Enterprise ICT Architecture

Elk River

Authors:

Albertini - Boero - Faraone - Lafronza - Ranieri

ICT Enterprise Architecture

Elk River

Index:

1. Executive Summary

People and ecosystems require an adequate quality as well as a steady equilibrium of the environment that surrounds them in order to live and prosper. Historically, many ecosystems on Earth have suffered severely because of human activities and have changed drastically, or even permanently, because of poor pollution management and failures to understand the severity of the problem, also due to the lack of information, by policymakers. Luckily, nowadays, thanks to the huge progress that have been made both in the field of database management systems as well as in the field of IoT systems, policymakers and environmental protection agency have a powerful tool to measure constantly hundreds of parameters and act accordingly.

The Elk River is a large geographical area in the United States, known for its distinctive flora and fauna. This report focuses on describing the problems that afflict this geographical area, with a special focus on the three most important elements of a riverside environment: water, air and soil (with a focus on noise levels). Following these considerations, a proposal for an information system (IS) that deploys different type of sensors has been presented. The data will be gathered by the sensors and stored in a database, which can be accessed by the policy makers in order to help them make informed decisions about corrective measures that might be needed for the area.

2. The Elk River Valley

A state agency wants to improve the Elk RIver geographical region

This paper aims to analyze a public administration scenario of a state agency responsible for the oversight and administration of a large geographical area located in the Appalachian region of the Southern United States. The Elk River region is a rather distinctive area for its territory properties as well as for the industries that have their operations in the region. It is useful therefore to describe all the defining characteristics of the area and its history, to better understand the proposal of the new information system solution, which is intended to improve the health of the territory through constant monitoring and corrective actions.

2.1. The Elk River Valley

The Elk River is a 277-kilometers long affluent of the Kanawha River, which drains about 4.000 square kilometers in central West Virginia (WV) in the United States and is part of the watershed of the Mississippi River. The river rises in western Pocahontas County and flows northward and then eastward across the state of West Virginia, where it finally empties into the Kanawha River near Charleston (WV), the capital and largest city of the state. Some sources claim that the river owes its name to the river's original Native Americans name "Tis-chil-waugh" (meaning "Plenty of fat elks"), although consensus has been reached around the explanation that today's English name is simply owed to the numerous herds of wild elk that used to roam the stream's bank in pioneer days (Kenny, 1945).

Elk River is a great resource for the inhabitants and it needs to be preserved

Although very secluded from the rest of the nation, the Elk River has proven to be a great resource for the people that settled in its valley. In the late 1800s and early 1900s, the water stream proved to be the preferred route to transport logs with large rafts out of the forests of central and eastern West Virginia to the paper mills. The river lost its importance in this respect after the advent of railroads, such as the West Virginia Coal & Coke Railroad and the Baltimore & Ohio Railroad, but remained an important asset for its territory (West Virginia Explorer, 2019). The Elk River serves as the source of water for 1500 miles of pipelines that carry its water to customers in central and southwestern West Virginia (Botelho, 2014).

Elk River is the only place on earth where you can find some species

The 'Flood Control Act' of 1938 by the US Congress authorized the construction of the Sutton Dam just upstream of Sutton (WV), a small town settled by Elk River 163 kilometers above the mouth of the river. The dam, which was built primarily for flood control on the Elk, Kanawha and Ohio Rivers, gave form to the Sutton Lake, a 23-kilometers long water reservoir surrounded by over 10,000 acres of public land. The absence of residential homes and commercial activities on the lake grant the visitors a pleasant experience surrounded by nature (Sutton Lake Marina, 2019).

According to the Elk River Wildlife Management Area, the river, the lake, and the valley are host to a lot of different species of animals and plants. Even though elks were all exterminated during the 19th century, many mammals such as deers, groundhogs, foxes, and bobcat are still present on the territory. Tens of different varieties of fish make their home in the river, accommodating fishing for bass, trout, bluegill, catfish, walleye, and muskellunge. The Elk River is also the only place on earth where you can find the cave crayfish (Cambarus nerterius), as this species is endemic to the state of West Virginia, and more specifically to the caves of the Elk River (Zachary, 2010) (Figure 1).

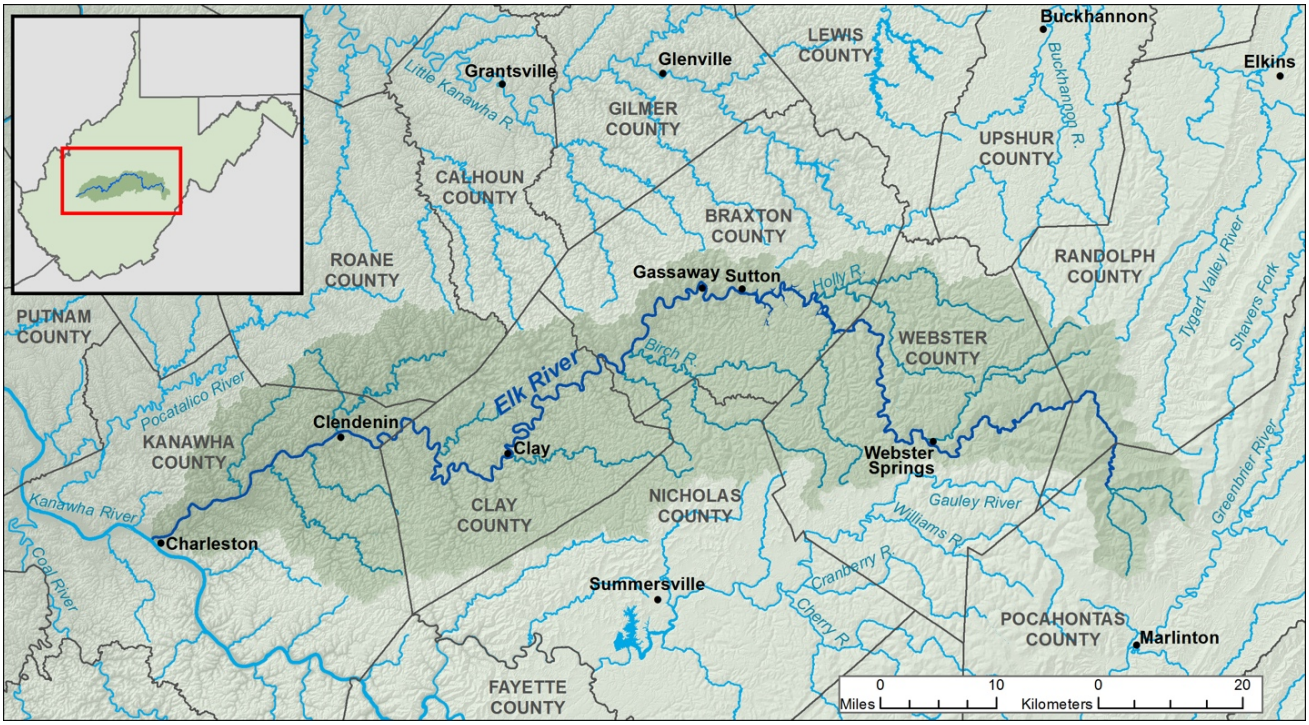


Figure 1: a map of the Elk River valley (river flowing from the right to the left of the map). (Wikimedia Commons, 2020)

2.2. Ancillary activities near the territory

Charleston hosts a number of historical buildings

In the territory of Elk River, there are many possible attractions to tourists. One of the most popular is the state capital, Charleston, the most populous city of West Virginia located at the confluence of Elk and Kanawha River. The city has a high number of old buildings that can refer to an enormous variety of architectural styles, such as the Charleston Baptist Temple or the Kanawha County Courthouse. Moreover, at least fifty places in Charleston are included in the National Register of Historic Places (NRHP), a United States federal government list that holds places, districts, sites or buildings that are worthy of being preserved due to their historical importance.

Besides the historical monuments, there are also important outdoor areas near Elk River. One of the most famous is the Kanawha State Forest, 11 kilometers away from Charleston, which is about 38 square kilometers. Even if it is classified as "State Forest" West Virginia decided to manage it as a "State Park". It is famous for the many outdoor activities practiced in there, such as camping, picnicking and hiking. There are also tourist-dedicated structures such as a playground, swimming pool, public shooting-range, and cross-country skiing structures. It is also famous to be a place with an extremely high concentration of "geocaches", objects used to play a GPS game called "Geocaching". Another outdoor attraction is the Magic Island, an island (that is now connected to the mainland) that now is used as a public park. The island is so-called because once the rise and fall of the river used to submerge it and it felt like it was disappeared, as if by "magic".

Elk River is a very fishy river and destination of many professional fishers

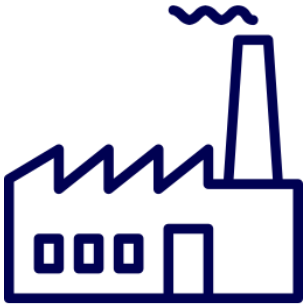
One of the most practiced activities in the Elk River area is the so-called "Fly Fishing", a method of angling that employs a long rod and an arbor reel that holds a heavy line joined to a lighter nylon leader. The lures are artificial flies made of hair, feathers or synthetic material designed to imitate the shape of a fish. The upper part of Elk River is considered a top destination for the professional fly fishers. It has a big population of brown trouts and rainbow trouts.

2.3. Industries operating in the territory

Charleston, the state capital, but in general the whole Elk River valley, lies in an industrial area that takes flinty pride in the nickname Chemical Valley. After decades of slow decline, the local industry has revived in recent years, owing to the boom in cheap natural gas, which has made America one of the world's most inexpensive places to make chemicals.

Main sectors are:

1. **Chemicals**: West Virginia is home to about 140 different chemical-related companies that provide over 12,800 jobs to the state. This makes West Virginia the location of one of the largest amount of chemical manufacturing companies in the world. Nearly a quarter of the state's international exports consists of chemicals and polymers.



The valley hosts many chemical, biotech, and energy industries

2. **Biotech**: Home to industry leaders like Dow, DuPont, Bayer, PPG, BASF an Proviron, West Virginia contains one of the largest concentrations of biotech industries. The main reason for West Virginia's success in this area is due to the fact that there are major research universities located in the state, which attract a wealth of resources and partnerships.

3. **Energy**: West Virginia is one of the primary regions in the U.S. for coal mining, according to Britannica. However, they indicate that the industry is being threatened due to environmental concerns about surface mining, known as mountaintop removal. Nevertheless, West Virginia produces about one-tenth of the amount of coal in the country. West Virginia is the top interstate electricity exporter in the country.

2.4. Description of the incident

On January 9th, 2014, Freedom Industries reported that one of their chemical storage tanks has been leaking for an unknown amount of time. The total amount of spilled chemical substance was more than 28.000 liters (Bernstein and Lenny, 2014). The chemical that leaked in the river was crude MCHM (4-Methylcyclohexanemethanol), an organic compound that is used by the coal industries to clean coal from impurities that increase the volume of pollutants

produced during combustion. Another use of this chemical agent is the cleaning of the coal plants to delete any residual impurity.

On January 13, 2014 (five days after the spill), the Governor of West Virginia declared immediately the state of emergency and people were ordered not to use tap water in their houses. The automated phone calls warning didn't work well, so not all the people were informed about this issue by the authorities and many of them received 'second hand' information by family members or friends. This issue affected more than 300'000 citizens since the affected area was very wide (nine counties) (Botelho, Greg, et al, 2014) and between the 9th and the 14th of January more than 150 people visited the local hospitals presenting symptoms of nausea and vomit, while more than 700 people called the 'West Virginia's Poison Control Center' reporting many different symptoms including nausea and rashes.

A chemical leak in 2014 polluted the river and the water supply

On January 21, 2014, the company found out that from the holed tank leaked out even some other mixture of substances that included propylene glycol phenyl ether (PPH) and dipropylene glycol phenyl ether (DiPPH). This mixture polluted not only the Elk River but also the drinkable water supply. The inappropriate reporting of this information to the population was considered a violation of the State law. The water supplier didn't declare the water drinkable until February 24.

Freedom Industries has declared bankruptcy because of the inability to cope with the high number of lawsuits filed against them. In fact, more than 2.000 people claimed against the business and reached a 2.5 million dollars settlement (Financier Worldwide, 2014).

3. Environmental Monitoring Entity

Elk River runs through various counties in the state of West Virginia, this means that the jurisdiction about its conditions and the measures adopted must be taken by mutual agreement. The goal of this paper is to simulate a state agency, created and founded by all those counties together, that has the authority to constantly measure and monitor the state and the pollution level of the waters of the river. Moreover, since the final objective of the counties is to valorize the

entire area of the river in order to make it more attractive for tourists, industries and citizens, it is important not only to monitor the waters, but also the conditions of three important elements, air, soil and noise. There are motivations behind the selection of those elements, even if they are all related to the same objective:

The regulation of Elk River is a matter of collaboration between different entities

- **Water**: The area of Elk River is well known in West Virginia to be one of the best areas to practice angling, especially with the technique called "Fly Fishing". So, it is important to monitor the nutrients present in the water, the overall fish population of the river and the pollution of the water. In fact, since the event occurred in 2014 the population is very sensitive about the conditions of the river, especially because it is the main water resource for the entire area and if it is polluted, the entire citizenship will suffer the consequences. It is also important to consider that Elk River is the only place in the world where it is possible to find a particular fish, the Cambarus Elkensis, also called the "Elk River crayfish".

- **Air**: Elk River's area is also famous to be a very productive area of the entire country, in particular the coal industry is one of the most active in the area. In order to stimulate the tourism (especially the one related to the camping near the banks of the river) it is important to constantly monitor the air pollution. Also, it is important to monitor other aspect that are unrelated to the pollution, such as weather conditions and air pressure that can be important, especially for citizens and tourist, in order to organize trips and day offs near the river.

The agency will need to monitor all four identified elements

- **Soil**: In order to re-evaluate an entire area not only is important to stimulate the tourism and the attractivity for visitors, but also to encourage more people to invest in the area. The event of 2014 cannot be considered only related to water pollution, but it has also affected the conditions of the soil. The area where the river flows in West Virginia is a big valley where agriculture should play an important role for the re-evaluation of the territory. Monitoring the conditions of the nutrients and the toxic agents that can be founded in the soil can be

fundamental for the counties in order to discover and indicate the best areas that could be dedicated to agriculture and breeding.

- **Noise**: The last element that is important to monitor is the noise pollutions. In fact, in proximity of the banks of the river are situated a lot of important cities of the state, among which it is possible to find Charleston, the state capitol city. Constantly monitoring the noise pollution, the time of the days where there are peaks and the most polluted areas is important in order to introduce new regulations in order to improve the livability in the cities. Also, since most of the industries are also situated near the river it is important to monitor noise pollution in order to find the most silent areas where is possible to create camping zones. Even if the pollutions of the cities can be considered directly on the cities jurisdiction can be useful to delegate the responsibility to the state agency in order to not create different systems for the measurement that would increase the costs and mine the quality of the data collected.

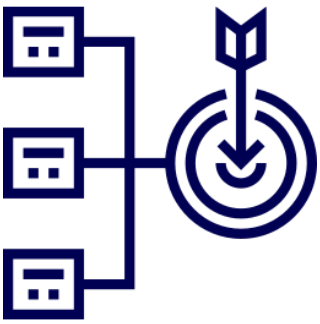
Data collecting will enable a better understanding of the environment's status

Collecting data from all those different sources, that at the same time are interrelated, should give to the state agency the possibility to monitor and have a better understanding of what are the main problems of the area. This is the first step in order to try to re-evaluate the entire valley.

3.1 Business Process

As stated before, the first objective of the agency is to use sensors in order to have important measurements. Even if the measurement is only a mean used by the entity in order to reach its final objective it is one of the most important part. In fact, there is a huge amount of data coming from the collection on a vast territory that have to be analyzed and stored. So, the main focus of this paper will be on the measurement and the storage of all data.

After that, it is also necessary to elaborate it in order to retrieve some important information that have to be used by the agency and the counties to improve the conditions of the territory and the livability of the whole area. Since the operations that this entity has to perform are various and involve many third



A well defined Business Process is paramount to meet the goals

parties it is necessary to have a well calculated Business Process in order to provide good information and policies. Business Process Management is essential to describe how everything in the agency should be done and performed in order to reach the final objective. Since a Business Process is a collection of inter-related events, activities and decision points it is fundamental to create a model that make all these interactions work correctly and in harmony to improve the overall performances of the system. In this paper the Business Process Management techniques will be used mostly to give the reader a general overview about the structure (both informatic and not informatic) of the society in order to describe the behavior of the entity in correlation with all the third parties with which it will work. Another goal that it is important to reach for the entity is the so-called "Process Automation", the ability to solve most of the tasks required without the necessity of the manual intervention of a human being but only with the interaction between different computers.

As it was stated above third parties are really important for the functioning of the entity for two main reasons:

- **Cost reduction**: The analysis of data coming from all the different measurements it is very time expensive and requires the efforts of an entire team that has that as its only goal. Building an entire department for the agency only for the analysis and the production of results will be very expensive and not as efficient as paying specialized laboratories or companies to do the analysis based on the data collected by the sensors.
- **Improvement**: Creating an ad-hoc team that has as its only goal to analyze the data collected by the sensors sometimes can be considered a good solution. However, entrusting the operation to a company that works with data analysis benefits of the learning economies and of the more means that the third part can give to its workers.

Data analysis will be outsourced to a specialized company

Generally, the only important downsize that can be considered of relying on a third-party society is that it could compromise the security of the data that are being transmitted. However, the data that the sensors measure cannot be considered as sensitive data and the security of the transmission has a lower importance compared to the necessity to have a high-quality analysis.

The last part of the process of the agency is the policy making. In order to do that policy makers will use an Executive Information System to produce reports that will be used to take the final decisions. However, these files are useful not only for taking decisions but also to monitor, using also the stored data, the effect of the policies. So, it is necessary to have a database that stores all the history of reports that are created by the information system that will be used for further comparisons. [N.B. The structure of this second database will be simple because all the checks on previous detections will be done on the relational database and this one will be used only as a secondary source of data that stores the official report produced].

Data will be used by policy makers to regulate the operations in the area

The policy making part of the process has to involve the human work because policy decisions cannot be taken by a computer, but it is necessary to have experts to operate the regulations. Moreover, it is also necessary for the agency to communicate with other third-party entities, in this case directly related to the counties, in order to communicate them the changes on the policies and the way they have to apply the new changes (e.g. the agency creates a new policy related to maximum levels of pollutions that a single industry can produce everyday has to be communicated to local authorities that have to intervene if the new standards are violated).

The agency will also give the possibility to other users, such as citizens or other interested people, to access the data if they want to have information about the status of the various areas of the Elk River valley. This functionality is useful in order to improve the participation of all the citizens to the activities of the county and their awareness about the actions and the results that the agency is doing in order to improve the livability of the area.

3.2 Business Process Model

In order to describe better the Business Process of the agency in this paper it is being utilized the Business Process Model and Notation (BPMN). It is a standard utilized by the Object Management Group (OMG), supported by numerous tools, that provides a graphical notation for specifying business processes. The graphic flowchart of the model, where the starting point is the collection of the data made by sensors and its transmission, will be proposed below (Figure 2).

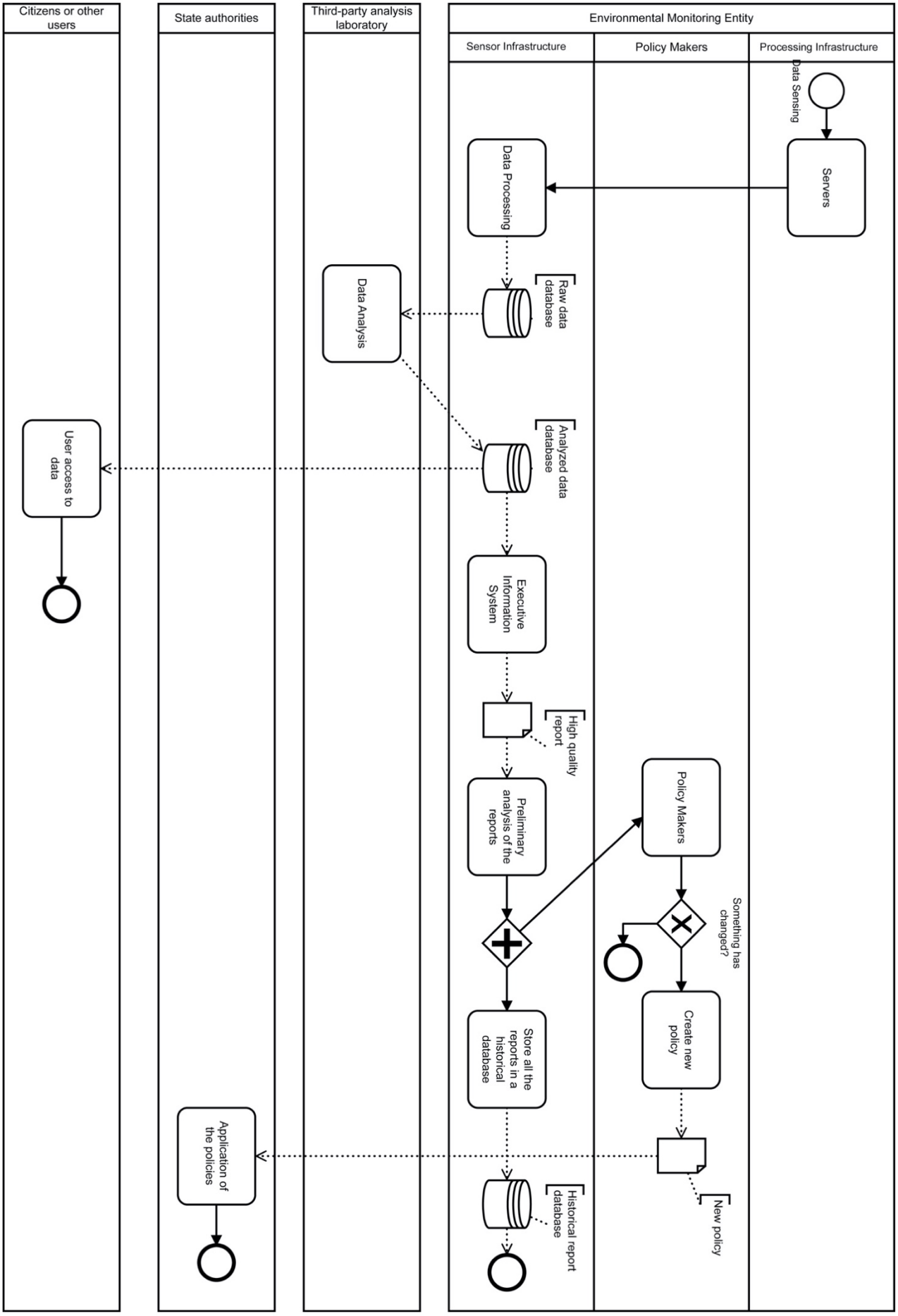


Figure 2: Visual representation of Business Process of the entity

4. Detection of pollutants

Different type of pollutants and their sources will be monitored by sensors

Given the morphology of the Elk River region and the objectives described in the previous paragraphs, a deeper analysis of the three main elements of the environment was deemed necessary. To build a more complete potential scheme, the counties should monitor the data related to the three main components of the Elk River's environment: water, soil, and air (with a special focus on noise levels). For each one of the elements, a series of potential mean of pollution and their sources will be identified, in order to better understand the impact that these might have on the everyday life of people, animals, and plants on the territory. Moreover, for each element, the most efficient sensor system, together with an efficient layout of the sensors, will be identified. Each sensor system must be deployed to record and then analyze the data, helping the policy makers address, in the most efficient and effective way, the different issues.

4.1 Water

Water is a very critical element in the analysis of the pollution of an environment. It has relevant effects on the health and lifestyle of a population, there are different kind of water pollution (Figure 3):

Water pollution is caused by many different sources

- Industrial, due to the release of production refuses by the industries.
- Urban, caused by human production of feces that flow in the sewerage.
- Agricultural, it derives from the use of fertilizers and pesticides in considerable quantities, and from the spreading of sewage from farms.
- Caused by hydrocarbons, caused by disasters such as the sinking of oil tankers.
- Thermal, linked to the industrial pollution, derives from the excessive use of water to cool industrial plants, especially in thermal power plants.

Starting from these different sources of pollution, considering the environment where the Elk River is located, the more relevant pollution agents are industrial refuses and thermal augmentation.

INDUSTRIAL POLLUTION

The factories that operate in the valley are one of the main contributors

Chemical industries, paper mills and sawmills are responsible for water pollutions. Companies belonging to these sectors produce refuses containing nitric acid, soda, phosphoric acid, ammonia, sulfuric acid, hydrochloric acid (Laws, 2018). All these substances can decrease the solubility of oxygen in water and have a relevant impact on the PH, causing pathological changes of disappearance of some living species or the development of unusual ones. Under normal conditions, the water is able to self-purify thanks to a certain amount of dissolved oxygen which transforms the substances, thanks to the aerobic decomposition (oxidation), in non-polluting compounds (such as carbon dioxide, nitrates, phosphates, sulphates).

In this category of pollution source flow even heavy metals (Quicksilver, Chromium and Lead) and heavy metals ions (such as Cr6+, Hg2+, Cd2+ and Cu2+), produced by metallurgical industries, that have big impacts not only on the water ecosystem but even on the human health. In theory every producer of this kind of chemicals should filter them before releasing them in the rivers and atmosphere.

Substances like these can have big impact on human life since they can contaminate water and food (fishes that live in contaminated environments and other animals that eat grass grew up using contaminated water). Heavy metals and their derivates block the catalytic action of the body's enzymes causing poisoning or death, both for animals and for humans.

THERMAL POLLUTION

The factories' cooling procedures are raising the water temperature

Thermal pollution is caused by the cooling procedures of the industrial plants. Elevated temperatures of the water can harm the everyday life in the water environment. Like the presence of some substances does, the increasing of water temperature make the self-depuration cycle of the water harder. A higher temperature of the water increases the metabolic rate, such as the enzymes activities, in aquatic animals making them consume more food (Goel, 2006). As result the food chain is compromised.

In addition, when a power plant first opens or shuts down for repair or other causes, fish and other organisms adapted to particular temperature range can be killed by the abrupt change in water temperature, either an increase or decrease, known as "thermal shock" (Laws , 2000).

Each zone of the river will have a precise number of water sensors

SENSORS

Elk river length is 172 miles, in order to understand which part of the river is more polluted, sensors should be grouped in 4 zones (each zone corresponding to 43 miles of the river's flow). On the average there should be 10 sensors for each zone, but since the first zone (Z1) is the fount one and few activities are developed there, it will be more useful to monitor Z1 with less sensors than the zones where the concentration of industries is higher.

Given the morphological structure of the territory and the distribution of the industries in the different areas a possible schema of the sensor is the following (Figure 3):

- Zone 1 → 5 sensors
- Zone 2 → 10 sensors
- Zone 3 → 11 sensors
- Zone 4 → 14 sensors

In order to monitor the water quality, the sensors that are required have to collect data about the PH, the concentration of heavy metals and the temperature of the water.

4.2 Soil

Soil pollution is caused mainly by the incorrect disposal of waste

Soil is the upper layer of earth in which plants grow, a black or dark brown material typically consisting of a mixture of organic remains, clay, and rock particles; its pollution is caused by xenobiotics, chemical substances that are not naturally produced but instead are the final output of human activities.

There are three main sources of soil contamination:

- Industrial activities, linked to the waste disposal in the rivers
- Agricultural chemicals, used by the agriculture industries that to kill pests and weeds (on pavements and railways)
- Improper disposal of waste coming from the humans' normal life activities and the incorrect disposal of them.

Elk river's valley is famous for its industrial activities, more specifically the chemicals and coal mining ones, that's why they represent the main sources of pollution that a coherent monitoring system must consider .

INDUSTRIAL POLLUTION

Industrial activities are able to generate effects on the environment such as acid rains that affect the soil quality.

The main sources of pollutions are the coal ashes generated by the combustion of coal. In the early 60s when coal was the main source of energy its ashes were spread in the ground producing huge pollution effects. In fact, coal is reach of lead and zinc and other heavy metals that in the ashes result to be more concentrated. As it is shown in the previous paragraph heavy metals in the ground have possible devastating effects on the health of the humans and of the animals.

Acid rain can have a big impact on soil pollution

Another important factor of pollution is represented by chemicals such as petroleum hydrocarbons and polynuclear aromatic hydrocarbons (such as naphthalene). These chemical agents are able to cause advanced and remarkable soil erosion.

SENSORS

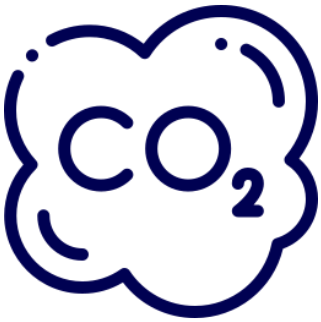
Soil pollution sensors must collect data concerning the concentration of heavy metals and the concentration of petroleum hydrocarbons in specific and relevant sites. As the previous sensors they will be distributed along Elk rivers sides divided in four zones, to better understand the level of pollution of each zone. In order to have an economical save, soil sensors will be implemented in the same structures of the air sensors.

The collection of soil data will be less frequent than the one that is made on the water and on the air, since the composition of the soil is less frequently subject to significant changes.

4.3 Air

Air pollution occurs when harmful or excessive quantities of substances including gases, particulates (both organic and inorganic), and biological

molecules are introduced into Earth's atmosphere. This kind of pollution is caused by:



Air pollution is one of the biggest threat for humans and needs to be monitored

- **Natural source**s, such as volcanic eruptions, fires and biological processes.
- **Anthropic sources**, such as vehicular traffic, domestic heating, industries and craft activities, off road vehicles (trains, tractors, quarry vehicles etc.), agriculture and other activities.

Due to the specific environment where Elk river is placed, the main pollution sources are the anthropic ones.

ANTHROPIC POLLUTION

The anthropic pollution is the pollution caused by human activities. The main families are vehicular traffic, home heating and industrial activities.

All these three families produce different pollutants depending on the kind of flue it is used by the citizens or during the industrial processes. The fossil fuels' combustion produces carbon dioxide (CO2) while diesel fuels produces different substances such as Carbon monoxide (CO), hydrocarbons (HC), nitrogen oxides (NOx) and sulphur dioxide (SO2).

Carbon dioxide is called "the leading pollutant" because of its role in greenhouse effect. It is a natural component of the atmosphere (Essick, 2019).

Industrial activities are not the only source for air pollution

Carbon monoxide is a colourless, odourless, toxic gas. It creates a smog type formation in the air that has been linked to many lung diseases and disruptions to the natural environment and animals (NHS, 2017)

Nitrogen oxides can be seen as a brown haze dome above or a plume downwind of cities

Sulphur dioxide is considered a pollutant due to the effects that its further oxidations can cause. In fact, its transformations produce the acid rains, that not only pollute the air but even soil and water.

SENSORS

Air sensors must collect data about Carbon dioxide, Carbon monoxide, Nitrogen oxides and Sulphur dioxide. In addition, they must be able to detect the environment temperature that can be influenced by the CO2 action.

These sensors should be placed relatively far, at least one mile away, from the Elk River since the presence of the river can influence the data such as the detected temperature.

As the previous sensor they should be divided into four zones, following the metrics used for the water sensors (Figure 3):

- Zone 1 → 5 sensors
- Zone 2 → 10 sensors
- Zone 3 → 11 sensors
- Zone 4 → 14 sensors

4.3 Noise

Noise is a source of pollution caused by human activities such as industrial production and vehicular traffic.

High noise levels can seriously affect the quality of people's life

Excessive noise seriously harms human health and interferes with people's daily activities at school, at work, at home and during leisure time. It can disturb sleep, cause cardiovascular and psychophysiological effects, reduce performance and provoke annoyance responses and changes in social behavior. Moreover, noise pollution has a severe impact on the living condition of the local fauna and could harm the entire ecosystem of the valley. It is really important to monitor the noise pollution both in the cities, in order to protect health of the citizens, and near the industries, since they are often situated near rural areas where most of the fauna lives (World Health Organization, 2018).

SENSORS

Noise sensors must collect data about the level of noise in a certain site. Noise sensors will be placed in the same location of the air and soil sensors. In order to make the deployment the noise sensors economically convenient, a good strategy is the implementation of few high-sensitive sensors, and many low-sensitive sensors. With these scheme, high-precision sensors serve to calibrate the less expensive ones (Cui, Yaokui, et al, 2018).

A possible proposal for the different zones is the one that follows (Figure 3):

- Zone 1 → 1 high-sensitive sensor and 4 low sensitive sensors
- Zone 2 → 3 high-sensitive sensors and 7low sensitive sensors
- Zone 3 → 4 high-sensitive sensors and 7 low sensitive sensors
- Zone 4 → 5 high-sensitive sensors and 9 low sensitive sensors

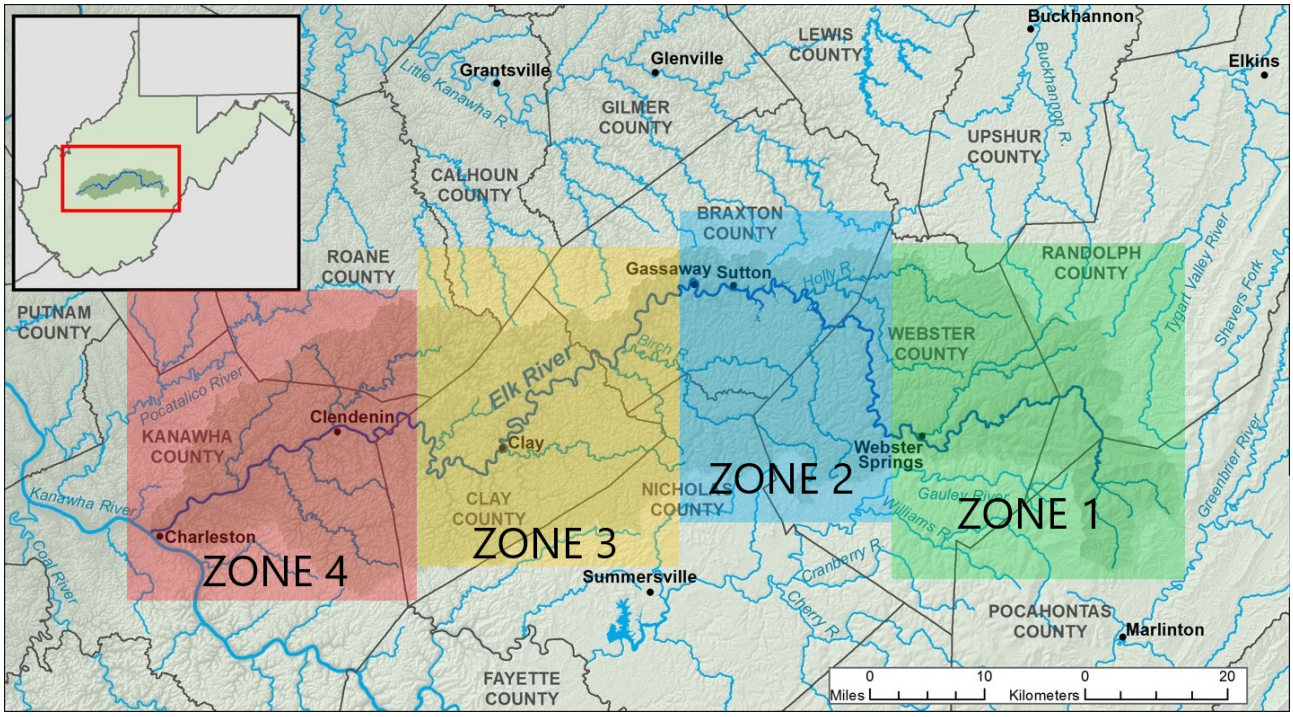


Figure 3: Zoning of the Elk River valley. (Wikimedia Commons)

5. Connection (data transmission)

All the data must be transmitted to the main server located in Charleston, the capitol city of West Virginia. The city is located at the end of the course of the river, where its water mixes with the ones of the Kanawha river. To make data available in an efficient and effective way a fast connection is essential. However, the most important characteristic that the connection must have is a wide area covering. This is fundamental because the distance between the main server and the furthest sensor is high and this could influence the quality of the data transmission. A good connection to accomplish this task is Narrowband. This technology is a radio technology that is capable to cover a wide range of cellular devices and services (Ublox, 2017).

The data will be transmitted from the sensors to the main server

It has low power consumption, excellent extended range in buildings and underground, easy deployment into existing cellular network architecture, network security & reliability, lower component cost.

The main server in Charleston is located in the main building of the Environmental Monitoring Entity where all the data is stored and then immediately transmitted, using the same technology, to the third-party entity that has to perform further computation on those data. Then, the analyzed data is transmitted again to the second main server. After this passage all the further computation and data transmission is done utilizing a LAN connection within the entity.

The last data transmission required is the one that has the task to transmit the new policies to the local authorities. However, the focus on this paper is not on this type of transmission that do not require further explanations.

6. Database structure

A relational database will answer all our needs for this specific project

To achieve the desired results, we decided to propose a relational database. We need a system that is able to store hydrological observations in order to enable the possibility of analyzing data in an effective manner. We implemented the model proposed by CUASHI Hydrologic Information System, known as Observations Data Model (ODM) that was created to store and analyze information gathered by different sensors, investigators, and sources. Its aim is to make easier the data analysis creating a common format to collect data. The ODM, around the data values, store a lot of ancillary information about the observations to allow the possibility of creating a traceable heritage from raw measurements to usable information unambiguously interpreted and used.

When you have to deal with this type of observations there are three main variables that must be considered:

- **Space**: the location in which the observation is made.
- **Time**: the date and time of the observation.
- **Variable**: the type of the observed variable (e.g. water's PH, heavy metals, temperature, hydrocarbon, etc.).

In addition to these dimensions, there are a lot of other distinguishing attribute that can be added to give more information to each observation. So, the main

Metadata are essential to contextualize the values registered by the sensors

decisions that we have to make in order to implement ODM involve the quantity of ancillary information to be added and the decision about the place in which store this information, same table or different tables with the use of foreign keys. An observation is an event that results in a value describing some phenomenon (Open Geospatial Consortium, Inc., Observations and Measurements, OGC Best Practices Document OGC 05–087r4, version 0.14.7, available at http://portal.opengeospatial.org/files/?artifact_id=17038). These values are not self-describing and we need to contextualize them using other relevant information, or metadata. Metadata is the descriptive information about data that explains the measurement attributes, their names, units, precision, accuracy, and data layout, as well as the data lineage describing how the data was measured, acquired, or computed [Gray et al.,2005].

6.1. ODM Logical Data Model

The logical data model for ODM is shown in Figure 3. At the center of the chart is located the table named **Data Values**. This table stores the numeric values of the observations and all the links (foreign keys) to all the other relevant attributes. To avoid the condition of data redundancy most of the attributes are stored in different tables. So, in the figure below, for each relationship is indicated the name and the directionality in order to simplify the understanding. **Data Values** contains the unique identifier, that is also the primary key, for each data value (*ValueID*), the numeric value of the observation (*DataValue*), and the unique identifier of the variable observed (*VariableID*). *ValueAccuracy* is a numeric value that indicates the measurement accuracy of the observation. Then, this central table stores information as the reference ID to the source table (*SourceID*) and about the time and the location of the observation (*LocalDateTime, UTCOffset, DateTimeUTC, SiteID*), about the method used (*MethodID*) and eventual sample (*SampleID*), the information on the quality of the measurement (*QualifierID, QualityControlLevelID*), the value of the offset (distance from a control point to the point at which data value is observed) and the reference to the OffsetType table (*OffsetValue, OffsetTypeID*). So, the **DataValues** table is related to many other one. A first set of relationships is the one that define ancillary information about the data, this set includes the fields *SiteID, VariableID, OffsetTypeID, QualifierID, MethodID, SourceID, SampleID, QualityControlLevelID*.

Sites table contains relevant information about the sites in which sensors are located, or different type of data are measured. This information regards the name of the site (*SiteName*), *Latitude*, *Longitude*, *Elevation* and optionally the local coordinates (*LocalX, LocalY, LocalProjectionID*), the last attribute is also one of the two foreign keys shared with SpatialReferences table. Sites table also include information about the *State,* and *County* of the sites, and any other additional *Comments*. As we said before, Sites table is linked with **SpatialReferences,** that contains additional information about the Spatial Reference Systems used for latitude and longitude as well as local coordinate system. So, this table will contain official information obtained from http://www.epsg-registry.org/.

Variables table is very important for our purposes because it contains information that describe in a more precise way the variables observed. We already saw that in the DataValues table there is the *VariableID*, that is the unique identifier of the variable observed. Then, in this table there are information about the code, the name, and the speciation of the variable. *VariableUnitsID*, is a foreign key that links Variables table with Units table. *SampleMedium* indicates the medium in which the sample was recorded, *ValueType* is the type of the recorded data value. Time information are stored in the attributes *TimeSupport* and *TimeUnitsID*. *GeneralCategory* identifies the category of the data values from the Categories table. *NoDataValue* is a numeric value that we can use to indicate when we have not data values for this variable.

Categories table is used to define categories for categorical variables, and this is required for all the variables for which DataType attribute is indicated as "Categorical". In fact, *VariableID* is the foreign key that references to the Variables record. *DataValue* is a numeric identifier for the category and *CategoryDescription* can includes the definition of the categorial variable value.

Units table is useful for storing *Units* and *UnitsType* associated with variables, time, and offsets. It contains four attributes, the primary key is *UnitsID*, the other specify name, type, abbreviation of each units.

OffsetTypes table is important because it contains all the information about each measurement offsets. Each offset is identified by *OffsetTypeID*, the primary key. *OffsetUnitsID* is an integer identifier that references the record in the Units table giving the units of the offset value. In addition, *OffsetDescription* provides a full text description of the offset type (e.g. "Above Ground Level").

Qualifiers table contains comments that qualify data. This table has three attributes, the primary key is *QualifierID*, then there is information about the text code that the organization choose when collects data (e.g. “a” for approved, “e” for estimated) and a more exhaustive explanation (*QualifierDescription*).

QualityControlLevel table is created with the purpose of containing quality control levels that are used for versioning data. *QualityControlLevelID* is the primary key of the table, and represents the quality control level (e.g. 0,1,2,3,4). *QualityControlLevelCode* attribute is important because is the value that is used to identify the level of quality control to which data have been subjected. *Definition* and *Explanation* attributes contain the definition and the explanation of quality control level.

Methods table comprises all the different methods that are used to collect data. Our work is focused on sensors, but that does not prevent the possibility that other type of data will be used in this database, indeed is created also the table Sources. Methods’ primary key is *MethodID* the unique identifier of each method. The other attributes are *MethodDescription* and *MethodLink*, the first one contains a text that describes the method and the second contains link to any type of additional material on the method.

Samples table must necessarily be included because many information requires to carry out analysis on physical samples. So, the table has four attributes: *SampleID*, primary key, that is the unique identifier for each physical sample. *SampleType* specifies the type of the sample, *LabSampleCode* is the tracking of the sample during the analysis in the laboratory. *LabMethodID* correspond to the unique identifier for the method used by the lab during the analysis, and is a foreign key linked to the following table.

LabMethods table contains all the specification of the laboratory methods used to analyze samples. *LabMethodID*, is the primary key, and represents the unique identifier for each lab methods (it is the foreign key used in the previous tab). The other five attributes (*LabName, LabOrganization, LabMethodName, LabMethodDescription, LabMethodLink*) extend the information about the method, there is also the possibility of including links to additional materials.

Sources table lists all the original sources of data, in a way that makes possible to reconstruct the data value from the original file if necessary. The primary key of this table is *SourceID* and is a number that uniquely identifies each data source. *Organization* attribute is the name of the organization that collected the data (e.g. the owner of sensors, or other research laboratories, etc.).

SourceDescription contains the description of the specific source of data and *SourceLink* can contain any link to additional materials. *ContactName, Phone, Email, Address, City, State,* and *ZipCode* of the contact person for the data source (optional). *Citation* is a text corresponding to the citation to be used when the data from each source are referenced. *MetadataID* is a foreign key referencing to the following table.

ISOmetadata table is fundamental for our database, it contains metadata information for compliance with international standards, such as the draft ISO 19115 or ISO 8601. The first four attribute must be non-empty to be compliant with the ISO standards. *MetadataID* is the primary key (linked to the previous table) and it is the unique identifier for each metadata record. *TopicCategory* attribute corresponds to the ISO topic category. *Title* and *Abstract* increase the information for a specific data source. Again, is optional to insert in the attribute *Metadatalink* any additional materials.

Groups table can be used to store data in groups (e.g. all data from 1/1 to 7/1 correspond to "first week of the year group" to which is associated the code 001). The table has only two attributes that are both foreign keys: *GroupID, ValueID* which correspond respectively to integer ID of each group and to the unique identifier of each value.

GroupDescriptions table describes all the group of the previous table. *GroupID* is the primary key (that correspond to the foreign key of the previous table). *GroupDescription* can contain the text description of each group.

SeriesCatalog table after the DataValues one is the most important of the entire database. This table is created to support common data discovery queries in identifying which variables have been measured at which time and location, we use the concept of "data series". Data series is a set of particular observations, that are measured at the same location using the same method from the same source. This table significantly simplifies and improves searching performance, because it is created to answer to some common queries, for example, "which variables are collected for the particular site xyz?" or "which location at which time as a particular value of the variable x?". Using this table, we can create unique combinations (variable/site) using unique combinations of the foreign keys (*SiteID, VariableID, MethodID, SourceID,* and *QualityControlLevelID*). The entire table should be programmatically derived and should be updated every time new data is added to the system. The primary key of the table is *SeriesID* that is the unique identifier for each data series.

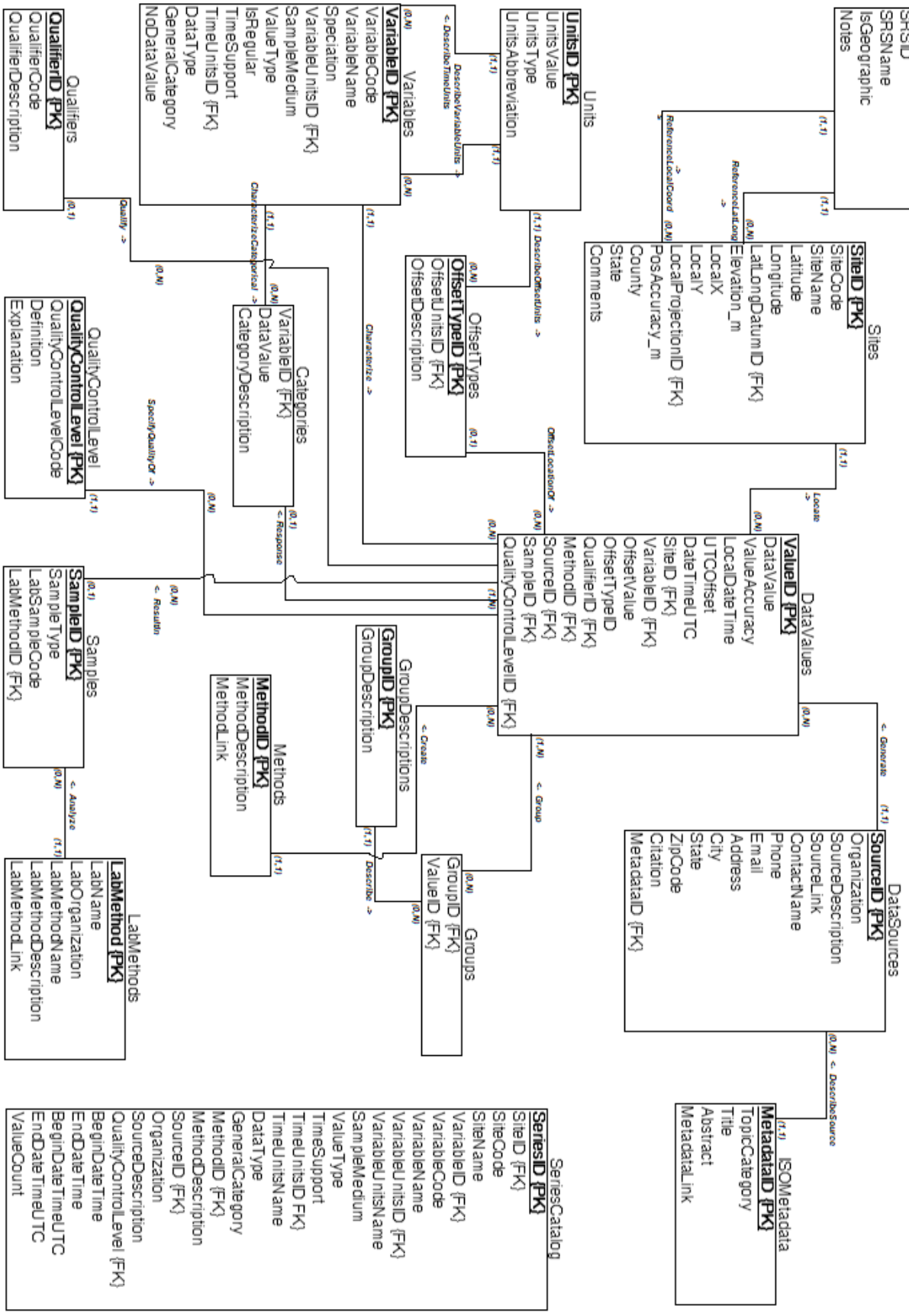


Figure 4: Logical Data Model of the Observations Data Model

6.2 Data manipulation (queries)

In the following paragraph some examples of queries that make use of the database we created will be shown. In order to accomplish this, all the tables that are going to be used have been defined as such:

```
tables used to make sample queries:

Categories(VariableID, DataValue, CategoryDescription)

DataSources(Organization, SourceDescription, SourceLink, ContactName, Phone, Email, Address, City, State, ZipCode, Citation, MetaDataID)

DataValues(DataValue, ValueAccuracy, LocalDateTime, UTCOffset, DateTimeUTC, SiteID, OffsetValue, OffsetTypeID, QualifierID, MethodID, SourceID, SampleID,
           QualityControlLevelID)

Groups(GroupID, VariableID)

GroupDescriptions(GroupID, GroupDescription)

Methods(MethodID, MethodDescription, MethodLink)

Samples(SampleID, SampleType, LabSampleCode, LabMethodID)

SeriesCatalog(SeriesID, SiteID, SiteCode, SiteName, VariableID, VariableCode, VariableName, VariableUnitsID, SampleMedium, ValueType, TimeSupport,
           TimeUnitsID, DataType, GeneralCategory, MethodID, MethodDescription, SourceID, SourceDescription, QualityControlLevelID, BeginDateTime,
           EndDateTime, BeginDateTimeUTC, EndDateTimeUTC, ValueCount)
Sites(SiteID, SiteCode, SiteName, Latitude, Longitude, LatLongDatumID, Elevation_m, LocalX, LocalY, LocalProjectionID, PosAccuracy_m, County, State, Comments)

Variables(VariableID, VariableCode, VariableName, Speciation, VariableUnitsID, SampleMedium, ValueType, IsRegular, TimeSupport, TimeUnitsID, DataType,
           GenaralCategory, NoDataValue)
```

The variables that are going to be used to write the queries are the following:

```
VariableIDs used to make sample queries:

temperature->Temp
PH->PH
mercury->Hg
lead->Pb
cadmium->Cd
selenium->Se
challium->Tl
nickel->Ni
silver->Ag
manganese->Mn
chromium->Cr
iron->Fe
barium->Ba
soda-> Soda
phosphoric acid-> H3PO4
nitric acid-> HNO3
ammonia -> NH3
gasoline->Gsln
kerosene->Krsn
fuel oil->Foil
mineral oil->Minoil
carbon dioxide-> CO2
carbon monoxide-> CO
nitrogen dioxide-> NO2
sulphure dioxide-> SO2
decibel->dB
```

```
SiteIDs used to make sample queries:

z1-z2-z3-z4

SourceIDs used to make sample queries:

name_of_zone + type_of_sensors + sequential_number (e.g. z1_water_001)
```

Firstly, the entities that have access to the database might be interested to understand which variables have been measured at which time and location. In order to do this, they can use the concept of “data series”:

```
---- DATASERIES

-- SeriesCatalog for water temperature for a particular day of the year (01/01/2020) in all the zones

SELECT *
FROM SeriesCatalog
WHERE VariableID LIKE 'Temp' AND SourceID LIKE '%_water_%' AND EndDateTimeUTC LIKE '01/01/2020%'

-- Which Location at which time has a particular value (2.70 g/cm3) of the variable silver (Ag)

SELECT SC.SiteID, SC.SiteName, SC.SourceID, SC.EndDateTimeUTC, DV.ValueID, DV.DataValue
FROM SeriesCatalog AS SC, DataValue AS DV
WHERE DV.VariableID = SC.VariableID AND DV.VariableID LIKE 'Ag' AND DV.DataValue >= 2.70

-- Which Variables are collected for a particular site (z1)

SELECT DISTINCT (VariableID, VariableName)
FROM SeriesCatalog
WHERE SiteID LIKE 'z1'
```

Given the diverse nature of the sensors, which measure different type of parameters for each of the identified elements, the following queries, for each of the four elements (in order: water, soil, air, and noise), have been identified as the most useful for the people that are going to run them:

```
---- WATER OBSERVATIONS

-- PH values for a particular site (z2) in particular datetime (01/01/2020 3:10:00 PM)

SELECT *
FROM DataValues
WHERE SiteID LIKE 'z2' AND DateTimeUTC LIKE '01/01/2020 3:10:00 PM' AND VariableID LIKE 'PH' AND SourceID LIKE '%_water_%'

-- avg temperature for a zone (z3) in a particular day (01/01/2020)

SELECT avg(DataValue)
FROM DataValues
WHERE VariableID LIKE 'temp' AND SiteID LIKE 'z3' AND SourceID LIKE '%_water_%'

-- heavy metals (Cd, Cr, Pb, Ni) above a certain threshold (...)

SELECT SourceID, DateTimeUTC, VariableID, DataValue
FROM DataValues
WHERE (SourceID LIKE '%_water_%' AND
      ((VariableID LIKE 'Cd' AND DataValue >= ...)
      OR (VariableID LIKE 'Cr' AND DataValue >= ...)
      OR (VariableID LIKE 'Pb' AND DataValue >= ...)
      OR (VariableID LIKE 'Ni' AND DataValue >= ...)))
```

```
---- SOIL OBSERVATIONS

-- petroleum hydrocarbons above a certain threshold (...)

SELECT SourceID, DateTimeUTC, VariableID, DataValue
FROM DataValues
WHERE ((VariableID LIKE 'Gsln' AND DataValue >= ...)
      OR (VariableID LIKE 'Krsn' AND DataValue >= ...)
      OR (VariableID LIKE 'Foil' AND DataValue >= ...)
      OR (VariableID LIKE 'Minoil' AND DataValue >= ...))

-- site with the heighest value of gasoline

SELECT SiteID, max(DataValue)
FROM DataValues
WHERE VariableID LIKE 'Gsln'
```

```
---- AIR OBSERVATIONS

-- carbon dioxide, carbon monoxide, nitrogen dioxide, suplhur dioxide in a particular site for a particular day

SELECT VariableID, SiteID, SourceID, DateTimeUTC, DataValue
FROM DataValues
WHERE VariableID IN ('CO2','CO','NO2','SO2') AND SiteID LIKE 'z2' AND DateTimeUTC LIKE '01/01/2020%' AND SourceID LIKE '%_air_%'

-- how many missing observations

SELECT count(VariableID)
FROM Variables
WHERE NoDataValue = 1
```

```
---- NOISE OBSERVATIONS

-- decibel particular site particular datetime

SELECT SiteID, VariableID, DataValue
FROM DataValues
WHERE SiteID LIKE 'z4' AND DateTimeUTC LIKE '01/01/2020%'
```

7. Historical reports database

The second database is used to store all the reports, that are considered as official documents, produced by the Executive Information System and used by policy makers to take decisions. In order to accomplish this task, the database structure that best fits is the document one. The NoSQL database that has been chosen is MongoDB, a type of document database that best fits for the structure type that is utilized by the entity.

7.1 Database structure

The database has a single collection that contains all the reports and it is structured as follows (Figure 5).

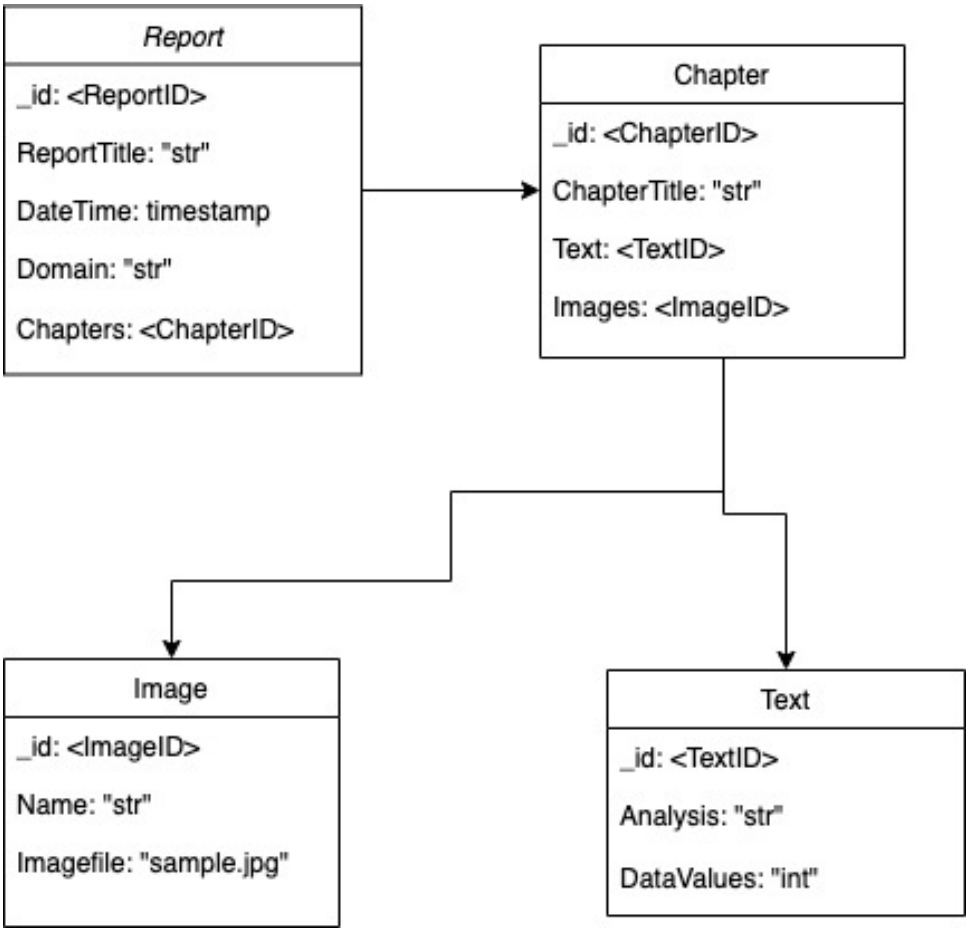


Figure 5: The report database structure

The structure contains four types of documents:

- **Report:** The document that represents all the reports that are stored into the database. Each Report has an associated _id and a ReportTitle field that are useful identifiers. However, the two most important fields are DateTime and Domain. The first one records the date on which the report was first produced and the second one shows to which one of the elements that are analyzed it refers. Each element has a letter associated and the field can contain a single element, if the report was dedicated only to one element, or more, if the report was created to give an aggregate view of all the measurement of a determined time. The possible unique domains,

associated with their related symbol, are: Water (W), Air (A), Soil (S) and Noise (N). The possible combinations of domains are: WA, WS, WN, AS, AN, SN, WAS, WAN, WSN, WAN, ASN and WASN.

This document also contains the field "Chapter", that contains the _id of another kind of documents, called Chapter. The _id is used as a kind of primary key to nest the different documents to create the structure of the report.

All the reports will be stored in a database

- **Chapter:** The Chapter document stores all the chapters of the different reports that are created and is used mainly to represent the structure of the different chapters. Each document has an associated _id and ChapterTitle. The first is crucial because it is used to associate a chapter to its own Report. The two other fields, Text and Images, contain the _id of two other categories of documents.
- **Image:** This type of document contains all the images that are used into the reports. In order to reduce the number of document types all the tables and graphs are stored as images and will follow this structure. This document contains the _id and Name of each image and then a field that contains the .jpg file of the related image.
- **Text:** This last type of document contains all the text of the reports. The _id field is useful to associate each text to the related chapter. The Analysis field is a string and contains the effective text that is written on the document. The DataValues field is used in order to store the data as integer into the document.

The documents are stored in a JSON format and an example of document is represented below:

```
Report: {  "_id": "str",
           "ReportTitle":"str",
           "DateTime": Timestamp,
           "Domain": "str",
           "Chapter":{"_id":"str",
                      "ChapterTitle":"str",
                      "Text":{"_id":"str",
                              "Analysis":"str",
                              "DataID":"str",
                              "DataValue":int},
                      "Image":{"_id":"str",
                               "Name":"str",
                               "ImageFile":.jpg}
                       }

        }
```

7.2 Data manipulation (queries)

In this paragraph are provided some examples of queries that an employee, that is looking for a specific set of repots, can perform using MongoDB as query language.

```
Find all the reports related to a specific Domain

db.collection.find({Domain:"W"})

Find all the reports related to a specific day

db.collection.find({DateTime:"01/01/2020"})

Find all the reports where there are images with specific titles in specific domains

db.collection.find({Domain:"WS","Chapter.Image.Name":"Last week variance of heavy metal concentration"})
```

8. Reporting, analytics & control

The environmental reports will be used by policy makers to implement new rules

The final output of the detection is a report that will be used by decision makers to choose among the policies to apply. That report will flow from the analyst to the upper levels with an ad-hoc information system. Also, the control on the plant will be implemented by an automated IS.

Information systems (IS) are formal, sociotechnical, organizational systems designed to collect, process, store, and distribute information. In a sociotechnical perspective, information systems are composed by four components: task, people, structure (or roles), and technology.

Why it is necessary to use an IS in a governmental organization?

An Information System (IS) can have a major impact on corporate strategy and organizational success. The involvement of decision makers in all aspects of information systems is a major factor for organizational success, including better results and lower costs. Some of the benefits business organization seek to achieve through information systems include better safety, competitive advantage, fewer errors, greater accuracy, higher quality results, improved communications, increased efficiency and productivity, more efficient administration, superior financial and managerial decision making. We often think that information systems are used only in corporate environment, but we forget that also policy and control institutions need information to make

decisions. Moreover, the control system of our institution is very similar to the one of a company: in its business it can be considered an enterprise.

A process control system will be used to monitor the sensors

So, for example, to monitor the state of the sensors we can use a process control system. A process control system can also be used to understand which sensors break more rapidly because are exposed to heavier streams. Also, when we're not able to understand the causes of a high level of pollution we add some other sensors; when we notice high level of pollution in a determinate area, we can decide to place other sensors near production plants in order to identify the faulty plants. The PCS will control if the sensor are working looking at the values they output in order to provide maintenance of hardware and communication systems; also, when we cannot understand clearly the reason of the pollution, we can add some other sensor to better identify the sources of pollution.

Another field that can be covered by information system technologies is decision making; indeed, decision makers need tools to monitor the outcomes of their choices. For example, following some scientific research, a policy maker could decide to move the production in the valley rather than at the source of the river. In this case the DSS will track the levels of pollution in the overall river; or, if they want to boost tourism in the valley, they can decide to stop the most polluting plants/ the nearest to the touristic areas: in this case an EIS can compare the effects on the different assets (working places, economy, etc.).

Analysts will analyze data from the sensors and create reports for different stakeholders

The EIS: executive information system

For the institution we are managing we designed basically two IS. The primary one will *transfer the data regarding the pollution levels of the river to the decision makers*. The *analysts* of our plant will produce the report that will be sent to the *policy makers' assistants*. *Analysts* will take data from the sensors in the plant and then organize them in an organic report that will focus on the different threats and the different *stakeholders* affected.

The PCS: process control system

The secondary one will, instead, simply *monitor if the sensors are working and if the controls are implemented properly*. It will be managed by the *technicians* that will replace the broken ones and, following the recommendations of the

analysts will then move sensors or install new ones to recognize the causes of exceptional pollution. The technology we will use is a **PCS.**

9. Policy

The local authorities will be responsible for the creation of the appropriate policies

The last task of the Environmental Monitor Entity will be the creation of policies that, based on the data collected and analyzed during the other processes of the company, has as main goal the restoration of the area and improve the tourism and business activities not related to industries. The high-quality report created by the policy makers using the Executive Information System of the entity has the aim to give a general overview of the current situation of the entire valley. Moreover, the Information System gives the possibility to its users to gather data also from other sources. This is useful in order to compare the data coming from the entity's database with the ones coming from other different sources in order to give the policy makers a broader view on the problem and the possible solutions that they can implement.

However, the intervention of these actors is not always necessary. In fact, if there are no significative changes on the data gathered by the sensors there will be no need to create new policies. Nevertheless, is important to produce reports not only when there are significant negative changes on the previous data but also good ones. In fact, policy makers must also take in account the possible positive effects of their intervention on the territory in order to direct the future actions and know the better path to follow.

Obviously, since the application of the policies cannot be directly controlled by the entity the functioning of the last part, and also the most crucial, will be handled by the local authorities.

In order to understand the concept here will be provided a short example of a policy created after the notification of significant changes on water data.

Using this query an employee of the entity retrieves the temperature of the water of the river on the 1st January 2020.

```
SELECT avg(DataValue)
FROM DataValues
WHERE VariableID LIKE 'Temp' AND SourceID LIKE '%_water_%' AND DateTimeUTC LIKE '01/01/2020%'
```

Using this data is possible to perform some researches and compare the temperature with the one of rivers that are situated in territories with similar air temperatures during this time of the year. A further check could be performing a comparison with the historical data of those rivers during the past years, in order to take in account, the possible changes due to changes in atmospheric temperature. [N.B. It is important to underline that in this example it is necessary to make comparison with other rivers due to the fact that the it is being considered the first year of data collection. When there will be enough historical data those comparison will be unnecessary or used only for further analysis]. Using the data retrieved with this research it is produced a report which shows that the temperature is significantly higher than the one that the river normally should have.

After the acquisition of this information policy maker have the possibility to perform further checks, such as monitoring the level of soda, phosphoric acid or nitric acid (typical refuses produced by the industries operating in the Elk River valley).

```
SELECT SourceID, DateTimeUTC, VariableID, DataValue
FROM DataValues
WHERE (DateTimeUTC LIKE '01/01/2020%' AND SourceID LIKE '%_water_%' AND
       ((VariableID LIKE 'Soda' AND DataValue >= ...)
       OR (VariableID LIKE 'H3PO4' AND DataValue >= ...)
       OR (VariableID LIKE 'HNO3' AND DataValue >= ...)
       OR (VariableID LIKE 'NH3' AND DataValue >= ...)))
```

If the level of these substances is significantly higher than the standards [N.B. These standards do not need previous analysis because they are valid for all the different water flows around the world and just point the level of these substances that can be considered harmful for people and environment] means that the increase of water temperature is due to the high level of polluting substances in the water. Having all this information it is possible to prescribe the temporary stop (or reduction of the activity) of all the industries that produce the refuses stated above. The last step will be the communication to the local authorities of the new directive and then, after it is application, check the new data in order to view the effects.

10. Conclusion

The creation of the Environmental Monitoring Entity should lay the foundation of a solid restoration project for the entire Elk River valley. Obviously, not only it is fundamental that the policies decided by the policy maker are aimed to enhance the quality of the life but also it is necessary that all the counties, and their related authorities, cooperate in order to have a common application of all the directions given by the entity. If everything is performed correctly the Environmental Monitoring Entity will be able to improve prosperity of the entire valley and also give an important boost to all the activities that are connected to the environment. This should have an important impact especially on the tourism, agricultural and lodging sector while only partially damaging the most pollutant industries.

In addition to this, the development of a mobile app to let the citizens access to the detected values of pollutants in the water, the soil and the air, will make possible the spread of alarms in case of a disaster in an efficient and effective way. Following this path, the State of West Virginia will be able to avoid all the diseases caused by the lack of communication and information in January 2014. Moreover, the development of a mobile application is useful to give citizens an easier access to the database. This is an important step to improve the citizens awareness about the conditions of the environment in which they are living and also modify their lifestyle in a more sustainable way.

11. References

- By Tim Kiser (User:Malepheasant) - Own work, data from w:The National Map, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=26871476>
- Kenny, Hamill (1945). West Virginia Place Names: Their Origin and Meaning, Including the Nomenclature of the Streams and Mountains. Piedmont, WV: The Place Name Press. p. 229.
- "Elk River." West Virginia Explorer, 11 Dec. 2019, <https://wvexplorer.com/attractions/rivers-streams/elk-river/>
- Botelho, Greg, and Tom Watkins. "Chemical Levels in West Virginia Water Drop, but Still No End in Sight." CNN, Cable News Network, 11 Jan. 2014, <https://edition.cnn.com/2014/01/09/us/west-virginia-contaminated-water/>
- "About Sutton Lake: Sutton, WV: 304.765.2120." Suttonlakemarina, <https://www.suttonlakemarina.com/about-sutton-lake>
- Zachary J. Loughman and Stuart A. Welsh "Distribution and Conservation Standing of West Virginia Crayfishes," Southeastern Naturalist 9(sp3), 63-78, (1 January 2010). <https://doi.org/10.1656/058.009.s304>
- Botelho, Greg, et al. "Bottled Water for West Virginia Residents Plagued by Chemical in Water Supply." CNN, 11 Jan. 2014, <https://web.archive.org/web/20140111152737/http://www.cnn.com/2014/01/11/us/west-virginia-contaminated-water/>.
- Bernstein, Lenny. "Chemical Spill into W.Va. River Spurs Closures, Scramble for Bottled Water." Washington Post, 10 Jan. 2014, <https://web.archive.org/web/20140111033321/http://www.washingtonpost.com/national/health-science/chemical-spill-into-wva-river-spurs-closures-run-on-bottled-water/2014/01/10/a6ec518a-7a0e-11e3-b1c5-739e63e9c9a7_story.html>.
- "Freedom Industries Files for Chapter 11 Bankruptcy." Financier Worldwide, <https://www.financierworldwide.com/freedom-industries-files-for-chapter-11-bankruptcy#.XfaY0-hKg2w>.
- Laws, Edward A. Aquatic Pollution: an Introductory Text. John Wiley & Sons, Inc., 2018.
- Goel, P.K. Water Pollution: Causes, Effects and Control. New Age International, 2006.
- NHS Choices, NHS, 17 Oct. 2017, www.nhs.uk/conditions/carbon-monoxide-poisoning/.
- Essick, Peter. "Air Pollution Causes, Effects, and Solutions." Air Pollution, Facts and Information, 25 June 2019, www.nationalgeographic.com/environment/global-warming/pollution/ .
- "Noise." World Health Organization, World Health Organization, 2018, www.euro.who.int/en/health-topics/environment-and-health/noise.

- “Narrowband IoT (NB-IoT).” UBlox, 31 Oct. 2017, www.u-blox.com/en/narrowband-iot-nb-iot.
- Cui, Yaokui, et al. “Global Water Cycle and Remote Sensing Big Data: Overview, Challenge, and Opportunities.” Big Earth Data, vol. 2, no. 3, Mar. 2018, pp. 282–297., doi:10.1080/20964471.2018.1548052.
- Horsburgh, J., Tarboton, D., Maidment, D. and Zaslavsky, I. (2008). A relational model for environmental and water resources data. Water Resources Research, 44(5).
- Tarboton, D., Horsbourgh, J. and Maidment, D. (2008). CUAHSI Community Observations Data Model (ODM) Version 1.1 - Design Specifications.