Future Generation Computer Systems 163 (2025) 107501

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

A multi-level IIOT platform for boosting mines digitalization

Raúl Miñón [a,b,*], Juan López-de-Armentia [a], Lander Bonilla [a], Aitor Brazaola [c], Ibai Laña [c], M. Carmen Palacios [c], Szymon Mueller [d], Michal Blaszczak [d], Herwig Zeiner [e], Julia Tschuden [e], Mohammad Yusuf Quadri [e], Ignasi Garcia-Milà [f], Andrea Bartoli [f], Norbert Gormolla [g], Alberto Fernández [h], Pablo Segarra [h], José A. Sanchidrián [h], Philipp Hartlieb [i]

[a] Digital, TECNALIA, Basque Research and Technology Alliance (BRTA), Albert Einstein, 28, 01510, Vitoria, Spain
[b] Facultad de Ingeniería, University of Deusto, Avenida de las Universidades 48007, 48007, Bilbao, Spain
[c] Digital, TECNALIA, Basque Research and Technology Alliance (BRTA), Astondo Bidea, Building 700, Derio, 48160, Bilbao, Spain
[d] Poznan Supercomputing and Networking Centre, Noskowskiego 12/14, 61-704, Poznań, Poland
[e] Joanneum Research, Forschungsgesellschaft mbH, Leonhardstrasse 59, 8010, Graz, Austria
[f] WorldSensing SL, Viriat st, 47 10ª plant, 08014, Barcelona, Spain
[g] DMT GmbH & Co. KG, Am TÜV 1, 45307, Essen, Germany
[h] Universidad Politécnica de Madrid – ETSI Minas y Energía, Ríos Rosas 21, 28003, Madrid, Spain
[i] Mining Engineering and Mineral Economics, Montanuniversitaet, Erzherzog-Johann-Straße 3, 8700, Leoben, Austria

ARTICLE INFO

Keywords:
Mine digitalization
Big data
Edge
Fog
Cloud continuum

ABSTRACT

This paper presents an innovative IIoT multi-level platform tailored to address the specific needs of the mining domain. The platform has been conceptualized and built in the context of the illuMINEation European project. For this purpose, mining specific use cases have been designed such as promoting underground safe areas, performing efficient mining operations or boosting predictive maintenance approaches. Then, specific requirements have been identified and, as a result, the platform has been developed. It consists of four-level layered platform: (1) edge devices layer to manage several sensors deployed in the mines; (2) edge box layer to provide in-mine operations such as filtering, streaming and processing; (3) fog layer which offers an overall perspective of each mine; and (4) cloud layer to centralize the data of all the mines and to provide powerful processing capabilities. In addition, the platform is robustly secured in terms of protecting communications confidentiality and access control and also provides a toolbox aimed at manipulating 3D complex images to obtain operable mine-domain novel user interfaces. Finally, a platform validation is proposed where three different use cases are explained to better show and demonstrate the capabilities of the platform.

1. Introduction

1.1. Motivation context

Digitalization [1] has been recognized as a major trend with the potential to reshape societal and business norms in the short and medium-term future. It implies a conversion from the analog format to the digital which offer numerous advantages. Among them, it enables quick access to information, efficient storage, automation of processes, enhanced security mechanisms or efficient operations. In addition, it is worth mentioning that the analysis of digital data provides valuable insights, aiding informed decision-making and identifying patterns.

In this direction, the mining industry is facing several challenges in order to improve the productivity and be more socially and environmentally aware. Sanchez and Hartlieb [2] highlight key technologies like robotics, smart sensors, artificial intelligence (AI) analytics or integrated remote operation centres to drive the change over the diverse mine aspects.

Consequently, a set of innovative European projects have been carried out in this area of mining digitalization. For instance, *Sustainable Intelligent Mining Systems for the Global Mining Industry (SIMS)* [3],

* Corresponding author.
E-mail addresses: raul.minon@tecnalia.com, raul.minon@deusto.es (R. Miñón), juan.lopezdearmentia@tecnalia.com (J. López-de-Armentia), lander.bonilla@tecnalia.com (L. Bonilla), aitor.brazaola@tecnalia.com (A. Brazaola), ibai.lana@tecnalia.com (I. Laña), maricarmen.palacios@tecnalia.com (M.C. Palacios), mael@man.poznan.pl (S. Mueller), mblaszczak@man.poznan.pl (M. Blaszczak), herwig.zeiner@joanneum.at (H. Zeiner), julia.tschuden@joanneum.at (J. Tschuden), yusuf.quadri@joanneum.at (M.Y. Quadri), igarciamila@worldsensing.com (I. Garcia-Milà), abartoli@worldsensing.com (A. Bartoli), norbert.gomolla@dmt-group.com (N. Gormolla), alberto.fernandez@upm.es (A. Fernández), pablo.segarra@upm.es (P. Segarra), ja.sanchidrian@upm.es (J.A. Sanchidrián), philipp.hartlieb@unileoben.ac.at (P. Hartlieb).

https://doi.org/10.1016/j.future.2024.107501
Received 26 January 2024; Received in revised form 17 June 2024; Accepted 25 August 2024
Available online 27 August 2024

produced results in robotics, automation and 5G connectivity underground. It also addressed the areas of fully electrical underground mining machines, virtual reality for training, aerial underground mine inspection, application of intelligent rock bolts and finally, the underground navigation system onboard. The project *Real-time optimization of extraction and the logistic process in highly complex geological and selective mining settings (Real-Time-Mining)* [4] was aimed at developing a real-time process feedback control loop linking online data acquired during extraction at the mining with an sequentially updatable resource model associated with real-time optimization of long-term planning, short-term sequencing and production control decisions. *Sustainable Low Impact Mining solution for exploitation of small mineral deposits based on advanced rock blasting and environmental technologies (SLIM)* [5] aimed at non-linear rock mass fragmentation by blasting models, airborne particulate matter, vibration affections and nitrate leaching mitigation actions for exploitation of small mineral deposits through a new generation of explosives and an advanced automatic blast design software based on improved rock mass characterization and fragmentation models for optimum fragmentation and minimum rock damage and far-field vibrations. *Sonic Drilling coupled with Automated Mineralogy and chemistry On-Line-On-Mine-Real-Time (SOLSA)* [6] deals with (1) integrated drilling optimized to operate in the difficult lateritic environment with the challenge of a mixture of hard and soft rocks -extensible also to other ore types-, (2) fully automated scanner and phase identification analyser and software -usable as well in other sectors. SOLSA combines, for the first time, the non-destructive sensors X-ray fluorescence, X-ray diffraction, vibrational spectroscopies, 3D and hyperspectral imaging along the drill core.

1.2. Research problem

However, despite the significant advances achieved in the diverse mining digitalization projects, additional research must be accomplished in these area. For instance, those projects lack of a unified mining-specific data management platform which in this level of technological maturity becomes of capital importance. It is clearly noteworthy the necessity of a data platform capable of managing the data generated by sensors and exposed by other sources, in addition to provide the ability to easily execute AI algorithms boosting low-latency in-mine decisions. In this way, aspects like the mine security, the predictive maintenance or the sustainable resource extraction could potentially be enhanced.

In this context, the European H2020 illuMINEation project - Bright concepts for a safe and sustainable digital mining future - is being developed. This project provides a Multi-level reliable distributed IIoT platform and Innovative user interfaces tailored specifically for the mining domain in order to mainly satisfy the following objectives:

- Boosting secure mining operations with the concept of secure areas. To this end, rock stability is assessed, the environment is monitored to identify possible issues in the air quality, workers are geo-localized and drones can be used as first-response units.
- Advanced predictive maintenance and efficient operations: drones as agile inspectors, supervision of the state of the batteries of the diverse machines, additive manufacturing of spare parts and telematics data analysis of the equipment.
- Sustainable and intelligent mining resource extraction. This includes advanced drilling techniques, measurement while drilling and optimization in rock fragmentation and resource management.

illuMINEation platform is based on three principal key enabler technologies: (1) Wireless sensor networks (WSN) as a way to monitor large deployments by measuring many points on a single environment. In turn, three specific technological advances enabled this concept: micro-controllers, wireless communications and sensors [7]. (2) Big Data technologies to deal with the data life-cycle [8] of the vast amount of data coming from the diverse mines as well as the provided by third-party entities. (3) Cloud architectures [9] involving the management of the underlying layers: hardware, infrastructure, platform and application.

1.3. Contributions

The contribution of this paper is focused on explaining this innovative IIoT platform proposed specifically for the mining domain.

Concretely, the building of a novel Industrial Internet of Things (IIoT) platform that connects the physical mining world with a robust multi-level IIoT platform. The data of high-quality and low-cost sensors installed throughout the mines is wirelessly distributed from an edge layer into the mines to fog and cloud computing layers. In addition, with the implementation of advanced algorithms as services a powerful, sophisticated management is implemented, thus ensuring safer and more efficient mining operations. Moreover, novel User Interfaces and dashboards are utilized for operational control and support of a virtual mining environment.

1.4. Structure

The rest of the paper is structured as follows. Section 2 describes the related work. The multi-level IIoT platform for the mining domain is detailed in Section 3. Section 4 presents the platform validation to demonstrate its capacities. Finally, conclusions and future work are respectively presented in Section 5 and Section 6.

2. Related work

In this section, several architectures and platforms related to IIoT platform implemented in the illuMINEation project are presented. In addition, Table 1 shows a summary of the key features taken into consideration for this analysis. Concretely, the characteristics considered for the comparison are the following:

- Diverse sensors: it can ingest data from diverse sensors.
- Several layers: it has the ability to transfer sensors' data from the mines to a centralized cloud system, passing through fog intermediate layers and following the cloud continuum [10] approach.
- Streaming: the platform supports sending and processing data in streaming mode.
- Open-source solution: it is an open-source solution
- License required: a license is mandatory to use the platform.
- Integration of 3rd party modules: it can integrate external analytics models, data transformations and applications from stakeholders.
- Data import/export: the platform allows exporting and importing data from/to third-party systems.
- Data discovery: it allows managing and discovering datasets and metadata associated to them.
- Mine-domain interfaces: it provides an advanced user interfaces for the mining domain.
- Holistic security: it integrates robust security mechanisms.

There are various platforms that address some of these characteristics. FIWARE [11] is an open-source cloud platform with a collaborative and mature ecosystem of developers, innovation Hubs and accelerators. FIWARE Foundation became the main European provider of reference architecture aligned to the idea of multi-layer communication. It was designed to support complex distributed systems, where multiple data sources, like sensors, actuators or external services are processed, integrated and managed by dedicated applications. However, it is not adapted to the mine domain, so it does not have advanced interfaces for visualization of the mines themselves. On the other hand, it does not have a dataset discovery layer, which difficult to know the

Table 1
Related work compared against the main functionalities of the illuMINEation platform (Yes: ✓; No: ✗; Partial: ∼; Not known: -)

	Diverse sensors	Several layers	Streaming	Open Source	License required	3rdparty modules	Data import export	Data discovery	Mine domain interfaces	Holistic security	Implemented
FIWARE [11]	✓	✓	✓	✓	✓	✗	✗	✗	✗	∼	✓
CDH [12]	✓	✓	✓	✓	✓	✓	✓	∼	✗	✓	✓
Pivotal Data [13]	✓	✓	✓	∼	✓	✓	✓	∼	✗	∼	✓
Azure HDInsight [14]	✓	✓	✓	✓	✓	∼	✓	✓	✗	✓	✓
SAP HANA [15]	✓	✓	✓	✗	✓	∼	✓	✓	✗	✓	✓
Databricks [16]	✓	✓	✓	✗	✓	✓	✓	✓	✗	✓	✓
Elastic Cloud [17]	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓
IoMT [18]	∼	✓	✓	–	✗	–	–	✗	✗	✗	∼
Cao et al. [19]	✓	✓	✓	–	✗	✓	∼	✗	✓	∼	✓
Certiq [20]	∼	–	✓	✗	✓	✗	–	✗	✗	✓	✓
OptiMine [21]	✓	–	✓	✗	✓	–	–	–	✓	✓	✓
Sustainable mining IoT [22]	✓	✗	✓	–	–	–	✗	✗	✓	✗	✗
MIoT [23]	✓	✓	✓	–	–	✗	✗	✗	✗	✗	✗
illuMINEation platform	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓

available datasets and their structure and, in turn, the building of data pipelines.

Regarding commercial platforms, CDH [12] (Cloudera Distribution Hadoop) is a multi-environment analytics platform based on open-source technologies. It offers an enterprise data cloud to execute scalable and elastic workloads. In addition, it offers Edge and AI support. 1010data [24] unifies data and analytics on its platform allowing users to perform analysis on data in the same place as it is stored. Pivotal Data suite [13] is a platform which provides a set of open-source big data technologies. The platform can be deployed in on-premises or in public clouds. Azure HDinsight [14] is a Platform as a service solution offered in the Azure cloud to execute open-source big data technologies. SAP HANA [15] is the in-memory database for SAP's Business Technology platform with strong data processing and analytics capabilities that reduce data redundancy and data footprint, while optimizing hardware and IT operational needs to support business in real time. Available on-premise, in the cloud and as a hybrid solution, it performs advanced analytics on live transactional data to display actionable information. Finally, it is important to mention, DataBricks [16] a data analytics platform based on Apache Spark ecosystem and Elastic Cloud [17] based on the ElasticSearch ecosystem.

On the one hand, these platforms are based on a license to be used and, on the other hand, are not flexible enough since they have their specific set of technologies preselected. Conversely, illuMINEation platform, despite having a set of base technologies deployed, it has been conceptualized from scratch to be able to include other technologies if necessary. For instance, let us suppose that a Graph database is required for analysing data with to many kind relationships. In this commercial platforms, this will not be viable if not offered such kind of technology by default while yes in illuMINEation platform. As it will be explained in Section 3, this is possible because of the use of Docker container management technologies alongside Platform as a Service (PaaS) and Infrastructure as a Service (IaaS) technologies. Furthermore, this platform has been designed to deal with specific mining requirements unlike these commercial platforms which are general purpose. For instance, the infrastructure and architecture of the platform can be adapted taking the mine characteristics into consideration, as it is detailed in Sections 3.2 and 3.4. Moreover, novel mine-specific user interfaces are enabled as Section 3.8 explains.

Hernandez et al. [18] define an end-to-end 3-tier architecture combining edge-fog and cloud computing layers for the IoMT (Internet of Moving Things) data-intensive applications. Their main aim is to manage the life-cycle of data generated continuously. They propose the creation of an execution model to manage workflows where data sequences can be defined. Nevertheless, it is not fully implemented and it does not contemplate security aspects, nor does it mention anything about import/export mechanisms, or about integrating third-party modules, such as analytical models, data transformations or applications. It also does not support advanced data visualization, such as 3D mine renderings with real-time data embedded in these images.

Cao et al. [19] present another edge-fog-cloud architecture for processing continuous data in the domain of smart cities. It provides Grafana for visualizing real-time data, and Superset for historical data, but it is not designed to prepare and display 3D visualizations, which is very interesting to be able to see the value of the sensors in their real location in an intuitive way. Related to security, the architecture they propose uses Wazuh [25], an intrusion detection system to detect threats in real time. Nevertheless, illuMINEation platform, in addition, uses an identity management system to authenticate and authorize users in different services.

With a focus on IIoT platforms in the mining industry, the following works highlight several solutions and research projects. The Certiq telematics solution [20] from Epiroc is a vendor specific solution for machine monitoring. It collects various parameters from Epiroc machines equipped with sensors. The collected data is transferred into the cloud and lets the equipment owners monitor the machines in real-time. Data is displayed on Web-dashboards, summarized in reports, and warnings can be sent as text messages. This allows equipment owners to monitor their whole fleet, to produce statistics, to optimize drill plans, do proactive maintenance, and finally reduce costs. However, this approach is only focused on the telemetry over the diverse machine-equipped sensors. Moreover, the solution is only available for Epiroc machines and, unlike illuMINEation platform, it does not allow the integration of external analytics models, data transformations and applications from stakeholders.

Sandvik OptiMine [21] is a digital solution for analysing and optimizing underground hard rock mining production and processes. It is not limited to Sandvik devices only, it allows to integrate between mining systems and equipment from different providers. The solution builds upon IBM's Watson IoT and consists of various components devoted to analytics, scheduling, task management, 3D model visualization, tracking and drill planning. One of the main differences with our approach is the possibility of not only tracking such elements but also the mine structure by the use of intelligent rock bolt sensors. Additionally, our solution, among other things, enables the data import from third-party sources to the platform to boost more powerful analytics by the use of external data.

There are some studies on the use of IoT and modern technology in the mining industry. In Zhou et al. [26], researchers examined IIoT applications to identify issues in the mining industry and determine whether an underground mine can support IIoT systems. Young et al. [27] presented a review of digital transformation in mining. In de Moura et al. [28] an IIoT and Advanced Analytics framework was proposed. They provided a layer-based architecture for analytics that can be used as a guide and facilitator for IIoT adoption in the

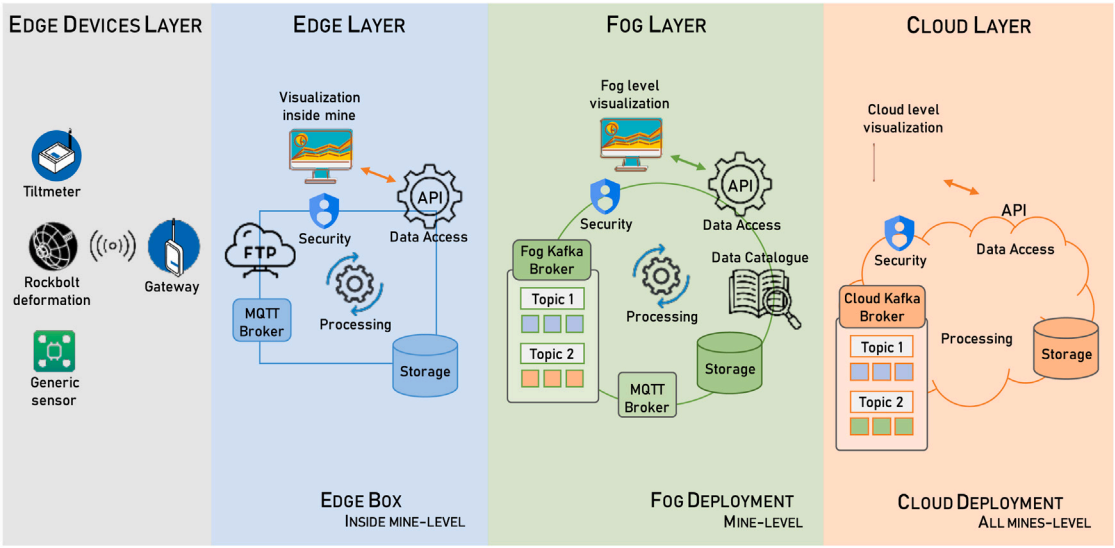

Fig. 1. High level illuMINEation IIoT Platform architecture.

mining industry, but they did not implement any platform. Salam [22] presented an IoT architecture and technology to enable the creation of a digital mining platform highlighting the exploration of the rock-fluid environment. However, the completion of the implementation of the platform is not clear. Zhang et al. [23] presented the basic structure of a Mine IoT (MIoT) system based on a three-layer IoT architecture, classifies types of sensors commonly used in underground mines by specific application, and introduces available wired and wireless communication technologies and network topologies that can be applied in underground mines. The main difference with our work is that illuMINEation platform is not only an architecture, but it is also implemented and deployed within the context of the project. Other difference with our proposed platform is that we integrate the possibility to discover and manage the datasets that are in the platform, which facilitates the work of analysts working in the data processing pipelines. On the other hand, in the article they talk slightly about data visualization, but do not present anything about advanced 3D visualization interfaces for the mining domain. Finally, concerning a key field, such as security, we ensure the confidentiality, integrity and availability of the whole platform, while Zhang et al. [23] do not elaborate on this issue at all.

2.1. Discussion

Among the architectures and platforms reviewed (see Table 1), none of them completely covers the key characteristics considered at the beginning of this section. Among the commercial or more robust solutions, some are not mine domain oriented, and those that are, are not open to integrating external modules (analytical models, data transformations or applications). While among the open-source solutions or architectures proposed by other authors, most of them remain in very early stages or in a theoretical framework. For this reason, we believe that the illuMINEation IIoT platform we propose can play an important role for the scientific community interested in the digitalization of mines.

It is worth noting that is complex to make a quantitative analysis in terms of aspects like latency, throughput or ingestion rate without knowing more detail of the rest of the platforms. Additionally, the platform has been designed to be scalable (see Section 3.9 for more details) and the resources required can significantly vary when considering each mine characteristics. Consequently, the comparison of specific measurements for a specific mine configuration goes beyond the scope of this article.

Table 2
Results of the calculation of a quantitative KPI proposed.

	Quantitative KPI
FIWARE [11]	52,22%
CDH [12]	65,56%
Pivotal Data [13]	56,67%
Azure HDInsight [14]	61,11%
SAP HANA [15]	56,67%
Databricks [16]	61,11%
Elastic Cloud [17]	65,56%
IoMT [18]	17,78%
Cao et al. [19]	86,67%
Certiq [20]	47,78%
OptiMine [21]	82,22%
Sustainable mining IoT [22]	13,33%
MIoT [23]	17,78%
illuMINEation platform	100,0%

However, at least a KPI has been elaborated indicating in which extent these platforms satisfy these requirements for data mining platforms. The creation of this metric could be differently elaborated but it reflects the main criteria considered when defining this platform. To this end, the provision of *mine-domain user interfaces* feature has been weighted with a value of importance 0,3 over 1, because it is indeed which makes the difference between a general purpose platform and a mine-specific one. In addition, it is also of paramount importance whether the platform has been implemented or is just a definition. Consequently, this aspect has also been pondered with a 0,3 over 1. The remaining 0,4 points are calculated considering the arithmetic average of the rest of the aspects but considering the *license required when is not fulfilled*. The Eq. (1) shows the formula proposed and Table 2 the resulting values obtained after the calculation of this KPI. These results clearly state the superiority of illuMINEation platform over the rest of them as these criteria is concerned.

$$\begin{aligned}\text{Quantitative KPI} = (0.4 * (&\text{Diverse sensors} + \text{Several layers}+\\ &\text{Streaming} + \text{Open-source} + \text{NOT(License required)}+\\ &\text{Integration of 3rd party modules} + \text{Data import export}+\\ &\text{Data discovery} + \text{Holistic security})/9+\\ &0.3 * \text{Mine-domain interfaces} + 0.3 * \text{Implemented}) * 100\end{aligned} \tag{1}$$

3. Multi-level internet of things platform for the mining domain

3.1. Overall perspective

The platform was initially created to deal with three main use cases of four different mines that participated in the project. Nevertheless, it was conceptualized to not only cover such use cases but future ones as well. These preliminary use cases were: establish safe zones for mining operations, boost sustainable and intelligent mineral resource extraction and perform advanced predictive maintenance and efficient operations. Once analysed in detail, the requirements stated at the beginning of Section 2 were defined. Subsequently, a four-layer platform (see Fig. 1) was conceptualized. Below, the main function of each layer is presented and the following subsections examine them in detail as well as other key aspects:

- Edge devices layer covers the elements physically installed in the mine, which includes devices like intelligent rock bolts, sensor boxes, drones and other equipment as well as gateways and edge boxes.
- Edge layer provides limited processing and storage capabilities with unstable local connection, but allows for rapid local alerting of dangerous situations, including simple GUIs.
- Fog layer – represents the mine perspective. It provides a set of tools and solutions provided by illuMINEation to perform local operations and storage to support local mine. Moreover, it also provides a dedicated mine operation dashboard showing status of the operation and applications running within the mine.
- Cloud layer – a single instance made available for whole illuMINEation IIoT platform, provides cloud infrastructure for Big Data, AI/ML capabilities, hyper-scaling processing, performant storage for the project.

3.2. Edge devices layer

This layer aims at integrating of IoT devices to send data, in an uniform way, using LoRa and LoRaWAN protocols. Such edge devices may have external low-cost sensors or be itself sensors (ex. Tiltmeters, Rock bolt heads). IoT gateways, which transform LPWAN messages to structured JSON, are used to transfer such collected data to the Edge Layer where the EdgeBox device is the responsible to aggregate data from Edge devices. The edge device layer can be deployed by using three different architectures regarding the context of use.

The first architecture is for mine environments where the deployment would be simple and there is no need for network coverage redundancy. This architecture integrates all the required LPWAN network components inside the gateway, and only sends structured data through an MQTT broker to the required edge integration consumer. The presented architecture has been used in a specific mine for air quality and mine stability monitoring. Fig. 2 shows this architecture.

The second architecture enhances the network message processing capabilities and is devoted for mining environments where network connectivity is low, there is a need to integrate LoRaWAN devices or the number of devices will be significant (see Fig. 3).

The last architecture provides a novel pattern to support several hops using LPWAN gateways to re-send the radio messages in a tree topology. This approach can be observed in Fig. 4.

3.2.1. Key flows

Periodic sensor sampling

The IoT edge device is configured to wake-up on pre-defined periods to sample, process and transmit the collected sensor data. This operational mode is applied to ensure a long-lasting battery usage, allowing to reach years of operation without any type of maintenance. The edge devices are extremely efficient in the energy usage due to the ultra-low power design of its hardware and the operation mode, as only when

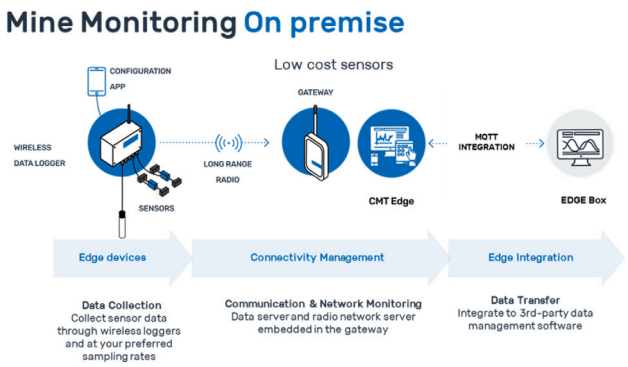

Fig. 2. First architecture: no network coverage redundancy required.

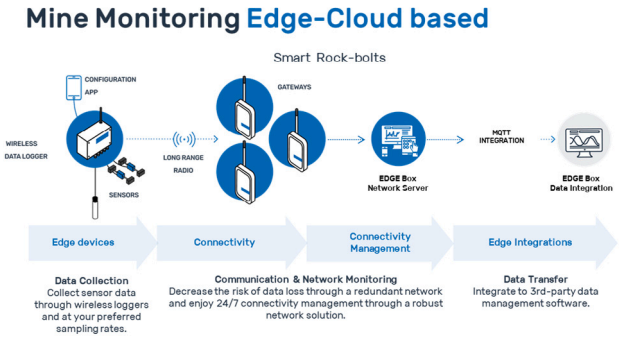

Fig. 3. Second architecture: redundancy is relevant.

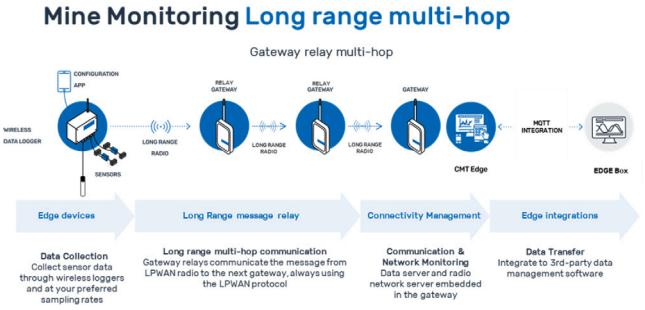

Fig. 4. Third architecture: radio messages can be re-send by the use of hops in gateways.

the device is awake is fully powered, able to read data from sensors, transmit data and receive configuration messages. The limitation to the operation mode described is the availability of the reconfiguration of the devices, as it will only happen when the device is awake and after having transmitted a message.

Long-Range Data re-transmission

As stated before, the main enhancement of the Edge Device layer consists of the multi-hop ability between gateways. Long-Range (LoRa) radio protocol is designed to support a star topology, using direct connections between IoT edge devices and gateways. Such topology is very efficient in line-of-sight outdoors communications, as coverage range can reach up to 15 km. The new multi-hop solution supports a tree topology, supporting IoT edge devices in any gateway level, and the re-transmission of messages between gateways. This solution

allows up to seven layers of gateways to re-transmit radio Long-Range messages.

IoT gateway to EdgeBox data transmission

All data from sensors connected to IoT Edge devices is sent to the main gateway which is responsible for the storage of all sensor data from the devices at the network. The main gateway is also responsible for re-transmitting the data to the upper layers, through an FTP or MQTT connection. FTP is aimed to send in an asynchronous way all data in file batches, while MQTT is aimed to re-send data in real time as soon as it is available at the gateway.

3.3. Edge layer

The Edge Analytics Box (short Edge Box) is used in the illu-MINEation platform for analytical tasks inside the mine. This software computing unit is part of the infrastructure in the mine. Technically, the Edge Box is part of an intermediate layer between the edge devices layer and applications in the mine and the fog above the mine. The Edge Analytics Box is also able to analyse data due to the demands of underground mining. The main functionalities can be summarized as:

- Processing of incoming data near where it is generated which implies a drastic reduction in the latency. The possibility of being able to perform analytics inside the mines permits to take quick decisions within the mine. For instance, if there was a dangerous situation in a specific mine zone, the low latency of the edge layer would be critical to evacuate such area.
- Data can be processed, filtered, and aggregated at the edge. Typical data filtering and pre-processing can be performed. This includes event filtering based on conditions such as value ranges, string matches, etc. Also part of these initial steps is data cleansing through the setting of defaults and the handling of *null* values through the use of defaults or functions. This can reduce the amount of data to process at the upper layer, and consequently, a reduction in networks and machines saturation and a price reduction in cloud services when applicable.
- Events can be written to the stream processing logic using time-aware queries to ingest, process and parse events from multiple input data sources in multiple formats, integrate with multiple input streams, and publish output (for example, a sensor stream) to multiple endpoints including MQTT and Kafka.
- Performing complex queries involving time-sensitive patterns like time windows. It also supports pattern and sequence recognition.

An overview of the Edge Box is shown in Fig. 5. The core of the Edge Box contains an open-source complex event processing engine to provide the streaming facilities. It also contains a local storage based on PostgresSQL which, by using specific plugins, is compatible with time-series and geo-referenced data. Moreover, the Edge Box supports the registration and management of different sensors. Finally, a logging facility is implemented as well.

3.3.1. Key flows

Integration of data received from mines

In order to transmit data from several mines to the EdgeBox, a number of gateways has been deployed. They employed the LoRaWAN communication for this data transfer. Once the sensor data is received, the gateway forwards it through MQTT to translation script of EdgeBox. This Python script extracts the data from the LoRaWAN instance and transforms it into a relevant data format. Subsequently, there is the option to either store it or transmit it for further utilization by using MQTT or Kafka producers (see Fig. 6).

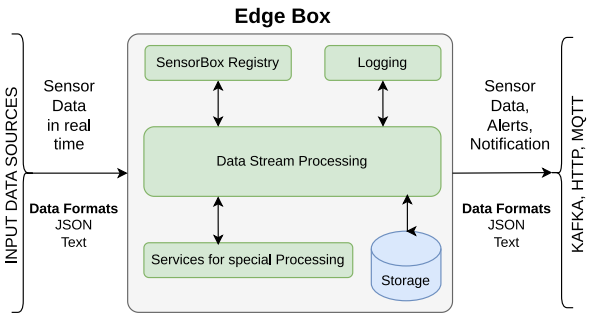

Fig. 5. Edge Box - overall structure.

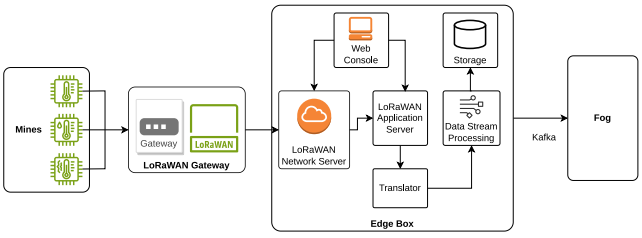

Fig. 6. Edge key flows diagram.

3.4. Fog layer

Fog computing layer is located at the mine level. There is a dedicated fog cluster for each of the involved mines, which aims at providing each mine with its own dedicated perspective, enabling the possibility to build mine-level aggregations. Being located between edge and cloud computing layers, one of its main responsibility is ensuring an adequate inter-layer data flow.

Fig. 7 illustrates the main layers and blocks of the fog architecture.

- **Automation layer:** The creation and configuration of the infrastructure have been conducted by following the Infrastructure as a Code paradigm and DevOps principles to automate it. In this way, Fog infrastructures can be easily replicated in several machines while minimizing human errors and in a more efficient way. Terraform [29] has been utilized to automate the creation of virtual machines for each Fog cluster. Then, Ansible [30] configures the virtual machines, installs the required software packages and libraries, deploys the pertinent applications, in native mode, and creates a Docker Swarm [31] cluster where also deploys on it the needed applications as containers. The use of Docker Swarm enables the stack life-cycle management by using containers and the easy integration and deployment of applications created by the uses cases.
- **Data Flow layer** securely manages the ingestion of data from the Edge, the data transmission to the cloud and the streaming communication between Fog elements. It deals with the different required formats and protocols and supports a schema mechanism. To this end, an MQTT Mosquitto [32] broker and a Kafka [33] cluster have been integrated in the stack. Regarding security, client and server certificates are utilized.
- In the **Persistence Layer** different data storage structures are supported. Specifically, HDFS [34] manages files storage in a distributed fashion, ElasticSearch [35] enables the persistence of semi-structured data in JSON format while providing advanced search capabilities and PostgreSQL [36] allows the storing of structured data in a table format. In addition, the PostGIS plugin of PostgreSQL is used to manage spatio-temporal data which is necessary for the mines characterization.

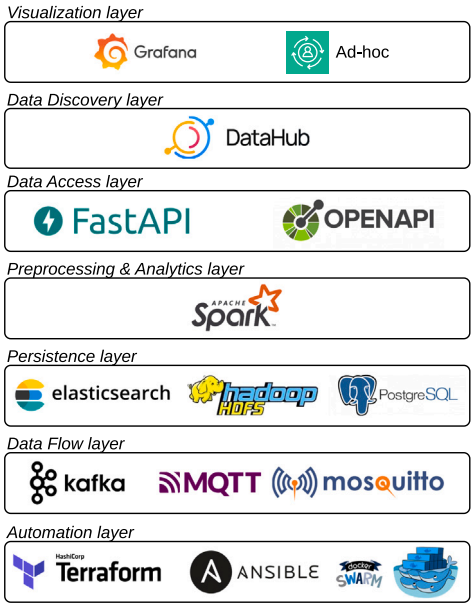

Fig. 7. Fog architecture.

- **Preprocessing & Analytics layer** aims, at the one hand, at cleaning and preparing the data for more complex processes, such as aggregations, analytics or simple visualizations, and, at the other hand, at performing such complex operations to create aggregations at the mine level or executing ML algorithms. For this purpose, an Apache Spark [37] cluster has been deployed to allow the execution of both batch and streaming operations and the inclusion of ML distributed libraries. Moreover, additional preprocessing or analytic processes could be deployed into containers outside of the Spark context by using underlying Docker Swarm [31] installation.
- **Data Access layer** provides a set of APIs based on the OpenAPI syntax for interacting with the underlying storage technologies. In this way, stored entities are not directly exposed to internet delegating to the API authorization and authentication policies. In addition, this approach enables a unified access for all the storage technologies instead of having to master each access mechanism. To this end, FastAPI [38] has been utilized which is a Python framework devoted to build APIs.
- **Data Discovery layer** is focused on managing and discovering the data. To this end, LinkedIn DataHub [39] acts as a catalog for an organization's data assets, making it easier for data consumers to search for and discover relevant datasets for their analytical and business needs. It allows to store, organize, and manage metadata related to the data assets. Metadata includes information about the data's source, quality, structure, lineage, and more. This helps mine professionals understand and work with data more effectively.
- **Visualization layer** enables the creation of interactive dashboards to exploit the data processed by integrating Kibana [40] and Grafana [41] which are an open-source tools for observability, allowing users to create custom dashboards to visualize their data in a way that is meaningful to them. In addition, for specific purposes ad-hoc web clients can be deployed in Docker containers and making use of the API to get access to the data.

3.4.1. Key flows

Getting Keycloak token

Some of the interactions with the Fog layer requires the interaction with its API. As a consequence, before performing any action, a Keycloak [42] token must be retrieved to unequivocally identify who is performing the action.

Persistence structures creation

In the context of the illuMINEation platform, a *device-type* is an entity representing the data schema from a specific device. In consequence, some data structures must be prepared in the platform before ingesting data from a new *device-type*. Concretely, a Kafka schema associated with such device type must be created and persisted in a Kafka Schema Registry [43]. This schema is set in the producer to enforce the utilization of the defined fields. Moreover, an ElasticSearch [35] index is created to store the ingested raw data in JSON format and a PostgreSQL [36] table in tabular format. For this purpose, an API endpoint has been created to deal with the creation of such required structures. Thus, when ingesting data to the fog layer, all the registries having the same *device-type* (included in the topic name) are stored together, but they can be filtered by using other properties like the edge-id or the location.

Data Ingestion

The ingestion of data to the Fog layer can come from two different types of sources: the edge layer or data imported from an external source like a third-party platform or a local environment. The main entry point to ingest data to the Fog is the Kafka broker. Once the previously described steps for creating the persistence structures have been carried out, data should be produced to the corresponding topic. To this end, the first step is to define the topic name which for data coming from the edge is a composition of the name of the project, the level of processing, the edge id, the location and the device id and type and for external data the project or source id, the processing level and the schema name.

Consequently, the name of a topic for ingesting data from the edge layer could be: *illu. raw. edge1. tunnel1. Sensorbox1. rockbolt-device* and the name for ingesting data from the another project called Piacere: *external. piacere. raw. pressure*.

Then, a Kafka producer must be adequately prepared. After following these indications, the production process can start to ingest data to the Fog layer. Subsequently, the flow represented in Fig. 8 is triggered in the Fog layer: (1) the layer receives the data in the Kafka broker topic, (2) a Spark Generic Ingestor process consumes the data from the Kafka broker, processes it and stores the ingested data in both PostgreSQL and ElasticSearch in a structured way. In this way, the raw data ingested from the edge is available for further required processes. Moreover, this process also redirects the data to two different Kafka broker topics: one the Fog's broker which enables the possibility to execute ad-hoc streaming processes and the other to the Cloud's broker for further processing. These ad-hoc process in the Fog can deal with any specific requirements of the use cases. They can be natively deployed in Spark jobs or in any other piece of code by using Docker containers.

Data Discovery

The integration of DataHub enables the discovery of data, both by using its user interface and programmatically since it exposes a GraphQL [44] API to interact with the underlying content. In this project, the DataHub API is not externally exposed but wrapped in the FOG API to expose only the subset of functionalities relevant for the project. For this purpose, an endpoint for programmatically querying to DataHub about the underlying datasets has been created. This method can be parametrized selecting the number of items to retrieve, the offset from where start obtaining the data and a term to filter the results. Fig. 9 provides an example of the output involving datasets from PostgreSQL [36], Kafka [33] and Elasticsearch [35]. It can be observed that each dataset type shares the *URN* and platform properties but also have specific ones related to the source type. Then, it is possible

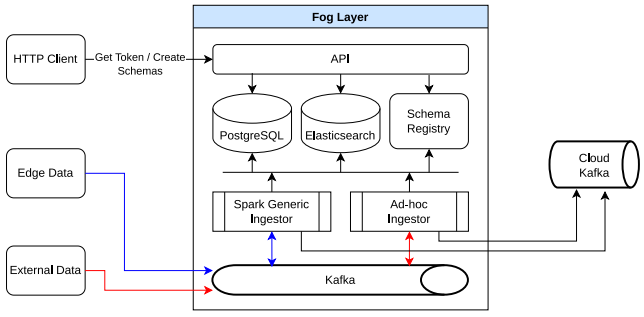

Fig. 8. Ingestion flow.

Fig. 9. An example response when querying the Datahub programatically by using the API.

to interact programmatically with each dataset regardless of its type by using the obtained *URN*.

Querying Persistence Layer

Once an URN has been retrieved in the discovery process, specific queries can be performed to operate the underlying persistence technologies. Regarding the target technology, these queries vary to be accommodated to the correspondent intrinsic nature of each one. Specifically, ElasticSearch supports its query DSL syntax, PostgreSQL the SQL language and Kafka enables to retrieve data from a specific topic and optionally be filtered by start and end dates.

3.5. Cloud layer

Cloud layer's main responsibility is to provide high availability and throughput of computing and storage resources for the IIoT platform. While other layers are focused on services for singular mine operation, the cloud layer provides solutions on global scale: algorithms relevant to processes running in the mines, comparing and cross-referencing data coming from various sources as well as monitoring infrastructures and services running on several mines. It can also serve as a computational and storage proxy for lower layers, where it provides an IaaS-like capabilities. Moreover, to provide support for application developers relying on aggregated data, cloud layer provides a PaaS solutions allowing easy deployment of applications in the platform. Similarly to fog layer, cloud can also be divided into six main layers with very specific responsibilities, whose main goals can be summarized as follows and illustrated in Fig. 10.

- **PaaS/IaaS technologies** is a backbone of the Cloud deployment and provides resources to configure and run servers, backend applications as well as applications of the users and developers. All machines providing cloud functionalities in illuMINEation are configured and running under OpenStack – a private cloud computing infrastructure management system. It allows easy creation, setting up and monitoring of resources. Deployment of the applications is done in a hybrid PaaS/IaaS. Consequently, most of the backend services, as well as some applications and algorithms developed during the project, are deployed directly on machines created and operated by OpenStack. The applications are most typically containerized as docker images and deployed by utilizing one of the three approaches: (a) directly running as docker image (e.g. Consul Agents); (b) run and governed by docker compose (e.g. Kafka brokers); and (c) run and governed by Kubernetes (e.g. DataHub).
  The second platform provided to run applications is OpenShift, which provides a PaaS approach on the cloud layer. It is provided by an OKD upstream implementation built on top of OpenShift, which focuses on continuous application development and deployment.
- **Data Transport layer** allows data to be transferred into and out of the cloud layer of the platform and includes data transformation and replication between the repositories. The ingress point of data coming into the cloud layer is Apache Kafka [33].
- **Stream Processing layer** provides tools to develop and deploy real-time applications operating on streams of data. One of the main focus of the cloud layer has been to provide applications with means of accessing, processing and acting on real-time streams of data coming into the cloud data layer. The first utility available to developers is Kafka Streams technology [33]. It allows to build not only highly reactive applications operating on incoming data, it also allows aggregating, joining multiple streams, access data in a table-like manner as well as producing new streams of data. In addition, there is the ksqlSB, a standalone purpose-built stream database. Its functionality is based on Kafka Stream API, but moves all the computations done on the streams from the applications themselves into the ksqlDB process. The applications can then access results of those transformations by utilizing SQL language.
- **Persistence layer** provides long term storage of data. Kafka itself can store the data infinitely in the topics, but cloud offers dedicated deployments of Big Data stack persistence solutions including MongoDB, ElasticSearch, and InfluxDB for timeseries data.
- **Data and Service Discovery layer** – allows applications and services to search for relevant data or other services within the cloud. There are two main discovery services deployed within this layer. The first one is DataHub, already utilized in the Fog layer. Alongside data, the other crucial element of the cloud offering is its array of services and applications that are deployed on various frameworks within the stack. To ease their discovery and development process and to ensure secure communication between the services, a HashiCorp Consul has been installed.
- **Visualization layer** is responsible for hosting tools required by the project to visualize different kinds of data that is obtained by mine use cases, focusing on global and multi-mine aspects. The main generic-purpose visualization application provided is Grafana [41].

3.5.1. Key flows

Deployment of application in OpenShift

OpenShift allows easy provisioning, configuration, running, scaling and monitoring of containerized applications. Developers wishing to include one of those databases in their stack can just easily spin-up pods, configured with relevant persistent volumes, and connect to them using internal or externally configured routes. Other applications, including those developed within the project, can also be deployed in OpenShift.

The image should also be uploaded to some recognized registry for images and then, manage it from the *Application Deployment section*. There, a YAML configuration can be provided to orchestrate the new deployment. In this YAML configuration file, the image, and the desired resources (CPU and memory) for the application are specified as well as the necessary persistent volumes or the replicas running the application. The full definition example can be found in Listing 1.

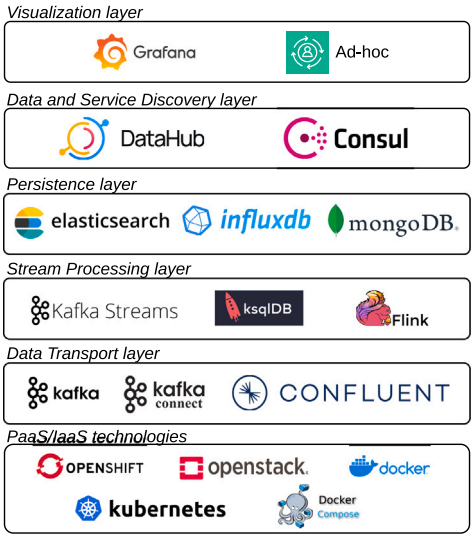

Fig. 10. Cloud architecture.

Listing 1: YAML configuration file to orchestrate a new deployment.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: illudemoconsul
  spec:
    replicas: 1
  selector:
    matchLabels:
    app: illudemonconsul
    template:
      metadata:
        labels:
          app: illudemonconsul
    spec:
      containers:
        - name: app
      image: img.registry.com/illuconsul:1.0.0
      resources:
          limits:
            cpu: "500m"
            memory: "521Mi"
          requests:
            cpu: "250m"
            memory: "256Mi"
}
```

Upon registering a new deployment, an application is started. By means of the Openshift portal, the status of the application is displayed and its configuration can be edited. Moreover, the execution history of the application as well as details of currently running replicas can also be displayed. It allows to display logs from the application as well as real-time adjustment of the settings, such as scaling number of running pods providing application.

Streaming application

This key flow uses Kafka Streams API to connect to Kafka broker, attach to a topic, monitor it for incoming data, perform aggregation and finally send results to another topic.

ksqlDB example

ksqlDB can also be used to create mapping and processing structures, but with much more familiar SQL-like statements. ksqlDB is running its processing topology in a dedicated service. It provides access by API as well as CLI, which will be used in this example. A named stream representation of a topic is created, where also internal time stored within the payload is used as a timestamp for the stream. See first step of Fig. 11 for the case of rockbolts monitoring.

In the second step of Fig. 11, a new table is created, which will be populated with latest deformation status for each rockbolt. Finally,

```
ksql> CREATE STREAM DEFORMATION_UPDATES(
>       id VARCHAR KEY,
>       date VARCHAR,
>       double DEFORMATION) WITH (
>       KAFKA TOPIC = 'illu.silver.rockbolt-device',
>       VALUE_FORMAT = 'JSON',
>       TIMESTAMP = 'payload.datetime',
>       TIMESTAMP_FORMAT = 'dd-MM-yyyy HH:mm:ss'
>       );

ksql> CREATE TABLE LATEST_DEFORMATION_STATE AS
>       SELECT ID, DEFORMATION, MAX(ROWTIME) AS DATE
>       FROM DEFORMATION_UPDATES
>       GROUP BY ID
>       EMIT CHANGES;

ksql> SELECT * FROM LATEST_DEFORMATION_STATE ID='Sensorbox1';
```

1 2 3

Fig. 11. KSQL operations(create stream, create table and query).

the table can be then queried as any other table in other databases, by providing a simple SQL statement, retrieving latest value for specified id, see third step of Fig. 11.

3.6. Security aspects

Ensuring confidentiality, integrity and availability is one of the key requisites of illuMINEation platform. To build a resilient implementation of the previously defined infrastructure, security is one of main pillars. Consequently, a threat analysis has been carried out to ensure security aspects are adequately handled in all the layers of the mining systems. Such analysis, which is aligned with the ENISA [45] and Cloud Security Alliance (CSA) [46] standards for IIoT, has been crucial to discover the following findings:

- As this project contemplates a system divided in different layers, some of them physically distant, first of all the secure reachability must be ensured.
- The Cloud Layer is the less critical layer for the mine business operation.
- Secure communications, including wireless communications which are common in the mines, should be ensured through the use of secure REST APIs at least in the fog to cloud channel. It is also recommended to implement input validation in all REST services so as to prevent malicious injection attacks.
- Users of the illuMINEation IIoT platform services should be appropriately authenticated in order to avoid intruders (unauthorized users) accessing the illuMINEation services which may pose risks and cause undesirable damages to the production itself.
- Some sensing and edge devices already include some security protections at the hardware or internal firmware level. For instance, when a LoRaWAN of the platforms' nodes is reconfigured, the firmware binary is also encrypted and signed.
- It is highly recommended to develop security monitoring agents for the edge, i.e. capable to perform detection of anomalies and intrusions directly in the edge, without the need of first transmitting the data to the back-end. This will allow fully availability of some detection capabilities in the mine and maintaining at all times a minimum protection level of the IoT ecosystem.

In summary, in order to protect the wide range of use cases running on top of the different technology stacks, the security aspects address risks related with the following points which are depicted in subsequent subsections:

- Protecting the confidentiality of the information in transit between layers.
  Mutual authentication in Transport Layer Security.
- Ensuring the authentication of users and granting only the adequate assigned permissions.
  Centralized Identity Provider.

- Protecting the administration interfaces from unauthorized entities.
  Single Sign-on.
- Detecting potential cyberattacks to edge devices affecting the availability of the platform.
  Security auditioning and threat modelling on hardware devices.

3.6.1. Identity management and SSO

To protect the system information and integrity from unauthorized users, it is a mandatory requirement to implement an Identity Management solution. This tool serves as a database for users and roles and is responsible of verifying the user identity to grant only the needed privileges in the system where is requesting the access. The Identity Management tool will also be the gatekeeper to all system and services through the platform, unifying the access roles and providing a simple and user-friendly Single Sign On process.

Although the identity management needs in the project initially are focused on the project staff and their associates, the system must be scalable enough to enable federating new partners credentials repositories or adding new identity information providers to the authentication schema. Therefore, the Identity Management service solution chosen for illuMINEation is based on the open-source project Keycloak [42] which is an Identity and Access Management solution for modern applications and services. It provides user federation, strong authentication, user management, fine-grained authorization, and more. The solution supports OpenID Connect, OAuth 2.0, and SAML authentication mechanisms.

3.6.2. User, groups, and roles

The authorization and identification schema for the platform is structured around each different mine, abstracting the specific capability permissions to the tool configuration itself and defining a few generic roles to be mapped by the application administrators.

In addition, there is also more groups of users that does not belong to any of the previously physical-defined groups that needs to be contemplated in the schema, the illuMINEation members, a data exchange, and a potential public group of testing.

3.6.3. Platform communications

An Identity Provider (IDP) tool provides a guided interactive flow to ask the users for its credentials, and depending on the access privileges, granting or not accessing the system or performing changes in it.

However, other communications are not being produced by a direct action of an identified user. For this situation, Mutual Transport Layer Security (mTLS) [47] is particularly relevant for securing API communications, especially in microservices architectures. With mTLS, each microservice can authenticate itself to others, reducing the risk of unauthorized access to sensitive APIs. In this platform, Kafka and MQTT brokers are secured by using this approach. Each service connecting to the broker is issued a certificate signed by CA. Access rights to topics in Kafka are specified by creating Access Control Lists (ACLs) on per topic basis.

3.6.4. Security audit and monitoring

To audit the security of the edge, gateway devices have been tested against some attack and discovery vectors to find how vulnerable can be and discover what kind of attack patterns can be detected based on the device logging capabilities.

The restricted nature of the software placed on the device makes it hard to install additional auditing tools that are commonly used in traditional systems to capture network traffic or internal operating system processes in order to find an attack.

The device where the tests were performed is a CMT Edge LS-G6-KIO-GW868 [48]. In addition, the device has been tested working paired via LoRa [49] with a tilt sensor like the one used in the mines, more specifically of the model LS-G6-INC15 [50].

Table 3
Attack Tactics and Techniques tested against the gateway.

Tactic	Technique ID	Technique Title
Discovery	T1046	Network Service Discovery
Impact	T1498	Network Denial of Service

To emulate the working conditions of an actual environment, the gateway was configured to send the captured data from the sensor to an FTP server created for this purpose in a laboratory laptop. This way, the gateway simulates to push the measured metrics in real time to the next processing layer.

The attack scenario proposed, as the use case, for performing the analysis is based on the MITRE ATT&CK Matrix for Enterprise tactics and techniques [51]. In this case, we assume that the attacker has already gained physical access to the site and is in the same network as the gateway. Tactics and techniques chosen for the use case are in the Table 3.

To perform the attack Kali Linux and NMAP were successfully used, specifically the thc-ssl-dos tool that allows to exploit the processing part of the RSA algorithm that occurs on the server side, so that the gateway can be saturated causing a performance peak in the system and blocking its operation.

In addition, as a security countermeasure, an intelligent security agent based on anomaly detection was trained and tested. This agent monitors the gateway status to detect any attack as soon as possible. This model learns patterns that represent normal behaviours in order to identify abnormal activities at production time. Therefore, relevant signals have been monitored such as a CPU usage metric, a key Indicator of Compromise (IOC) for the deny of service (DoS) attacks detection. For doing so, on the one hand, the moving average technique has been applied to reduce some fluctuations. And subsequently, the training data have been standardized. On the other hand, a One-class SVM (OC-SVM) model was trained. This unsupervised machine learning technique aims to find the smallest hyper-sphere containing the normal data. While for test samples, OC-SVM only judges whether they belong to that class. Since only normal data is required for training, it is effective for detection of unknown attacks. Finally, the performance of the anomaly detection model was evaluated against the test dataset and the model works properly as it detects abnormal behaviours, while the remaining samples are considered regular values.

As result of this exercise, the gateway was found as vulnerable to DoS attack via its administration web interface. To reduce the attack surface, the following actions in deployment are recommended:

1. Disable the HTTP web interface just after finishing the initial device configuration.
2. Physically block the ethernet port or place the gateway in a protected box to prevent an attacker to gain access plugging an ethernet cable directly to the device port.
3. Isolate the network where the gateway network adapter is connected to prevent attackers scanning open services or reaching the device via network interface.
4. Train and implement an intelligent security agent based on anomaly detection that allows operators to be warned in case of attack.

3.7. Data import and export mechanism

The platform provides a mechanism to import and export data from and to third-party systems. In this way, data produced from other platforms or systems can be utilized to obtain an added value and other platforms can exploit the data generated by using the illuMINEation

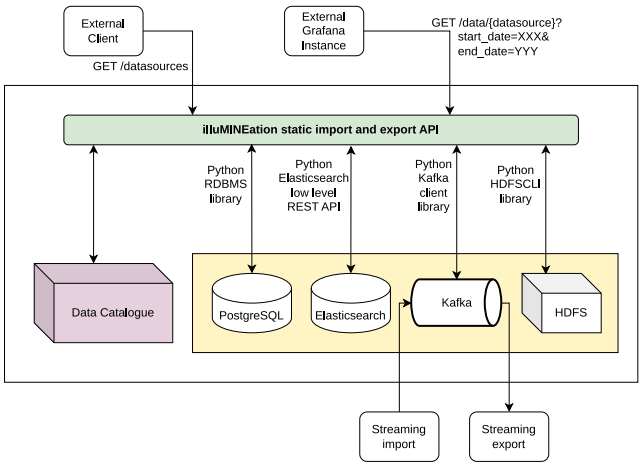

Fig. 12. Fog architecture.

platform. For this purpose, two distinct modes are proposed: streaming and batch.

The streaming ingestion to import data from external platforms is conducted identically as ingesting data from the edge but modifying the topic name. That is to say, a specific data structure must be built for each device type and then, the data can be ingested to a Kafka topic which name is normalized by concatenating relevant meta-information. After that a Spark process manages the incoming information and persists as well as redirects it. Regarding streaming exportation, specific consumer clients can connect to the Fog and Cloud Kafka brokers. Once having the pertinent credentials, DataHub can be utilized to discover the available Kafka topics which according to their data processing level can be labelled as raw, silver or gold. The importing and exporting of batch data must be carried out by using the API. As explained in Section 3.4, the API offers a set of endpoints to insert and get data from the underlying persistence technologies. DataHub can be then utilized to discover data both by using its user interface and programmatically. Finally, once the target persistence entity has been identified, specific ElasticSearch, PostgreSQL and Kafka queries can be executed. In this way, the import and export of batch data is guaranteed by using the corresponding query mechanism. It is worth noting that Kafka does not allow to insert data in batch mode, only retrieve it by optionally filtering it by dates (because this functionality has been programmed explicitly for illuMINEation). The insertion of data to Kafka must be done in streaming by using a correspondent producer. Once the data have been discovered, Grafana can also be used to interact with it by the instance deployed in the stack or by external Grafana instances. Fig. 12 summarizes the ideas of the import and export mechanism.

3.8. ToolBox for modelling 3D visualizations

3D models of the mines were provided as LIDAR data and as CAD construction data (dxf). These data formats are not suitable for being displayed in a web browser. Therefore, they must be reduced in resolution and converted to a suitable format using open-source programs. Subsequently, the 3D models can be displayed in the web browser and augmented with the streaming data to provide a real-time perspective of the mine state.

To this end, the illuMINEation Toolbox has been conceptualized aimed at providing, in an unified fashion, the toolset necessary for preparing the 3D visualizations. As a result, the necessary open-source applications have been integrated into a Docker container with desktop support for managing such applications in a usable environment. Specifically, this toolset is composed of the following applications: (a) *Blender*: it is a 3D content creation suite that encompasses various aspects of 3D production, such as modelling, sculpting, texturing, shading rendering, etc.; (b) *CloudCompare*: it is a 3D point cloud and mesh processing software, which is designed for working with large datasets of 3D point clouds; (c) *Drawio*: it is a free web-based diagramming application that allow users to create a wide range of diagrams; and (d) *Meshlab*: it is a portable and extensible system for the processing and editing of unstructured 3D triangular meshes. It provides a set of tools for cleaning, editing, inspecting, rendering, and converting meshes generated from 3D scanning, computer-aided design (CAD), or other sources.

3.9. Scalability

It is worth noting the ability of the platform to be adapted in terms of scalability to diverse contexts where different resources can be necessary. Regarding the edge device layer, three different architectures have been presented covering scenarios with diverse necessities and mine topology. Edge Box layers can be adjusted by deploying more gateways to when required.

As earlier mentioned, the fog layer provides the mine perspective and a dedicated cluster manages the mining flows. This cluster can be prepared with humble resources for small mines or with a bigger number of machines to deal with large mines with tons of sensors emitting data. This is possible because, on the one hand, the Docker Swarm technology enables the easy integration of new machines to the cluster on demand to achieve more computing and storage power and, on the other hand, at the application level the utilization of Docker Swarm permits to quickly replicate the required services like the API to support more heavy workloads. Moreover, the use of big data technologies like Kafka or Spark allows to distribute the computing requirements between the diverse nodes and applications and ElasticSearch or Haddoop permits the distribution of the storage. Finally, technologies like Terraform and Ansible boost this scalability by managing, in an automate fashion, this required provisioning of new machines and configuration of applications.

Similarly, the cloud layer enables the easy integration of new machines by using OpenStack, the management of the applications and their required replicas with Open Shift. Additionally, as commented, the cloud is also based on various Big Data technologies which enable the distribution of computing and storage among the diverse machines.

In this direction, several studies have accomplished to verify the scalability of these technologies in diverse contexts like Choundhary et al. (2022) [52], Bazai et al. (2021) or Ileana et al. (2024) [53].

3.10. Environmental impact

All the technologies shown in Section 4 (intelligent rock bolts, MWD, predictive maintenance) are considered being essential tools and improvements for the mining industry. Their joint application and comprehensive consideration via tools like the IIoT-platform presented in this paper are vital for further improving the efficiency, sustainability, environmental performance and safety of the mining industry. For detailed results on how the individual technologies contribute to the respective topics please refer to (Nöger et al. (2023) [54]; Nöger et al. (2021) [55] and Varelija and Hartlieb (2024) [56].

4. Platform validation

Section 2 presented several functionalities required for the platform. Derived from such functionalities the following hypotheses have been articulated in order to demonstrate how illuMINEation platform validates them:

- Hypothesis 1: The platform is able to manage data from sensors.
- Hypothesis 2: The platform is able to transfer the data through the diverse layers in streaming mode.

Table 4
Recap of where each hypothesis has been validated.

Hypothesis #	Validation place
Hypothesis 1	Section 3 and Scenario 1
Hypothesis 2	Section 3 and Scenario 1
Hypothesis 3	Section 3
Hypothesis 4	Sections 3.4 and 3.5 and Scenario 2
Hypothesis 5	Section 3.7 and Scenario 3
Hypothesis 6	Sections 3.4 and 3.5 and Scenario 1
Hypothesis 7	Section 3.8, Scenarios 1 and 2 and Section 4.2
Hypothesis 8	Section 3.6

- Hypothesis 3: The platform utilizes open-source technologies and does not require a license.
- Hypothesis 4: The platform enables the integration of 3rd party modules.
- Hypothesis 5: the platform allows exporting and importing data from/to third-party systems.
- Hypothesis 6: the platform boost data discovery.
- Hypothesis 7: The platform provides mine-domain interfaces.
- Hypothesis 8: The platform integrates a holistic security mechanism.

Hypotheses 1 and 2 are deeply addressed in Section 3 during the explanation of the diverse layers. In such section, Hypothesis 3 is also validated because it explains the set of open-source technologies utilized in the diverse layers. In addition, currently, there is no associated license since the platform has been developed in the context of an European project. The inclusion of 3rd party modules, services algorithms or applications (Hypothesis 4) is managed by using Docker containers. This technology is integrated in the Fog layers by using Docker Swarm and in the Cloud layer by using Open Shift (Check Sections 3.4 and 3.5 for additional details). In addition, Hypothesis 5 is tackled in Section 3.7 and data discovery (Hypothesis 6) in Sections 3.4 and 3.5. Moreover, the mechanisms for providing mine-domain interfaces (Hypothesis 7) is exposed in Section 3.8. Finally, Section 3.6 has deeply detailed the security approach which validates Hypothesis 8.

However, to better validate some of these hypotheses and the remaining ones, this section presents three successfully conducted scenarios:

1. The former demonstrates how the platform can be used to move data from sensors in streaming and perform the required operations through the distinct computing layers. Then, there is an example of how the data managed can be discovered. Moreover, a mine-specific 3D user interfaces are presented. Thus, Hypotheses 1, 2, 6 and 7 are addressed.
2. The second illustrates the use of novel user interfaces in the mining domain by utilizing the toolbox module developed in the context of the project (see Section 3.8 for additional information). This way, Hypotheses 4 and 7 are validated.
3. The latter presents a scenario for importing data to the platform to enable a better exploration using the platform tools, in consequence, Hypothesis 5 is validated.

Additionally, to highlight the satisfactory results of the mine-domain interfaces Section 4.2 presents the feedback provided by the users when interacted with them. As a summary, Table 4 recaps where each of the hypotheses has been addressed which implies that they can be considered as validated.

It is worth noting that the measurement of performance aspects like ingestion rate, data processing velocity, latency or throughput are beyond the scope of this paper. This is due to the fact that, as explained in Section 3.9, the Fog layers resources should be scaled considering each specific mine characteristics and, in addition, the platform has been conceptualized to be flexible enough to be easily scaled on demand. That is to say, if any performance metric would be inadequate for the specific requirements, the provision of additional hardware resources and a proper big data applications configuration would easily amend the issue. However, as a matter of illustration, below, a summary of the resources utilized in the fog layer to accomplish these scenarios is provided:

- Fog layer: two clusters of five virtual machines (with 1 TB of hard disk, 8 CPUs and 8 Gb of RAM each machine) each cluster with Ubuntu 22.0.4 installed. HDFS ans Spark are installed natively and theb rest of the applications are managed by Docker Swarm as containers with a replication factor of three. Additionally, there is 5 more virtual machines for each cluster to work as Spark workers each one with 16Gb of RAM.

4.1. Application scenarios

4.1.1. Intelligent rock bolts

One of the successful scenarios where the architecture has been evaluated is in the context of managing streaming data coming from intelligent rock bolts. This technology allows to measure the deformation on a rock bolt which are essential tools for securing the rock mass in underground mining. Besides information on the deformation of the bolt, they provide the rock bolt identifier and the timestamp. In order to enhance the mine security, IlluMINEation mines have been equipped with these intelligent rock bolts. In this way, the mine can be constantly monitored enabling the possibility to detect deformation on the rock bolt and accordingly movement of the rock mass, contributing to a better understanding of the underground situation and enabling enhanced planning and safety features.

Initially, the intelligent rock bolts are integrated to the edge device layer by using its LORA interface. This layer uniforms the data structures in JSON format from different devices acting as a gateway and publishes it to a MQTT broker specific topic. Then, the Edge Box process the data, adds the tunnel and mine identifiers and submits the records to the Kafka Broker on the Fog layer by using a Python client.

Prior to start of the data flow, the required persistence Fog structures should have been previously created. The result of this is the creation of the corresponding *rockbolt-device* schema, *silver-rockbolt-device* ElasticSearch index and the *silver.rockbolt_device* PostgreSQL table.

Then, the Kafka topic *illu. raw. edge1. tunnel1. Sensorbox1. rockbolt-device* is utilized to send in streaming the data from the edge box to the Fog layer. The Spark Generic ingestor process consumes this data, stores it structured in ElasticSearch and in PostgreSQL, and redirects it to the fog and cloud Kafka brokers to the topic with the name *illu. silver. edge1. tunnel1. Sensorbox1. rockbolt-device*.

In addition, an ad-hoc Spark process has been created to provide rock bolt deformation aggregations. This way, statistical information is generated and properly stored in an ElasticSearch index called *gold-rockbolt-device-agg* and a PostgreSQL database called *gold.rockbolt-device-agg*.

Fog Layer also provides users a data catalogue, Datahub, which facilitates the search and discovery of relevant datasets for further analysis or business needs.

Finally, some Grafana dashboards have been created to visualize this information. Fig. 13 shows a created dashboard that allows users to easily and quickly observe the rock deformation detected by the intelligent rock bolts.

The cloud layer, which can be set up to receive the rock bolt data (both raw and aggregations) from each fog instance (i.e., from different mines), allows to create additional, real-time operations on the data. Expanding on the streaming application key flow of the Cloud layer, the streams of data from the topics can be filtered by specific location (tunnel1, tunnel2 etc.) and calculating windowed aggregates similar to the example. Moreover, two streams of data from each location can be

Fig. 13. Grafana dashboard presenting rocks deformation.

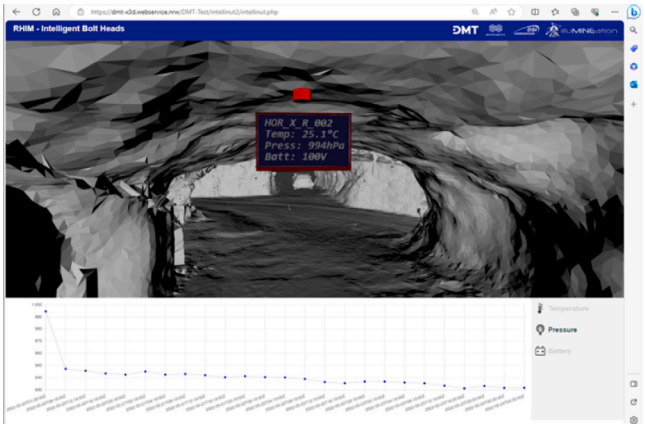

Fig. 14. Processed LIDAR data in a browser with Rock Bolt data.

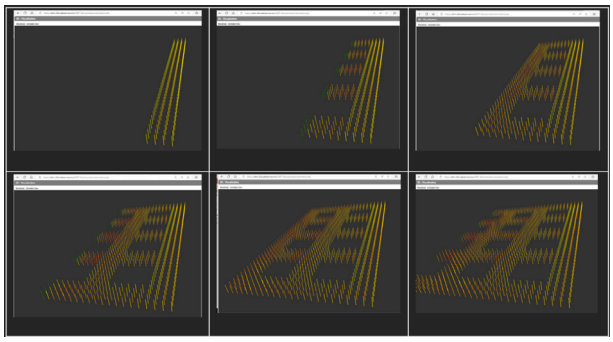

Fig. 15. Screenshots of rock bolts load at different times.

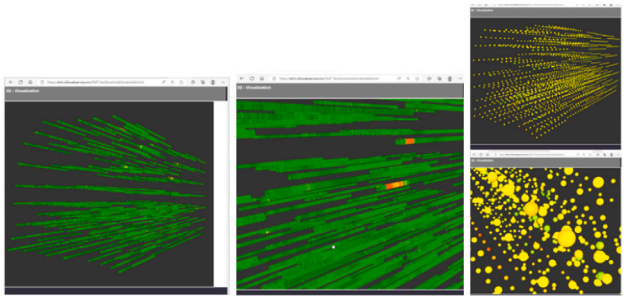

Fig. 16. 3D Visualization of boreholes.

joined together, allowing for the processes to simultaneously operate on both of them.

For a large number of visualizations, the Grafana server is used. However, visualizations with 3D modules are delivered via an HTML server. For the 3D models the open-source framework *x3dom* with its runtime unit for 3D graphics is used. For this the Javascript components are loaded in the web page and, after that, there are commands for loading *x3d* files and the 3D operating commands.

In order to link the rock bolt environment data to a 3D model of the mine, the previously explained Toolbox (see Section 3.8) has been used to convert the LIDAR to mesh data and then to store it in X3D format. For this conversion, MeshLab and Blender applications of the Toolbox are used. This toolbox runs on a remote server, whose hardware must be supported by powerful graphic units. In this way, this data, reduced in volume, can be displayed over the network and the response time of the website is acceptable. Fig. 14 shows the 3D model in a browser, where the head of the rock bolt is highlighted in red and a virtual display with some information about sensors is attached.

The resulting Rock Bolt data can also be used to create 3D models that grow with the excavation. The colour of the rock bolt follows the load. So, you can see the locations where the pressure in the rock changes. Fig. 15 shows screenshots of an animated version in the browser.

4.1.2. Novel integration of MWD visualization in the dependable IIoT platform

The Measurement-While-Drilling (MWD) technique is commonly used to gather drilling data on operational factors such as penetration rate and drill bit rotation speed. Despite being a mature technology, MWD presents a significant challenge in calibrating to changing rock conditions to ensure data reliability. Accurate drill hole spatial location is crucial for successful excavation via blasting, ensuring minimal over break and waste dilution, achieving desired rock fragmentation, enabling highly selective mining and more.

The data of MWD comes from based on legacy drilling rigs. Local persistence of sensed data in the drilling rigs will not be required and the analysis can be done in the with the IIoT Platform. Such a seamless integration/interaction of the IIoT platform and its applications with other legacy (proprietary) systems already installed – and widely used in mines – is also an issue as the IIoT has to exchange data/information with them. This is also the case with mine planning systems.

As part of IlluMINEation platform, innovative visualization techniques have been developed using web technology. In the case of MWD, the operating data from the drill rig is measured and transmitted to the platform. The rig also records the position of the boreholes in relation to each other and the penetration depth of the drill. This allows a property to be represented as a cylinder for each individual drilling operation. Fig. 16 shows the cylinder of the borehole with the coloured area being the calculated value of the rock hardness.

4.1.3. Predictive maintenance

The predictive maintenance use case is aimed at taking advantage of extensive amounts of sensor data captured in different machines of the Boliden and Breitenau mines. Conceptually, these data can be used to train machine learning models that are able to anticipate degradations and failure models, so maintenance operations can be performed in a preemptive way, in front of the usual reactive and planned maintenance types. The task involves capturing, storing, dealing and visualizing a large amount of data, a convenient application case for the illuMINEation platform. Sensor data from the *boomer* machines and event data from their operating system were obtained and ingested to the Fog layer and presented in the visualization tool to improve their exploration.

The steps performed have been the following: firstly, the different necessary data schemas alongside the underlying persistence have been created by using the API. Next, it has been defined a representative Kafka topic name following the specified pattern convention like *external.boliden.raw.maintenane-records*. Afterwards, the data have been submitted setting special attention to the date format and processed by Spark module which persists them in PostgreSQL and ElasticSearch and

Fig. 17. Dashboard with data from Boliden mine.

redirects them to the corresponding *silver* Kafka topic and to the cloud to perform ad-hoc further processing if necessary. After persisting data, interactive dashboards have been designed and created with Grafana to improve the exploration over such datasets. Fig. 17 shows an excerpt of such dashboard.

4.2. User interfaces feedback

This subsection presents the following feedback retrieved from a number of the users of the user interfaces which is worth mentioning in order to clearly validate the usefulness of the created mine-specific interfaces:

- *The mine Engineer uses the LIDAR based 3D visualizations shown in the web browser to demonstrate the installed IIoT system and get an overview of the rock stress and environmental situation at the mine During the recent period, the user has increased the number of sensors installed and expanded the use of the system. The possibility to easily adapt the visualizations and dashboards to the continuously changing situations at the mine is one of the most fascinating points of the system. The mine engineers are discussing to use the visualization also in the extraction area. The chosen solution with integrated open-source tools and the access via internet and web browser is for them comfortable and cost efficient. The already used more complex systems are not meant for a great number of users and involved the administrator of the mine company much more.*
- *Another mine uses the IIoT platform to get measurements from his preparation plant. The trend charts and the presentation of the large number of sensors in a dashboard make it easy to assess the situation in the analysed area. At the same time, the functionality of the system is displayed and documented. The risk of disturbing the production by adapting and changing the platform is eliminated by dividing control and the IIoT system. A continuous improvement of the dashboards and the algorithm to visualize the situation of the production is supported.*
- *The mining engineer uses the IIoT platform with integrated 3D LiDAR scan as a digital twin which shows in time data readings from installed sensors. Such installed sensors can give accurate status of the environmental measurements in a places with a high focus on the ventilation system. Such in time readings can optimize the ventilation system in terms of power consumption. Furthermore, with having alarms set for threshold values, it can dramatically increase safety for the working personnel. IIoT platform is easy to use and easy to update with new sensors installed. Additionally, it is quite flexible to organize according to a mining engineer or management wishes.*

5. Conclusions

This paper has presented the multi-level IIoT illuMINEation platform which is a robust multi-layer platform focused on providing support for the mining domain. The requirements from the novel platform were extracted from analysis key use cases for the mining domain like the concept of secure zones, sustainable resource extraction or advanced predictive maintenance.

Moreover, several general purpose and mine specific platforms have been analysed to clearly state the novelty of this approach in Section 2. This analysis has confirmed that there was no platform that has taken into account all the elements that we consider necessary, such as functionality, flexibility, security, performance and adaptation to the mining industry, as well as open-source and license-free.

Section 3 provides a general overview of the whole architecture where the diverse computing layers are illustrated. It describes the stack components of each layer as well as their key flows. They provide many novel key aspects for the mining domain like the long range multi-hop approach of the edge device layer which permits data re-transmission through diverse underground gateways. This is a sophisticated necessary approach inside the mines where coverage issues are common due to the heterogeneous morphology of the tunnels. Moreover, this section also presents three essential components: (1) the security mechanisms which is responsible for the identity management and the single-sing-on process, the management of users, groups and roles, the securitization of the internal and external platform communications and the audit and monitoring mechanisms. (2) the capacity to import and export data from and to third-party platforms in both modes streaming and batch. This functionality is not always found in big data platforms because several vendor-locking platforms restrict this possibility. (3) An integrated toolbox provisioned with open-source tools that enables users to prepare custom mine-specific dashboards from 3D models. Furthermore, this section finishes highlighting the scalability capability of the platform and environmental impact.

Finally, Section 4 presents a set of hypotheses, derived from the platform requirements, to validate the platform. To this end, it is detailed where each hypothesis is addressed. In addition, three successful scenarios are presented to better clarify and validate the platform functionalities. The first aims at demonstrating the suitability of the platform to move data through the diverse computing layers by managing data from intelligent rock bolts, as well as the possibility to discover programatically or graphically the data allocated on the platform to reduce data silos. Moreover, novel mine-specific user interfaces are illustrated. Subsequently, Section 4.1.2 presents a use case referred to data from the drill rig operation collected while drilling (MWD, 'Measurement While Drilling') that are used to assess the mechanical and structural characteristics of the rock mass, required e.g. for proper blast design and roof support. In this use case, the suitability of the novel user interfaces for the mining domain is highlighted. Finally, the last scenario exemplifies the process of importing data to the platform. The development of these scenarios has enable the possibility of evaluating the platform from different perspectives. At the end, Section 4 also provides the feedback retrieved from the mine-personal when interacting with the mine user interfaces to highlight its usefulness.

As a summary, we can conclude that the content of this paper demonstrates the appropriateness of the illuMINEation IIoT platform to successfully manage IIoT data from the mining domain. It has been clearly displayed that the platform is robust enough and secure to address the necessary requirements for boosting the mine digitalization.

6. Future work

As future work, the main idea is to create more mining related processing and analytic services to create a full ecosystem of utilities for the mining domain. To this end, the following specific features are planned to be integrated in the roadmap:

- A set of connectors to boost the integration with other systems like industrial ones. For instance, a connector supporting the OPC-UA protocol [57].
- Support for providing the underlying data to a mine-specific Data space ecosystem [58]. The content of the illuMINEation platform could be exposed for example through a IDSA [59] connector.
- The provision of more AI-based services to continue promoting the three original objectives. That is to say: promote secure mining operations, predictive maintenance and sustainable and intelligent mining. For example, it would be very interesting the provision of a system based on a Large Language Models (LLMs) [60] capable of guiding operators when dealing with complex tasks or facing hazardous situations.
- Inclusion of more mine-specific user interfaces (UIs) to enhance the operator duties and the provision of a template mechanism to boost the reuse of the UIs. For instance, a UI based on augmented reality could support operators when interacting with specific machines by explaining the purpose of each part.

CRediT authorship contribution statement

Raúl Miñón: Writing – review & editing, Writing – original draft, Software, Conceptualization. **Juan López-de-Armentia:** Writing – review & editing, Writing – original draft, Software, Conceptualization. **Lander Bonilla:** Writing – review & editing, Software, Conceptualization. **Aitor Brazaola:** Writing – review & editing, Writing – original draft, Software, Conceptualization. **Ibai Laña:** Writing – review & editing, Software, Conceptualization. **M. Carmen Palacios:** Writing – review & editing, Software, Conceptualization. **Szymon Mueller:** Writing – review & editing, Software, Conceptualization. **Michal Blaszczak:** Writing – review & editing, Software, Conceptualization. **Herwig Zeiner:** Writing – review & editing, Software, Conceptualization. **Julia Tschuden:** Writing – review & editing, Software, Conceptualization. **Mohammad Yusuf Quadri:** Writing – review & editing, Software, Conceptualization. **Ignasi Garcia-Milà:** Writing – review & editing, Software, Conceptualization. **Andrea Bartoli:** Writing – review & editing, Software, Conceptualization. **Norbert Gormolla:** Writing – review & editing, Software, Conceptualization. **Alberto Fernández:** Writing – review & editing, Software, Conceptualization. **Pablo Segarra:** Writing – review & editing, Software, Conceptualization. **José A. Sanchidrián:** Writing – review & editing, Software, Conceptualization. **Philipp Hartlieb:** Writing – review & editing, Software, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors are unable or have chosen not to specify which data has been used.

Acknowledgements

This work has received funding, in the context of the illuMINEation project, from the European Comission Horizon 2020 programme research and innovation program under grant agreement No 869379.

References

[1] Päivi Parviainen, Maarit Tihinen, Jukka Kääriäinen, Susanna Teppola, Tackling the digitalization challenge: how to benefit from digitalization in practice, Int. J. Inf. Syst. Project Manag. 5 (1) (2017) 63–77.
[2] Felipe Sánchez, Philipp Hartlieb, Innovation in the mining industry: Technological trends and a case study of the challenges of disruptive innovation, Mining Metallur. Explor. 37 (5) (2020) 1385–1399.
[3] SIMS Project, Sustainable intelligent mining systems for the global mining industry. SIMS European H2020 project, 2020, https://www.simsmining.eu/. (Accessed on 22 January 2024).
[4] European Union, Real-time optimization of extraction and the logistic process in highly complex geological and selective mining settings. Real-time mining european project, 2023, URL https://cordis.europa.eu/project/id/641989.
[5] European Union, Sustainable low impact mining solution for exploitation of small mineral deposits based on advanced rock blasting and environmental technologies. SLIM European project, 2023, URL https://cordis.europa.eu/project/id/730294.
[6] European Union, Sonic drilling coupled with automated mineralogy and chemistry on-line-on-mine-real-time. SOLSA European project, 2023, URL https://cordis.europa.eu/project/id/689868.
[7] Joseph M. Kahn, Randy H. Katz, Kristofer S.J. Pister, Next century challenges: mobile networking for "smart dust", in: Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking, 1999, pp. 271–278.
[8] Jingran Li, Fei Tao, Ying Cheng, Liangjin Zhao, Big data in product lifecycle management, Int. J. Adv. Manuf. Technol. 81 (2015) 667–684.
[9] Cisco, Global cloud index projects cloud traffic to represent 95 percent of total data center traffic by 2021, 2018, https://newsroom.cisco.com/c/r/newsroom/en/us/a/y2018/m02/global-cloud-index-projects-cloud-traffic-to-represent-95-percent-of-total-data-center-traffic-by-2021.html. (Last Accessed 26 October 2023).
[10] Luiz Bittencourt, Roger Immich, Rizos Sakellariou, Nelson Fonseca, Edmundo Madeira, Marilia Curado, Leandro Villas, Luiz DaSilva, Craig Lee, Omer Rana, The internet of things, fog and cloud continuum: Integration and challenges, Internet Things 3 (2018) 134–155.
[11] Flavio Cirillo, Gürkan Solmaz, Everton Luís Berz, Martin Bauer, Bin Cheng, Ernoe Kovacs, A standard-based open source IoT platform: FIWARE, IEEE Internet Things Mag. 2 (3) (2019) 12–18.
[12] Cloudera Inc., All your data. One platform. Limitless possibilities, 2021, https://www.cloudera.com/. (Last Accessed 26 October 2023).
[13] Broadcom, Pivotal Data Suite, 2024, https://network.pivotal.io/products/big-data/info. (Accessed on 22 January 2024).
[14] Microsoft, Azure hdinsight, 2021, https://azure.microsoft.com/es-es/products/hdinsight/. (Last Accessed 26 October 2023).
[15] SAP, SAP HANA cloud, 2021, https://www.sap.com/spain/products/technology-platform/hana.html. (Last Accessed 26 October 2023).
[16] Databricks, The databricks data intelligence platform, 2024, https://www.databricks.com/. (Accessed on 22 January 2024).
[17] B.V. Elasticsearch, Elastic cloud, 2024, https://www.elastic.co/es/cloud. (Accessed on 22 January 2024).
[18] Lilian Hernandez, Hung Cao, Monica Wachowicz, Implementing an edge-fog-cloud architecture for stream data management, in: 2017 IEEE Fog World Congress, FWC, IEEE, 2017, pp. 1–6.
[19] Hung Cao, Monica Wachowicz, An edge-fog-cloud architecture of streaming analytics for internet of things applications, Sensors 19 (16) (2019) 3594.
[20] Epiroc U.K. and Ireland Limited, Certiq telematics solution, 2024, https://www.epiroc.com/en-uk/products/parts-and-services/telematics/certiq. (Accessed on 22 January 2024).
[21] A.B. Sandvik, Sandvik, Maximize your productivity and profitability with our high performance mining and rock excavation products, 2024, https://www.rocktechnology.sandvik/. (Accessed on 22 January 2024).
[22] Abdul Salam, Internet of things for sustainable mining, Internet Things Sustain. Community Develop. Wirel. Commun. Sens. Syst. (2020) 243–271.
[23] Huili Zhang, Binghao Li, Mahmoud Karimi, Serkan Saydam, Mahbub Hassan, Recent advancements in IoT implementation for environmental, safety, and production monitoring in underground mines, IEEE Internet Things J. (2023).
[24] 1010data, Inc., See what you need, 2021, https://www.1010data.com/. (Last Accessed 26 October 2023).
[25] wazuh, Wazuh: The open source security platform, 2024, https://wazuh.com/. (Accessed 26 January 2024).
[26] Chenming Zhou, N. Damiano, Bruce Whisner, Missy Reyes, Industrial internet of things (iIoT) applications in underground coal mines, Mining Eng. 69 (2017) 50–56.
[27] Aaron Young, Pratt Rogers, A review of digital transformation in mining, Min Metallur. Explor. 36 (4) (2019) 683–699.
[28] Ralf Luis de Moura, Luciana de L. Farias Ceotto, Alexandre Gonzalez, Industrial IoT and advanced analytics framework: An approach for the mining industry, in: 2017 International Conference on Computational Science and Computational Intelligence, CSCI, IEEE, 2017, pp. 1308–1314.

[29] HashiCorp, Terraform, 2024, https://www.terraform.io. (Accessed 11 January 2024).
[30] Moshe Zadka, Ansible, in: DevOps in Python: Infrastructure As Python, A Press, Berkeley, CA, 2019, pp. 139–145.
[31] Docker Inc., Hykes, s. Docker swarm engine, 2024, https://docs.docker.com/engine/swarm. (Accessed 11 January 2024).
[32] Eclipse Foundation, Inc., Eclipse Mosquitto.
[33] Apache Software Foundation, Apache kafka, 2023, https://kafka.apache.org/. (Accessed on 22 January 2024).
[34] The Apache Software Foundation, HDFS architecture guide, 2024, https://www.mysql.com/. (Accessed on 22 January 2024).
[35] B.V. Elasticsearch, Elasticsearch, 2024, https://www.elastic.co/. (Accessed on 22 January 2024).
[36] PostgreSQL Global Development Group, Postgresql: The world's most advanced open source relational database, 2024, https://www.postgresql.org/. (Accessed on 22 January 2024).
[37] The Apache Software Foundation, Apache spark, 2018, https://spark.apache.org/. (Accessed on 22 January 2024).
[38] Tiangolo, FastAPI, 2024, https://fastapi.tiangolo.com/. (Accessed on 26 January 2024).
[39] LinkedIn, Linkedin datahub, 2024, https://datahubproject.io/. (Accessed on 26 January 2024).
[40] B.V. Elasticsearch, Kibana, 2024, https://www.elastic.co/es/kibana. (Accessed on 22 January 2024).
[41] Grafana Labs, Grafana, 2024, https://grafana.com/. (Accessed on 22 January 2024).
[42] Keycloak Authors, Keycloak, 2023, https://www.keycloak.org/. (Accessed on 22 January 2024).
[43] Confluent, Inc., Schema registry overview, 2024, https://docs.confluent.io/platform/current/schema-registry/index.html. (Accessed on 22 January 2024).
[44] The GraphQL Foundation, Graphql, 2024, https://graphql.org/. (Accessed on 26 January 2024).
[45] ENISA, Good practices for security of internet of things in the context of smart manufacturing, 2018, https://www.enisa.europa.eu/publications/good-practices-for-security-of-iot. (Accessed 07 June 2024).
[46] Cloud Security Alliance (CSA), CSA IoT security controls framework v2, 2021, https://cloudsecurityalliance.org/artifacts/csa-iot-security-controls-framework-v2/. (Accessed 07 June 2024).
[47] Internet Engineering Task Force (IETF), RFC 8446: The transport layer security (TLS) protocol version 1.3, 2018, https://www.rfc-editor.org/rfc/rfc8446.html. (Accessed on 22 January 2024).
[48] Geomotion Australia, LS-G6-KIO-GW868 4G rugged gateway, 2021, https://www.geomotion.com.au/wp-content/uploads/2021/04/geomotion_loadsensing_datasheet_4g_gateway.pdf. (Accessed on 22 January 2024).
[49] Internet Engineering Task Force (IETF), RFC 8376: Low-Power Wide Area network (LPWAN) overview, 2018, https://www.rfc-editor.org/rfc/rfc8376. (Accessed on 22 January 2024).
[50] Loadsensing, LS-G6-INC15 wireless tiltmeter, 2019, https://lora-alliance.org/wp-content/uploads/2019/09/Copy-of-LS-G6-Tiltmeter-Datasheet-v3.pdf. (Accessed on 22 January 2024).
[51] The MITRE Corporation, MITRE ATTCK enterprise matrix, 2023, https://attack.mitre.org/matrices/enterprise/. (Accessed on 22 January 2024).
[52] Chaitali Choudhary, Inder Singh, Manoj Kumar, A real-time fault tolerant and scalable recommender system design based on kafka, in: 2022 IEEE 7th International Conference for Convergence in Technology, I2CT, IEEE, 2022, pp. 1–6.
[53] Marian Ileana, Maria Ioana Oproiu, Constantin Viorel Marian, Using docker swarm to improve performance in distributed web systems, in: 2024 International Conference on Development and Application Systems, DAS, IEEE, 2024, pp. 1–6.
[54] Michael Nöger, Philipp Hartlieb, Michel Varelija, Marcin Paterek, Malwina Kobylanska, Maciej Pikula, Bartolomiej Bursa, Morgan Rody, Rafael Mosberger, Varelija Nöger, Bartlomiej Bursa, The safe zone concept: Improved safety and environmental performance, 2023.
[55] Michael Nöger, Philipp Hartlieb, P. Moser, T. Griesser, T. Ladinig, Dominik Dendl, The potential of a mine-wide digital rock mass condition monitoring system, in: Proceedings of the 5th International Future Mining Conference, Perth, Australia and Online, 2021, pp. 6–8.
[56] Michel Varelija, Philipp Hartlieb, Designing intelligent rock support systems to detect gravity-driven wedges, Rock Mech. Rock Eng. (2024).
[57] OPC Foundation, OPC foundation, 2024, https://opcfoundation.org/. (Accessed 10 June 2024).
[58] Boris Otto, The evolution of data spaces, in: Designing Data Spaces: The Ecosystem Approach To Competitive Advantage, Springer International Publishing Cham, 2022, pp. 3–15.
[59] International Data Spaces, International DATA spaces. The future of the data economy is here, 2024, https://internationaldataspaces.org/. (Accessed 10 June 2024).
[60] Yupeng Chang, Xu Wang, Jindong Wang, Yuan Wu, Linyi Yang, Kaijie Zhu, Hao Chen, Xiaoyuan Yi, Cunxiang Wang, Yidong Wang, et al., A survey on evaluation of large language models, ACM Trans. Intell. Syst. Technol. 15 (3) (2024) 1–45.

Raul Miñón finished the MEng. in Computer Science in 2008 at UPV/EHU. He worked on a consultancy creating web projects for two years in Bilbao. In 2015, he finished a Ph.D. at the Egokituz laboratory having done a research stay collaboration in the CNR (Italy). His thesis research topic was the automatic generation of accessible UIs considering people needs. Subsequently, he founded a company related to his thesis. Next, he worked two years in IK4-Ikerlan in the Big Data team. Currently, he researches and develops projects in the areas of Big data, DevOps, Edge computing and MLOps in Tecnalia

Juan López de Armentia Mendizabal is with Tecnalia since 2018, where he participates in Big Data, Cloud, Edge and Artificial Intelligence operationalization related projects. In addition, he has experience in tasks related to visualization techniques for big data volumes oriented to efficient monitoring of analytical processes and based on time series data. He finished his studies in Computer Engineering in 2008 at the University of Mondragon. During 2009-2011 he worked at Tecnalia Research & Innovation and completed a Master in Control, Automation and Robotics Engineering at the University of the Basque Country. In 2011 he joined DeustoTech-Deusto Institute of Technology of the Faculty of Engineering at the University of Deusto where he obtained his Ph.D. in Computer and Telecommunication Engineering in 2016.

Dr. Philipp Hartlieb is a senior scientist at Montanuniversitaet Leoben, Austria. He holds a M.Sc. in Applied Geosciences and did his Ph.D. in Mining. His research focusses on rock fragmentation and excavation technologies, including drilling&blasting, mechanical excavation and alternative excavation methods, as well as the efficient acquisition, use and analysis of data in the mining process. This leads to a holistic picture of the entire process, quantifying the effects of upstream processes (e.g blasting) to downstream (run of mine).