Data Analytics in Public Safety

There is great potential for data analytics to benefit public safety and keep citizens safer. This document examines the different concepts of data analytics and how they can be applied in the public safety sector.

eena

EUROPEAN EMERGENCY NUMBER ASSOCIATION

Data Analytics in Public Safety

EENA

European Emergency Number Association

EENA 112

Avenue de la Toison d'Or 79, Brussels, Belgium
T: +32/2.534.97.89
E-mail: info@eena.org

Authors that contributed to this document:

Authors:

Carlos Rosa, 112 Canarias
Iratxe Gomez Susaeta, Atos

Contributors:

Cristina Lumbreras, EENA
Francisco Nobre, Esri
Stuart Walsh, Creativity Software

1 | INTRODUCTION

Data analysis is not new. The first traces arise from Babylonian civilization (5000 B.C.), from which tablets written in sexagesimal notation have been found, containing lists of people, goods and quantities of food brought as offerings.

In the field of mathematics, statistical techniques arise to analyse and draw conclusions from a dataset. That is, to make decisions based on the observation of phenomena.

For many years, the main handicap for the evolution of this discipline had been the difficulty of handling large amounts of data. However, the evolution of computing, communications networks and the emergence of the Internet have allowed the field of work to be extended to a new context, including the ability to efficiently manage three main variables:

- The Variety of Data Sources
- The Velocity at which we process that data
- The high Volume of information managed

Those "three Vs", to which the Veracity and Value of the data are added, are those that give rise to the current concept of BIG DATA.

2 I KEY CONCEPTS OF DATA ANALYTICS

- **Data scientists**[1]

The Data Scientist is a professional profile that translates the large volumes of information available from all kinds of sources and varied formats into useful information. They work in any type of business and industry with the aim of obtaining reliable answers to everyday problems.

Common steps in this process of transforming data into value are:

- Extraction, apart from fonts or formats.
- Cleaning the data.
- Data processing using statistical methods.
- Redesigning the data, if necessary.

The profile of a Data Scientist combines mathematical and statistical knowledge with the mastery of statistical software and programming, as well as massive data analysis systems such as machine learning. They also need to master technology and databases.

This is the basic job of the data scientist[2]: to bring new ways of measuring and displaying data so that more awareness, clarity and direction is provided to those looking at it[3].

- **Descriptive Analytics**

This gives insight into historical data and yet gives trends to dig deep into the data for more details using descriptive statistics, interactive data exploration and data mining. Examples are summary statistics, clustering and association rules.

- **Diagnostic Analytics**

This is applicable when trying to know why something happened i.e. finding an answer to the "why did it happen question". It is employed by using data mining and data warehousing techniques. Examples are diagnostic assessment, churn reaction analysis and customer health score analysis.

[1] *Source:* https://www.iebschool.com/blog/cientifico-datos-big-data/
[2] *Source:* Data Scientists: The New Rock Stars of the Tech World (techopedia.com)
[3]*Source:* KayodeAbiodun, Oladapo, 2018. An Overview of Data Analytics in Emergency Management. I*nternational Journal of Computer Trends and Technology, 63 (1).*

- **Predictive Analytics**

It consists of the use of mathematical models to forecast future events. It is the most widely used part today in virtually all fields: risk analysis, market evolution, call behaviour in a call centre, etc.

- **Prescriptive Analytics**

It helps in picking the best solution among a variety of choices based on known parameters and then suggests how to take advantage of future opportunities or mitigate a future risk.

- **Geo-Analytics**[4]

"GIS uses datasets with a spatial aspect or component which can be defined as Geodata (also "spatial data", "geographic data", "geographic data sets", "geoinformation", or "G data"). Geodata has a spatial, temporal, IS and thematic aspect, can be linked to other data sources and represents the core value of geographical information systems. Hence, spatial queries, analysis and simulations can be conducted if such larger and complex data structures can be managed. It is appropriate to distinguish between spatial base data and thematic data as a subset of geodata."

"[Big Data spatial analysis] is an area of GIS analysis that supports the analysis of information fostered by IOT (Internet of Things) mega trends. The time-enabled data from sensors and social media tends to grow exponentially in any given incident. This implies that the GIS analysis will need to support big data."

- **Forecasting and planning analysis**[5]

"Planning and forecasting is the managerial process of mapping out corporate actions based on past and present data trends. Planning encompasses everything from high level corporate plans to strategic plans, to operational, HR, expense, capacity, sales and operational, balance sheet, profitability, capital, cash flow planning and more. Some may be long term, like strategic planning, while others are short term, like operational planning. To create a plan is to detail KPIs and events that should lead up to reaching a specific objective. Since forecasts are predictions of future events, plans often use forecasts in order to inform the decision making process."

[4] Source: https://eena.org/knowledge-hub/documents/spatial-information-for-emergency-services/
[5] Source: Planning and Forecasting - Glossary | CCH Tagetik

- **Graph data analysis**[6]

  "Graph analytics, also known as network analysis, is an exciting new area for analytics workloads. To some extent, the business driver that has shone a spotlight on graph analysis is the ability to use it for social network influencer analysis. Marketing managers in particular are interested in identifying social network influencers because they are potential targets for marketing campaigns trying to trigger chain reactions among social network communities to buy products and services. However, many more potential use cases for graph analysis exist, including several widely used business use cases:

  - *Detecting financial crimes such as money laundering*
  - *Spotting fraud, which applies to fraudulent transactions and applications in banking, benefits fraud in government, applications and claims fraud in insurance and fraudulent activities in telecommunications*
  - *Preventing crime and performing counterterrorism*
  - *Applying influencer analysis in social network communities*
  - *Performing grid and network quality of service such as identifying weaknesses in power grids, water grids and transportation networks as well as helping prevent cybercrime attacks on computer networks*
  - *Optimizing routes in the airlines and retail and manufacturing industries as well as for supply distribution chains and logistics*
  - *Conducting research in life sciences (bioinformatics) including medical research, disease pathologies and so on."*

- **Spatial data analysis**

Spatial analysis includes all transformations, manipulations, and methods applicable to geographic data to add value, support decisions, and identify patterns and anomalies. It uses concepts such as location, distance, and spatial interactions as central elements of the data by applying specialised methods and software to analyse, visualise, and apply learning to spatial use cases.

- **Time-series analysis**

Time Series Analysis is the statistical study of samples of variables collected sequentially over time. It is one of the most relevant areas in the field of emergency services for the study of behaviours over time of different categories of incidents.

- **Flow analysis**

Data-flow analysis is a technique used by software engineers to analyse the way values of variables change over time as a program is executed.

[6] *Source: What is graph analytics? | IBM Big Data & Analytics Hub (ibmbigdatahub.com)*

3 I DATA ANALYTICS APPLIED TO PUBLIC SAFETY

Data analytics has been part of the activities of Public Safety organisations for a long time, but a more and more data-centric world and the evolution of technology has brought it into the spotlight.

3.1 Relevant studies

STUDY Nº 1: Pre-Crime (beyond movies…)

A case that shows the possibilities of data analysis is the case study led in 2018 by Santiago González, IT Director of the strategic consulting area of LUCA (Telefónica), in collaboration with the City of New York and Columbia University.[7]

Using a variety of open data sources, such as Open Data New York[8] and similar open-access data sources, it was aimed at predicting possible crimes in certain areas of the city. Multiple analytical models were applied to achieve the goal of having sufficiently reliable predictions in a short space of time. The conclusion in this case was that the best results were obtained using XGBoost models. These models are based on generating many simple sequential models in such a way that each new model uses the result of the previous model, so that the results in each iteration of the model are increasingly accurate.

This information opens the way for optimising the distribution of security resources and minimising the chances of certain criminal activities materialising.

STUDY Nº 2: Social Media analytics related to crisis – monitoring & prediction[9]

Social media analytics can be useful in several stages of a crisis. A lot of studies have shown that a crisis can be detected through social media analytics even before the incident is communicated officially.

Using social media analytics would provide the possibility of responding to crises more quickly and sending help earlier. Data collection methods and even social media analytics tools already exist.

For example, the American Red Cross uses social media to extract information provided by the public. Social media was used as a participatory media during Hurricane Katrina (2005).

For detailed exploration of this topic, two important sources are: Social Media Analysis in Crisis Situations: Can Social Media be a Reliable Information Source for Emergency

[7] *Source: https://empresas.blogthinkbig.com/luca-talk-prediciendo-el-crimen-en/*
[8] *Source : http://opendata.cityofnewyork.us/*
[9] *Source : Stieglitz et al., 2018. The Adoption of Social Media Analytics for Crisis Management – Challenges and Opportunities. European Conference on Information Systems (ECIS). Project: The Potential of Social Media for Emergency Agencies.*

Management Services? [10]; The Adoption of Social Media Analytics for Crisis Management – Challenges and Opportunities.[11]

3.2 Availability of open datasets

Open data can and should be used by linking it to data from Public Safety organisations, and the current pandemic situation has proven more than ever the relevance of such an approach.

Some of the data sources available during the current pandemic came from entities such as the World Health Organisation, the European Centre for Disease Prevention and Control, the National Health Commission of the People's Republic of China and the US Centres for Disease Control and Prevention, among many others.

Datasets from trusted sources, combined with internal information, can help organisations enrich their analysis, and therefore provide better information to their stakeholders.

3.3 Location analytics

Location Analytics is becoming more and more important for Public Safety agencies. The demands for greater planning and faster reaction to public safety events are being driven by governments and organisations like the European Emergency Number Association (EENA). In the modern and fast-moving world, we are required to plan for public safety events (Disasters, Terror, Pandemics, etc.) and react to these events in a fast and efficient manner.

In order to carry out Public Safety Location Analytics, there are broadly three elements that need to be considered:

3.3.1 Location data sourcing

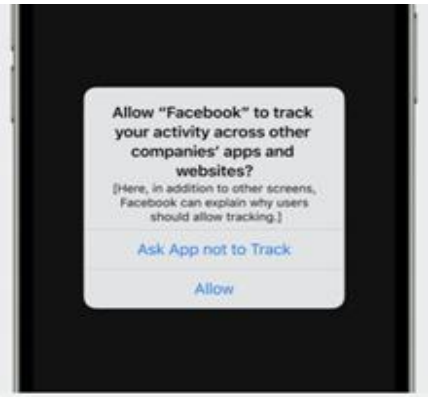
One of the biggest challenges for Public Safety analysts is getting access to accurate and timely location data for analysis. Typical sources vary from cloud sources (web and mobile apps) to communication networks. The reality is that all sources of location data need to be consolidated to provide the best picture for Location Analytics.

3.3.1.1 Cloud Sources

- These sources are wide and varied. Typically, location data gathered from mobile apps makes up a large portion and provides the widest picture of location intelligence.

[10] *Source: Lazreg et al., 2018. Social Media Analysis in Crisis Situations: Can Social Media be a Reliable Information Source for Emergency Management Services?. 27th International Conference On Information Systems Development (ISD2018 LUND, SWEDEN). Project: Social Media for Integrated Emergency Management.*
[11] *Source : Stieglitz et al., 2018. The Adoption of Social Media Analytics for Crisis Management – Challenges and Opportunities. European Conference on Information Systems (ECIS). Project: The Potential of Social Media for Emergency Agencies.*

- Mobile apps typically use GPS from the mobile phones to determine location but other sources are utilised such as WiFi nodes.
- As the main operating system providers in the market today (Google, Apple) are focusing on ensuring location data is not shared, it is envisaged that this source may become fractured in the coming years.



3.3.1.2 Communication Networks

This source comes from the various communication network providers such as mobile networks, WiFi providers and Internet Service Providers. This approach can be less accurate but more reliable as the data is generated from the communication network itself and does not rely on open sources. Location data is generated typically by the user's proximity to network nodes (Mobile Cells etc.). The challenge here is that you have to work with a variety of companies and attempt to coordinate and rationalise the location data you receive from them, it also relies on government pressure to get the companies to generate and release the data.

The challenges in getting this data will increase exponentially with the advent of 5G and IoT, they will allow for greater accuracy and the shear amount of data. If handled correctly it can become a great source for Public Safety organisations[12].

3.3.2 Location data treatment and storage

3.3.2.1 Treatment

This breaks into two primary tasks.

- Anonymisation (when required)
  - This has become a priority with the advent of rules like GDPR[13].

[12] *Source:* https://go.forrester.com/blogs/predictions-2021-technology-diversity-drives-iot-growth/
[13] *Source:* https://www.itgovernance.co.uk/dpa-and-gdpr-penalties

- Any identification of an individual's location will need to be removed.

- Location calculation and verification

  - All sources of data will need to be complied and turned into reliable/accurate location information[14].

  - Artificial Intelligence (AI) offers some great capabilities to both calculate accurate locations and identify trends in people's movements which would take humans much longer to carry out.

  - Indoor location calculation is one of the greatest challenges for Location Analytics.

    A leading authority on this has been the Federal Communications Commission in the U.S[15].

3.3.2.2 Storage

With the sheer number of data points available for Location Analytics, picking the right storage solution to allow easy and timely access for analysis is key. Depending on the length of time required (3-5+ years) and the population of that country, the data stored could become truly enormous. Solutions exist to handle PETABYTEs of data and beyond.

Cloud storage solutions are starting to become acceptable mechanisms for storing vast quantities of data reliably and securely[16].

3.3.3 Location visualisation and analysis

Once you have your location data stored, having the right tools is important to allow analysis to take place which will aid the authority in their planning exercises and in responding to events.

As discussed in previous sections of this document, GIS and mapping technologies combined with Location Intelligence offer ideal mechanisms to analyse where people are located over time and also during a specific event.

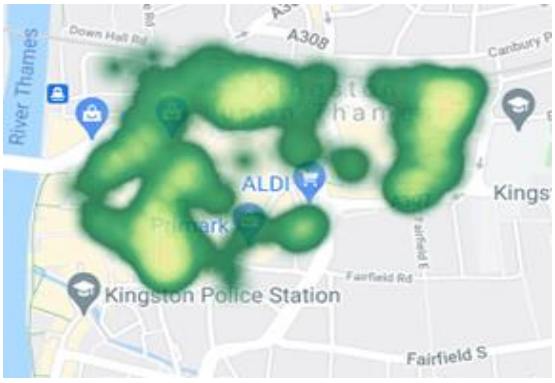
3.3.3.1 Safety planning

Key to public safety is having the right data to ensure you have the right support and services available to respond to emergencies and above all prevent disasters.

[14] *Source:* https://www.sciencedirect.com/topics/engineering/location-estimation
[15] *Source:* https://www.fcc.gov/public-safety-and-homeland-security/policy-and-licensing-division/911-services/general/location-accuracy-indoor-benchmarks
[16] *Source:* https://www.comparethecloud.net/articles/blockchain-technology-the-future-of-storage/

Examples:

- By being able to visualise where people spend their time during busy periods, you can deploy your services to be nearby in case of emergencies.
- This could also include examining trends over time as people migrate to different areas, ensuring you are constantly planning ahead.



- You can also carry out historical analysis to understand how people reacted in an emergency situation and learn lessons from their movements to ensure infrastructure and services are in place for future events.

- During a pandemic, you can analyse if people are adhering to lockdown or curfew rules

3.3.3.2 Responding to an event

Location Analytics can also be utilised in relative real-time to help Public Safety agencies to focus their resources in the right place and message the public appropriately.

Examples

- During a terrorist incident, you can utilise analytics to understand if people are moving towards the danger area and utilise public warning systems to message them to stay away or take refuge in an appropriate location.

- During a pandemic, it is important to react quickly and in a focused manner. Using analytics, you can react to an outbreak and take action to ensure it is contained.

3.4 What already exists

The examples below represent some data analytics solutions already in use.

3.4.1 Data-centric approach of the LFB

On 11/5/2020 EENA hosted a webinar on "Data analytics in emergencies"[17]. There were several very interesting presentations, including one about the data-centric initiatives of the London Fire Brigade, presented by Apollo Gerolymbos, who is in charge of data analytics for the London Fire Brigade.

Some key aspects to highlight from his presentation are:

- The London Fire Brigade has an interdisciplinary team with "data literate people with data skills"
- They collaborate with all the relevant stakeholders.
- They have systems in place to interrogate datasets.
- They publish AND pull in open data.

3.4.2 112 Canarias: Statistical analysis of the correlation between COVID-19 calls to 112 and positive cases

In the Canary Islands, which pioneered in Spain in the implementation of a telephone service line dedicated to the Sars-CoV-2 virus, a study was developed on the calls received and the cases detected later on.[18]

This service was organised on three levels:

- Level 1: General consultations and no reference to medical symptoms
- Level 2: Attended by healthcare professionals and focused on calls that refer to symptomatology
- Level 3: Transferred to the 112 service as emergency cases

[17] *Source :* https://eena.org/knowledge-hub/webinars/data-analytics-emergencies/
[18] *Source :* *https://elpais.com/sociedad/2020-05-07/la-epidemia-que-avisa-por-telefono.html*

The statistical analysis carried out allowed the establishment of a correlation between the calls managed by Level 2 and the positive cases detected approximately five days later.

The professor responsible for the study, Beatriz González López-Valcárcel, Professor of Economics, found that for every 100 calls received at Level 2, approximately 2.4 confirmed cases could be expected in the next 5 days.

In this case, the simplicity of the approach contrasts with the practical use that allows to better model the health resources that will be compromised in the short term.

3.4.3 Geo-analytics capabilities provided by Copernicus EMS

The Copernicus Emergency Management Service[19] (CEMS) provides several tools involving geo-analytics for use by regional or national Public Safety organisations worldwide.

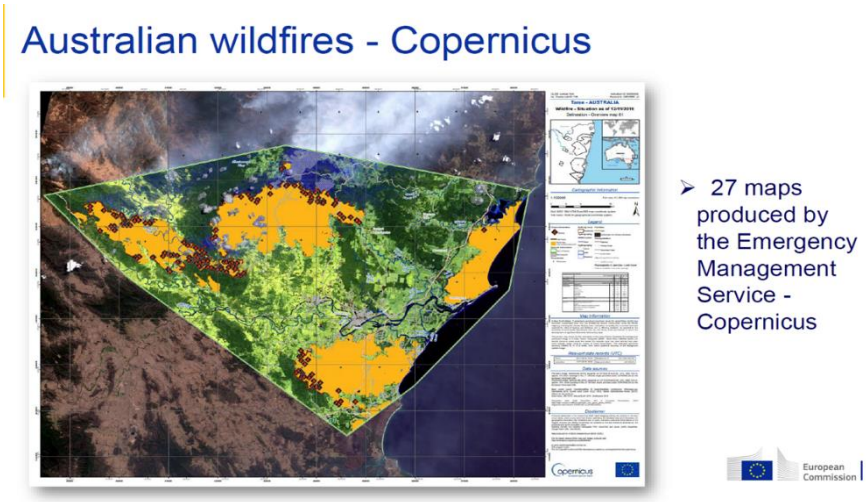
CEMS supports these actors by providing them with information from satellite and complimentary data sources for their decision-making processes, including:

- On-demand Mapping, with a portfolio including both Rapid Mapping and Risk & Recovery Mapping, which can be triggered only by or through authorised users.

Source: *https://twitter.com/CopernicusEMS/status/1324285860242739200?s=20*

The example below was presented in the EENA webinar on 23rd June 2020 about the lessons learnt from the Australian wildfires 2019-2020[20] (the service is extensively used for flood mapping and other disasters in Europe and beyond). A very relevant statement was given by another presenter in the same webinar, County Fire Authority CEO/CO Steve Warrington: "*We will not understand the full impact of the fires until we gather and analyse all the fire-event data*."

[19] *Source :* *https://emergency.copernicus.eu/faq.html*
[20] *Source :* https://eena.org/knowledge-hub/webinars/lessons-learnt-from-the-australian-wildfires/



- Information about floods, through the European Flood Awareness System (EFAS) and the Global Flood Awareness System[21] (GloFAS), which make extensive use of the climate data store provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) for the two systems they produce.[22]

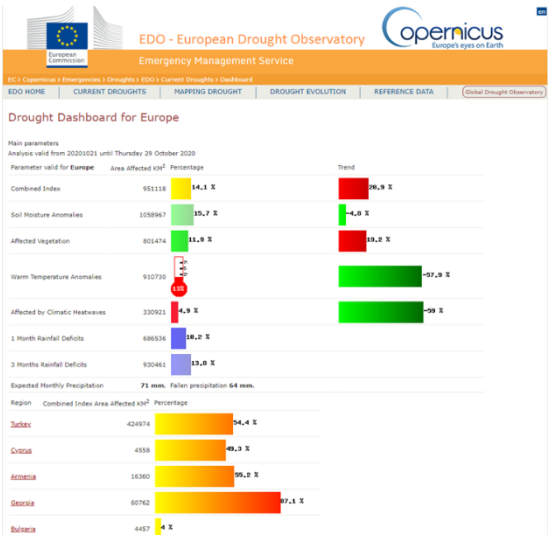
Source: *https://twitter.com/CopernicusECMWF/status/1324683102170525704?s=20*

[21] *Source :* https://www.globalfloods.eu/general-information/about-glofas/
[22] *Source:* *https://www.ecmwf.int/en/about/media-centre/news/2020/new-upgrades-deliver-step-change-improvements-flood-forecasting*

- Information about fires, through the European Forest Fire Information System[23] (EFFIS). This is very relevant as 2019 was the worst-ever year for forest fires around the world in recent history[24], with 2020 following very close.

Source: *https://twitter.com/CopernicusEMS/status/1322160597799194627?s=20*

- Information about droughts, through the European Drought Observatory[25] (EDO), providing maps of indicators derived from different data sources (e.g., precipitation measurements, satellite measurements, modelled soil moisture content). This includes irregularly published "Drought News" that give an overview of the situation in case of imminent droughts[26].



[23] *Source: https://effis.jrc.ec.europa.eu/*
[24] *Source: https://ec.europa.eu/commission/presscorner/detail/en/IP_20_1995*
[25] *Source: https://edo.jrc.ec.europa.eu/edov2/php/index.php?id=1000*
[26] *Source: https://edo.jrc.ec.europa.eu/edov2/php/index.php?id=1112#*

3.4.4 Bluetooth contact tracing apps for COVID-19

In the last year, several countries and regions have released contact tracing apps for citizens, which use the Bluetooth technology on mobile handsets, and make extensive usage of data analytics.

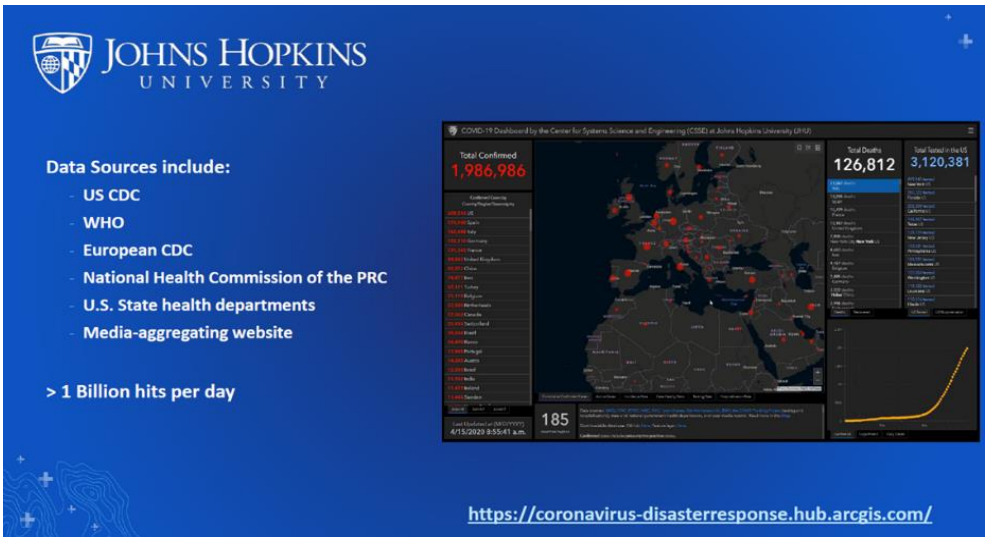
A recent EENA document about COVID-19 apps[27] presented some of the apps and other initiatives, such as the collaboration between Google and Apple that resulted in the privacy-preserving tracing project. With this collaboration, the companies hoped to harness the power of technology and help countries to slow the spread of COVID-19, so they released technical documentation for different frameworks and specifications: Bluetooth Specification, Cryptography Specification, Framework API.

3.4.5 Geo-analytics applied to COVID-19 response

In the same EENA webinar on Data Analytics as previously mentioned, ESRI presented their response to the COVID-19 pandemic for the Public Safety and Healthcare domains, focusing on spatial data analytics.[28]

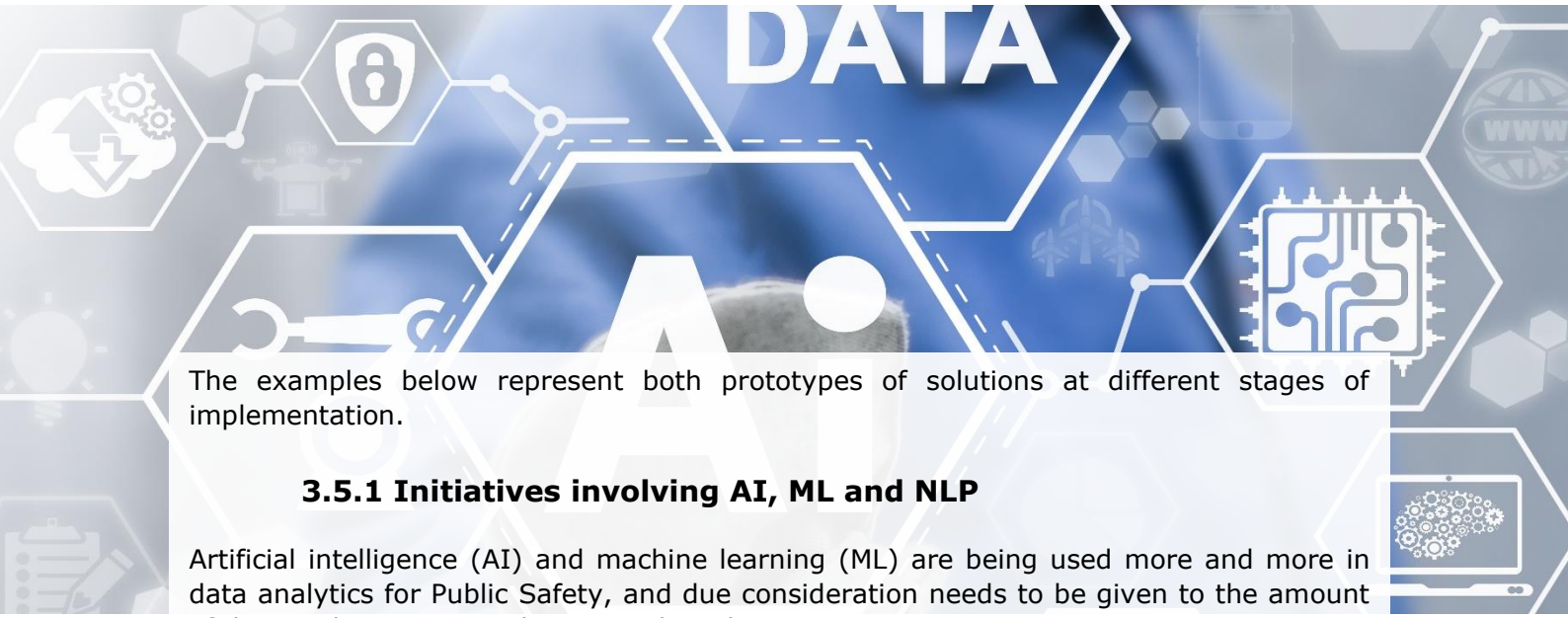
They listed several data types they work with, grouped as base data, thematic data, incident related data and emergency services related data. They also described several types of analysis - situational analysis, real-time and big data spatial analysis, indoor location and mapping, and 3D modelling.

Within their COVID-19 response package, several data sources were included, from the US and European Centres for Disease Control and Prevention, the World Health Organisation, and national health departments from China and US.



[27] *Source : https://eena.org/knowledge-hub/documents/covid-19-apps/*
[28] *Source : https://eena.org/knowledge-hub/webinars/data-analytics-emergencies/*

3.5 Examples of what is possible



The examples below represent both prototypes of solutions at different stages of implementation.

3.5.1 Initiatives involving AI, ML and NLP

Artificial intelligence (AI) and machine learning (ML) are being used more and more in data analytics for Public Safety, and due consideration needs to be given to the amount of data and time required to train algorithms.

A national public safety organisation in Europe wants to apply Machine Learning techniques in the analysis of the quality of the service provided by the Public Safety Answering Point, by looking into the correlation between the recordings of emergency calls and the data collected by operators. They would also like to have automatic speech to text transformation - including real-time translation of foreign languages - which requires analysing vast amounts of voice and data files to prepare the algorithms.

Also, the London Fire Brigade is carrying a proof of concept of free text analysis using Python and R to analyse unstructured text data (natural language processing), which they presented in the EENA webinar mentioned above. They are looking into potentially gaining new insights from text entries and commentary about incidents, health and safety events, transcripts etc.

3.5.2 Data analytics to improve health and safety at work in the Mid & West Wales FRS

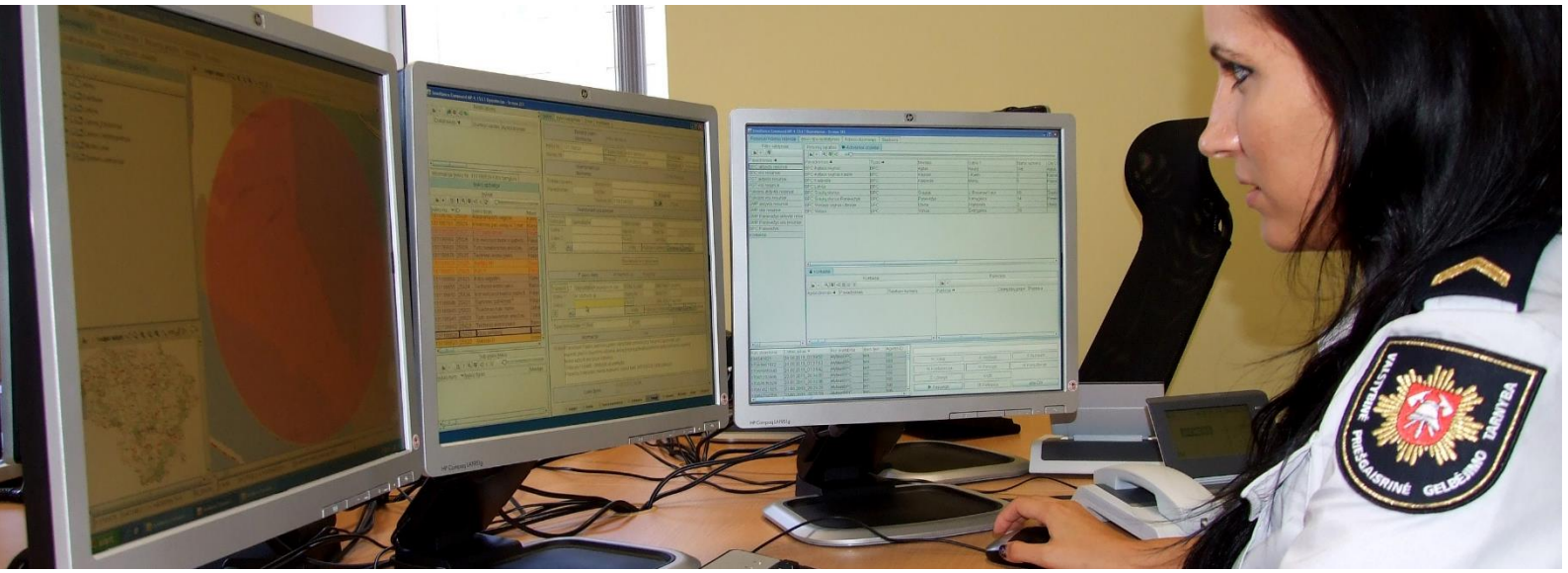
On April 2020, in an EENA webinar[29] about Health and Safety at Work, a very interesting presentation was given by Group Manager Steve McLinden from the Mid and West Wales Fire and Rescue Service. It gave us insight into some of the initiatives they have in collaboration with partners for the use of a data-centric approach for enhancing firefighter safety with the help of technology.

A couple of use cases linked to firefighter tracking data were presented:

- Using real-time tracking data for providing enhanced situational awareness.
- Combining tracking data with live images to potentially locating lost or distressed firefighters.

[29] *Source: https://eena.org/knowledge-hub/webinars/health-and-safety-at-work/*

4 l CONCLUSIONS & RECOMMENDATIONS



Public Safety organisations should engage data-literate personnel (i.e. data scientists) in their organisations. They should also work on Proof of Concepts, and consider using open data sources and Machine Learning techniques to enrich their activities.

Utilising Location Analytics, Public Safety agencies can be proactive and intelligently react to live events. Location data provides valuable insights for both historical and live trends. Combined with many of the Big Data techniques discussed in this document, Public Safety organisations can transform how they plan and react to emergency situations.

Ethical algorithms / Risks associated with algorithms;

According to the Ethics Guidelines for Trustworthy Artificial Intelligence[30] (AI) from the European Commission, AI should be:

(1) lawful - respecting all applicable laws and regulations

(2) ethical - respecting ethical principles and values

(3) robust - both from a technical perspective while taking into account its social environment

The ALTAI portal[31] is a European Commission tool providing an assessment list for Trustworthy AI, which should be used whenever ML techniques are to be applied in the activities of Public Safety organisations.

[30] Source : *https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai*

[31] Source :*https://futurium.ec.europa.eu/en/european-ai-alliance/pages/altai-assessment-list-trustworthy-artificial-intelligence*

5 I RESOURCES

EENA documents:

- Spatial Information for Emergency Services (2020)
- COVID-19 apps (2020)
- Artificial Intelligence & Machine Learning in Public Safety (2019)

EENA webinars:

- Data analytics in emergencies (2020)
- Health and Safety at Work (2020)
- Lessons learnt from the Australian Wildfires (2020)