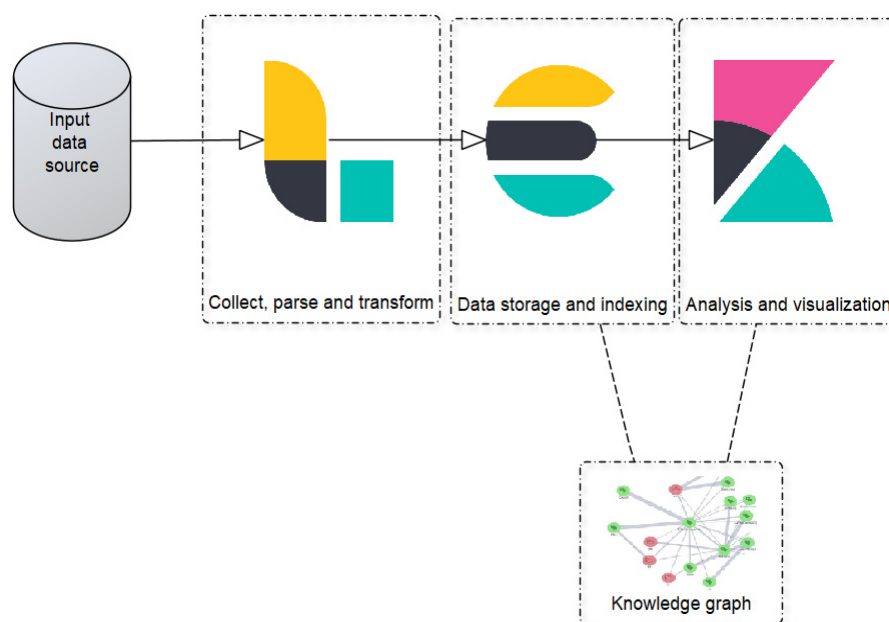

Figure 10. ELK and KG integration.

GRAPHYP belongs to the new, completely innovative category of Neural Graph Databases (NGDBs), which dramatically changes the interplay between strategic data discovery and reuse [47]. From the comparison in [40], we could see that queries in symbolic databases are optimized into a plan executed from indexes of storage DB, while in NGDB, the query is executed in a latent area of the underlying database graph. This is the case with the geometric hypergraph GRAPHYP, which follows this new possibilistic approach [48]. This conceptual framework opens wide avenues to a win–win partnership between ELK and GRAPHYP: we can find an overview of further developments with the recently proposed taxonomy of neural approaches for complex logical query answering. It shows, in the framework of NGBD, the dramatic extension of induced interactions between queries, modeling, and graphs (https://github.com/neuralgraphdatabases/awesome-logical-query (accessed on 9 April 2023)).

6. Conclusions

Big Data is so complex because it is based on two main factors: unstructured and connected data. In particular, the connections between the various elements, for example, between product and customer feedback or data from development and production, pose a challenge in the analysis for decision making. It is important to make the best possible data basis with high data quality, regardless of whether it is at a business or technical level. Data discovery is difficult due to unknown data sources, poor data quality, data silos, and compliance restrictions. These issues can be due to data being used or generated by a specific application and stored on an isolated data platform. Logging is one of many tools that can be used to identify and analyze errors. In times of microservices and distributed systems, tools such as ELK are becoming more and more important. With ELK, data can be processed quickly, no matter how many data points there are and from what sources they come. Many work processes can be automated with ELK and the data can also be continuously maintained, improved, and enriched. The combination of ELK and KGs such as GRAPHYP provides a powerful and customizable way to visualize the data. These dashboards can be used to quickly gain and share data insights. Filters can also be applied and potential anomalies examined. With Kibana, users can customize

dashboards and create appropriate visualizations that meet existing security, auditing, and compliance requirements.

A hybrid approach combining the capabilities of ELK and GRAPHYP is more than just a new decision-making solution that can be applied to different systems: its innovative results go further in the methodology of strategic thinking and end-user services. Additional research and testing need to be developed to identify the numerous services that could be derived from this win–win technology partnership.

Author Contributions: Conceptualization, O.A., R.F. and U.S.; methodology, O.A. and R.F.; software, O.A. and R.F.; validation, O.A. and R.F.; formal analysis, O.A. and R.F.; investigation, O.A. and R.F.; resources, O.A. and R.F.; data curation, O.A. and R.F.; writing—original draft preparation, O.A. and R.F.; writing—review and editing, O.A., R.F., U.S. and R.Q.; supervision, O.A., R.F. and U.S.; project administration, O.A., R.F. and U.S.; funding acquisition, O.A., R.F., U.S. and R.Q. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable; this study does not report any data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kalajdjieski, J.; Raikwar, M.; Arsov, N.; Velinov, G.; Gligoroski, D. Databases fit for blockchain technology: A complete overview. *Blockchain Res. Appl.* **2022**, *4*, 100116. [CrossRef]
2. Liu, R.; Fu, R.; Xu, K.; Shi, X.; Ren, X. A Review of Knowledge Graph-Based Reasoning Technology in the Operation of Power Systems. *Appl. Sci.* **2023**, *13*, 4357. [CrossRef]
3. Tian, L.; Zhou, X.; Wu, Y.-P.; Zhou, W.-T.; Zhang, J.-H.; Zhang, T.-S. Knowledge graph and knowledge reasoning: A systematic review. *J. Electron. Sci. Technol.* **2022**, *20*, 100159. [CrossRef]
4. Fabre, R.; Azeroual, O.; Schöpfel, J.; Bellot, P.; Egret, D. A Multiverse Graph to Help Scientific Reasoning from Web Usage: Interpretable Patterns of Assessor Shifts in GRAPHYP. *Future Internet* **2023**, *15*, 147. [CrossRef]
5. Rejeb, A.; Keogh, J.G.; Martindale, W.; Dooley, D.; Smart, E.; Simske, S.; Wamba, S.F.; Breslin, J.G.; Bandara, K.Y.; Thakur, S.; et al. Charting Past, Present, and Future Research in the Semantic Web and Interoperability. *Future Internet* **2022**, *14*, 161. [CrossRef]
6. Jaradeh, M.Y.; Oelen, A.; Farfar, K.E.; Prinz, M.; D'Souza, J.; Kismihók, G.; Stocker, M.; Auer, S. Open Research Knowledge Graph: Next Generation Infrastructure for Semantic Scholarly Knowledge. In Proceedings of the 10th International Conference on Knowledge Capture (K-CAP'19), Marina Del Rey, CA, USA, 19–21 November 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 243–246. [CrossRef]
7. Zeb, S.; Mahmood, A.; Hassan, S.A.; Piran, J.; Gidlund, M.; Guizani, M. Industrial digital twins at the nexus of NextG wireless networks and computational intelligence: A survey. *J. Netw. Comput. Appl.* **2022**, *200*, 103309. [CrossRef]
8. Azeroual, O.; Fabre, R. Processing Big Data with Apache Hadoop in the Current Challenging Era of COVID-19. *Big Data Cogn. Comput.* **2021**, *5*, 12. [CrossRef]
9. Sharma, I.; Tiwari, R.; Anand, A. Open Source Big Data Analytics Technique. In Proceedings of the International Conference on Data Engineering and Communication Technology, Pune, India, 15–16 December 2017; Satapathy, S., Bhateja, V., Joshi, A., Eds.; Advances in Intelligent Systems and Computing. Springer: Singapore, 2017; Volume 468. [CrossRef]
10. Thijs, B. Science Mapping and the Identification of Topics: Theoretical and Methodological Considerations. In *Springer Handbook of Science and Technology Indicators*; Glänzel, W., Moed, H.F., Schmoch, U., Thelwall, M., Eds.; Springer Handbooks; Springer: Cham, Switzerland, 2019. [CrossRef]
11. Ghosh, S.; Rath, M.; Shah, C. Searching as Learning: Exploring Search Behavior and Learning Outcomes in Learning-related Tasks. In Proceedings of the 2018 Conference on Human Information Interaction & Retrieval (CHIIR '18), New Brunswick, NJ, USA, 11–15 March 2018; Association for Computing Machinery: New York, NY, USA, 2018; pp. 22–31. [CrossRef]
12. Sharifpour, R.; Wu, M.; Zhang, X. Large-scale analysis of query logs to profile users for dataset search. *J. Doc.* **2022**, *79*, 66–85. [CrossRef]
13. Sanderson, M.; Scholer, F.; Turpin, A. Relatively relevant: Assessor shift in document judgements. In Proceedings of the Australasian Document Computing Symposium, Parramatta, Australia, 8–9 December 2010; RMIT Press: Melbourne, Australia, 2010; pp. 60–67.
14. Eke, C.I.; Norman, A.A.; Shuib, L.; Nweke, H.F. A Survey of User Profiling: State-of-the-Art, Challenges, and Solutions. *IEEE Access* **2019**, *7*, 144907–144924. [CrossRef]
15. Özsahin, T. Evaluation of the Visualization Grammar Vega as a Base for Visualizing and Managing Logstash Pipelines Based on the ELK-Stack. Ph.D. Thesis, Technische Hochschule Ingolstadt, Ingolstadt, Germany, 2020.
16. Turnbull, J. *The Logstash Book*; Turnbull Press: Brooklyn, NY, USA, 2013.

17. Durumeric, Z.; Adrian, D.; Mirian, A.; Bailey, M.; Halderman, J.A. A Search Engine Backed by Internet-Wide Scanning. In Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security (CCS '15), Denver, CO, USA, 12–16 October 2015; Association for Computing Machinery: New York, NY, USA, 2015; pp. 542–553. [CrossRef]
18. Liu, Z.H.; Hammerschmidt, B.; McMahon, D.; Chang, H.; Lu, Y.; Spiegel, J.; Sosa, A.C.; Suresh, S.; Arora, G.; Arora, V.; et al. Native JSON datatype support. *Proc. VLDB Endow.* **2020**, *13*, 3059–3071. [CrossRef]
19. Dhulavvagol, P.M.; Bhajantri, V.H.; Totad, S.G. Performance Analysis of Distributed Processing System using Shard Selection Techniques on Elasticsearch. *Procedia Comput. Sci.* **2020**, *167*, 1626–1635. [CrossRef]
20. Hirai, J.; Raghavan, S.; Garcia-Molina, H.; Paepcke, A. WebBase: A repository of Web pages. *Comput. Netw.* **2000**, *33*, 277–293. [CrossRef]
21. Kim, Y.; Callan, J.; Culpepper, J.S.; Moffat, A. Efficient distributed selective search. *Inf. Retr. J.* **2016**, *20*, 221–252. [CrossRef]
22. Gormley, C.; Tong, Z. *Elasticsearch: The Definitive Guide: A Distributed Real-Time Search and Analytics Engine*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2015.
23. Butakov, N.; Petrov, M.; Mukhina, K.; Nasonov, D.; Kovalchuk, S. Unified domain-specific language for collecting and processing data of social media. *J. Intell. Inf. Syst.* **2018**, *51*, 389–414. [CrossRef]
24. Kononenko, O.; Baysal, O.; Holmes, R.; Godfrey, M.W. Mining modern repositories with elasticsearch. In Proceedings of the 11th Working Conference on Mining Software Repositories, Hyderabad, India, 31 May–1 June 2014; pp. 328–331. [CrossRef]
25. Andhavarapu, A. *Learning Elasticsearch*; Packt Publishing Ltd.: Birmingham, UK, 2017.
26. Chhajed, S. *Learning ELK Stack*; Packt Publishing Ltd.: Birmingham, UK, 2015.
27. Zamfir, V.A.; Carabas, M.; Carabas, C.; Tapus, N. Systems monitoring and big data analysis using the elasticsearch system. In Proceedings of the 2019 22nd International Conference on Control Systems and Computer Science (CSCS), Bucharest, Romania, 28–30 May 2019; pp. 188–193. [CrossRef]
28. Bajer, M. Building an IoT data hub with elasticsearch, Logstash and Kibana. In Proceedings of the 2017 5th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW), Prague, Czech Republic, 21–23 August 2017; pp. 63–68. [CrossRef]
29. Atz, K.; Grisoni, F.; Schneider, G. Geometric deep learning on molecular representations. *Nat. Mach. Intell.* **2021**, *3*, 1023–1032. [CrossRef]
30. Bronstein, M.M.; Bruna, J.; Cohen, T.; Veličković, P. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv* **2021**, arXiv:2104.13478.
31. Santoro, A.; Raposo, D.; Barrett, D.G.; Malinowski, M.; Pascanu, R.; Battaglia, P.; Lillicrap, T. A simple neural network module for relational reasoning. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17), Long Beach, CA, USA, 4–9 December 2017; Curran Associates Inc.: Red Hook, NY, USA; pp. 4974–4983.
32. Satorras, V.G.; Hoogeboom, E.; Welling, M. E (n) equivariant graph neural networks. In Proceedings of the International Conference on Machine Learning, Virtual, 18–24 July 2021; PMLR: Cambridge, MA, USA, 2021; pp. 9323–9332.
33. Linardatos, P.; Papastefanopoulos, V.; Kotsiantis, S. Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy* **2021**, *23*, 18. [CrossRef]
34. Auer, S.; Oelen, A.; Haris, M.; Stocker, M.; D'souza, J.; Farfar, K.E.; Vogt, L.; Prinz, M.; Wiens, V.; Jaradeh, M.Y. Improving Access to Scientific Literature with Knowledge Graphs. *Bibl. Forsch. Prax.* **2020**, *44*, 516–529. [CrossRef]
35. Fabre, R.; Azeroual, O.; Bellot, P.; Schöpfel, J.; Egret, D. Retrieving Adversarial Cliques in Cognitive Communities: A New Conceptual Framework for Scientific Knowledge Graphs. *Future Internet* **2022**, *14*, 262. [CrossRef]
36. Latta, R. *A Critical Exposition of the Philosophy of Leibniz, with an Appendix of Leading Passages*; Oxford University Press: Oxford, UK, 1901; Available online: https://archive.org/details/cu31924052172271 (accessed on 6 April 2023).
37. Liu, Y.; Kale, A.; Althoff, T.; Heer, J. Boba: Authoring and visualizing multiverse analyses. *IEEE Trans. Vis. Comput. Graph.* **2020**, *27*, 1753–1763. [CrossRef]
38. Bell, S.J.; Kampman, O.P.; Dodge, J.; Lawrence, N.D. Modeling the Machine Learning Multiverse. *arXiv* **2022**, arXiv:2206.05985. [CrossRef]
39. Minango, N.R.; Maffei, A. Beyond assembly features: Systematic review of the core concepts and perspectives towards a unified approach to assembly information representation. *Res. Eng. Des.* **2022**, *34*, 3–38. [CrossRef]
40. Schüssler, J.; Karbstein, D.; Klein, D.; Zimmermann, A. Visualizing information for enterprise architecture design decisions using elastic stack. In Proceedings of the BIR-WS 2018: BIR Short Papers, Workshops and Doctoral Consortium Joint Proceedings, Co-Located with 17th International Conference Perspectives in Business Informatics Research (BIR 2018), Stockholm, Sweden, 24–26 September 2018; Volume 2218, pp. 1–10.
41. Rico, M.; Taverna, M.L.; Galli, M.R.; Caliusco, M.L. Context-aware representation of digital twins' data: The ontology network role. *Comput. Ind.* **2023**, *146*, 103856. [CrossRef]
42. Fabre, R.; Schöpfel, J. L'hypertexte et les sciences (1991–2021): Des voies navigables pour les routes de connaissances. Histoire de la recherche contemporaine. *La Revue Comité pour L'histoire CNRS* **2021**, *10*. [CrossRef]
43. Veličković, P. Message passing all the way up. *arXiv* **2022**, arXiv:2202.11097. [CrossRef]
44. Martinez-Gil, J. An overview of textual semantic similarity measures based on web intelligence. *Artif. Intell. Rev.* **2012**, *42*, 935–943. [CrossRef]

45.   Yu, C.; Wang, F.; Liu, Y.-H.; An, L. Research on knowledge graph alignment model based on deep learning. *Expert Syst. Appl.*
       2021, *186*, 115768. [CrossRef]
46.   Papadimitriou, G.; Wang, C.; Vahi, K.; da Silva, R.F.; Mandal, A.; Liu, Z.; Mayani, R.; Rynge, M.; Kiran, M.; Lynch, V.E.; et al.
       End-to-end online performance data capture and analysis for scientific workflows. *Future Gener. Comput. Syst.* **2020**, *117*, 387–400.
       [CrossRef]
47.   Ren, H.; Galkin, M.; Cochez, M.; Zhu, Z.; Leskovec, J. Neural Graph Reasoning: Complex Logical Query Answering Meets Graph
       Databases. *arXiv* **2023**, arXiv:2303.14617.
48.   Gao, C.; Wang, Y.; Zhou, J.; Ding, W.; Shen, L.; Lai, Z. Possibilistic Neighborhood Graph: A New Concept of Similarity Graph
       Learning. In *IEEE Transactions on Emerging Topics in Computational Intelligence*; IEEE: Piscataway, NJ, USA, 2022. [CrossRef]