AIGOV

Implementing ethical, trustworthy and fair Artificial Intelligence Systems in Public Sector

D1.2 The AIGOV Ecosystems

Version-Status:	V1.0 Final
Date:	31/3/2023
Dissemination level:	PU

The research project was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the **"2nd Call for H.F.R.I. Research Projects to support Faculty Members & Researchers"** (Project Number: 2412).

Deliverable factsheet

Project Number:	2412
Project Acronym:	AIGOV
Project Title:	Implementing ethical, trustworthy and fair Artificial Intelligence Systems in Public Sector
Principal Investigator:	Konstantinos Tarabanis
Scientific Area:	SA9. Management & Economics of Innovations
Scientific Field:	9.4 ICT enabled Innovation, Digitisation and Industrial Renewal
Host Institution:	University of Macedonia
Collaborating Organization:	Region of Central Macedonia, Greece

Deliverable title:	The AIGOV Ecosystems
Deliverable number:	D1.2
Official submission date:	31/03/2023
Actual submission date:	31/03/2023
Author(s):	Areti Karamanou, Dimitrios Zeginis, Maria Zotou, Antonia Psaltou, Evangelos Kalampokis, Konstantinos Tarabanis

Abstract:	This document is the second deliverable, entitled D1.2 "The AIGOV Ecosystems", of the first work package of the AIGOV project. The aim of T1.2 is to identify all the involved stakeholders (e.g., citizens, businesses, etc.) who may interact with, or impacted by, an AI-based solutions as well as the roles that these stakeholders could potentially undertake. In addition, T1.3 aims to propose a stable ecosystem and explain how this relates to innovating and transforming public sector processes, practices, policies, and services. This work follows an exploratory case study method to create the ecosystem. Specifically, the method that was followed in this deliverable comprises the following actions: • Create a case study that focuses on transport data

	• Conduct semi-structured interviews with the public servants Region of Central Macedonia, Greece • Disseminate a questionnaire to public servants These actions resulted in a set of artefacts that complementary constitute the AIGOV ecosystem, and the dependencies with each other. The ecosystem could serve as guidelines to adopt AI in Public Administration.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

The following table presents the acronyms used in the deliverable in alphabetical order.

Abbreviation	*Description*
AI	Artificial Intelligence
ML	Machine Learning
NLP	Natural Language Processing
PA	Public Administration
WP	Work Package

Executive Summary

This document is the second deliverable, entitled D1.2 "The AIGOV Ecosystem", of the first work package of the AIGOV project. The objective of the AIGOV project is to facilitate the implementation of fair, ethical, trustworthy, and robust, both from a technical and social perspective, Artificial Intelligence (AI) systems in Public Administration (PA) with a focus on users and those who may be affected.

D1.2 is the direct outcome of Task 1.2 and T1.3 and will describe the artefacts that constitute the AIGOV Ecosystem. The aim of T1.2 is to identify all the involved stakeholders (e.g., citizens, businesses, etc.) who may interact with, or impacted by, an AI-based solutions as well as the roles that these stakeholders could potentially undertake. In addition, T1.3 aims to propose a stable ecosystem and explain how this relates to innovating and transforming public sector processes, practices, policies, and services.

This work follows an exploratory case study method to create the ecosystem. Specifically, the method that was followed in this deliverable comprises the following actions:

- Create a case study that focuses on transport data
- Conduct semi-structured interviews with the public servants Region of Central Macedonia, Greece
- Disseminate a questionnaire to public servants

These actions resulted in a set of artefacts that complementary constitute the AIGOV ecosystem, and the dependencies with each other. The ecosystem could serve as guidelines to adopt AI in Public Administration.

1 Introduction

The aim of this section is to introduce the background of the work pursued within Task1.1 "Stakeholders' needs and Data sources" and Task1.3 "The AIGOV Ecosystem" of the AIGOV project. The scope and the objective that the current document has set out to achieve are presented in sub-section 1.1. The intended audience for this document is described in sub-section 1.2 while sub-section 1.3 outlines the structure of the rest of the document.

1.1 Scope

The present document is the deliverable "The AIGOV Ecosystem" (henceforth, referred to as D1.2) of the AIGOV project. The main objective of D1.2 is to document the results of Task1.1 "Stakeholders' needs and Data sources" and Task1.3 "The AIGOV Ecosystem" of WP1.

1.2 Audience

The intended audience for this document is public administration, policy-makers, and anyone interested in deploying Artificial Intelligence in the public sector.

1.3 Motivation

Public sector data represent a vast and diverse resource that holds immense value for society [36]. Within this resource, dynamic data, which are real-time data generated by sensors, are particularly important. Recently, the European Commission recognized dynamic government data as highly valuable, with enormous potential for the economy, the environment, and society. However, collecting and disseminating these data present a range of challenges, including high variability and rapid obsolescence [45].

Fortunately, recent advances in technologies such as Artificial Intelligence (AI) offer the possibility of creating value-added intelligent applications that can unlock the potential of government data. AI has tremendous potential for public administration, offering the possibility of saving up to 1.2 billion hours and $41.1 billion annually [43]. By automating routine tasks, public servants can focus on high-value work and make better decisions, detect fraud, plan infrastructure projects, answer citizen queries, adjudicate bail hearings, triage healthcare cases, and provide innovative, personalized public services to citizens [43] [92].

This work proposes a holistic ecosystem for the implementation and evaluation of AI technologies, such as machine learning, deep learning, and natural language processing, in public administration. We present a case study focused on dynamic government data, specifically transport data, to explore the ecosystem components and their dependencies. The case was selected because dynamic data is an underexplored form of government data.

1.4 Structure

The structure of the document is as follows:

- Section 2 provides in detail the method used to create this deliverable.
- Section 3 is an introduction to Artificial Intelligence.
- Section 4 presents the case study that is based on dynamic traffic data.
- Section 5 presents the analysis of the questionnaire’s data.
- Section 6 presents the AIGOV Ecosystem.
- Finally, section 7 draws conclusions.

2 Method

In this section we present the method that we follow in order to achieve the objectives of the second task of WP1.

This study aims to identify the components of an ecosystem for deploying Artificial Intelligence (AI) in public administration and explore their interdependencies through a single exploratory case study. Such a study is useful for gaining insights into a poorly understood phenomenon and generating new theory or propositions about it [103].

For the case study, we examined three open traffic data sets: one from the city of Thessaloniki in Greece, another from the Attica region in Greece, and a third from Switzerland. To identify the ecosystem components, we used ”snowballing” to identify AI algorithms, technologies, methods, and cases from technical and policy reports, government documents, and research articles.

We also conducted a series of research activities to explore potential AI applications for the public sector. Firstly, we conducted semi-structured interviews with employees of the region of Central Macedonia, which is the second-largest region in Greece. The interviews provided valuable insights that we used to create user stories.

To gain a deeper understanding of how public servants perceive AI, we designed a questionnaire (see APPENDIX A). The questionnaire was distributed to 16 postgraduate students of the Master in Public Management program at the University of Macedonia in Greece, and we analyzed the responses.

The method followed in this deliverable is illustrated in Figure 1.

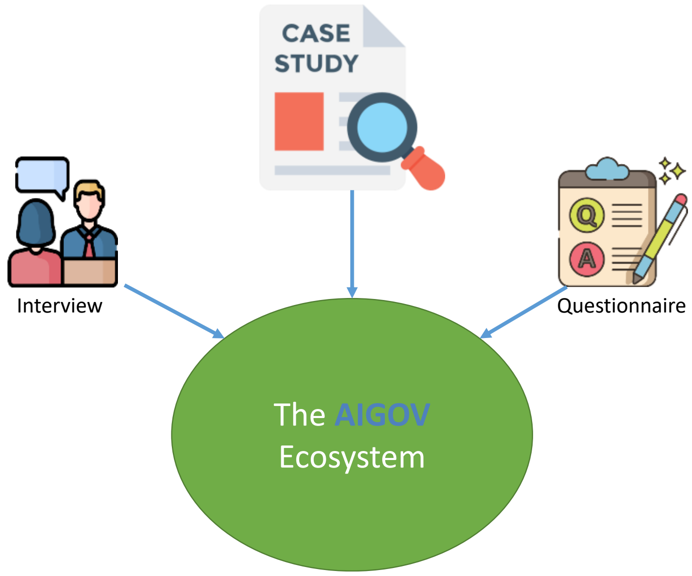

Figure 1 The method followed in this deliverable.

3 Background: Artificial Intelligence

3.1 Definition and Potential

Although Artificial Intelligence (AI) lack of an official definition, many definitions have been given by various global organizations over time. For example, one of the oldest definitions is the one that was given in 1956 by John McCarthy, who is considered to be the father of AI. John McCarthy define AI as *“the science and engineering of making intelligent machines”*. Since then more other definitions have also been proposed with many of them reflecting McCarthy’s approach. A typical example is the more detailed definition for AI systems provided by the Organisation for Economic Co-operation and Development – OECD) [70]:

“An AI system is a machine based system that can, for a given set of human defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. It does so by utilising machine and/or human based inputs to:

1. *perceive and/or analyse real and/or virtual environments;*
2. *abstract such perceptions/analyses into models manually or automatically; and*
3. *use model interpretations to formulate options for outcomes.*

AI systems are designed to operate with varying levels of autonomy.”

All around the world, 42 governments have endorsed this definition of AI, which views it as capable of mimicking human behavior and matching but not exceeding human intelligence [8]. In this context, literature occasionally refers to hypothetical general AI systems that could significantly outperform human abilities as "Artificial Super Intelligence," or "superintelligence" [7].

Additionally, the World Economic Forum defines AI systems in terms of their own system-based perspective as "systems that act by sensing, interpreting data, learning, reasoning, and choosing the best course of action" [100].

Another approach envisages AI from the perspective of designing systems: “the discipline of creating algorithms that can learn and reason” [70]. A similar definition of AI is provided by the High-Level Expert Group on Artificial Intelligence of the European Commission, which describes it as "systems that display intelligent behavior by analysing their environment and taking actions - with some degree of autonomy - to achieve specific goals" [19].

The British government also offers a definition that is broader in scope. According to this definition, AI is a "research field spanning philosophy, logic, statistics, computer science, mathematics, neuroscience, linguistics, cognitive psychology, and economics" that makes use of "digital technology to create systems capable of performing tasks commonly thought to require intelligence"[1]. Additionally, according to the Luxembourg regulatory authority, AI solutions should "focus on a small number of intelligent tasks and be used to assist humans in making decisions" [15].

As another option, the Institute of Electrical and Electronics Engineers (IEEE) Standards Association concentrates more on the Autonomous and Intelligence Systems (A/IS) than AI per se[2].

Although there is no consensus regarding the definition of AI, two perspectives are used to set the expectations on the level of AI's intelligence [8]:

- General AI. This first perspective is also known as "strong AI" or the "Artificial General Intelligence" (AGI) perspective.
- Narrow AI. This perspective is the more granular view of AI and is also known as "weak AI", "applied AI", or "Artificial Narrow Intelligence" (ANI).

All AI that has been used or that is being used now is Narrow, meaning that all AI algorithms and systems are accomplishing their tasks in a human-like way rather than being capable of performing or outperforming any task (intellectual or cognitive) a human can perform. Narrow AI makes use of the fact that while humans are still more adept at handling ambiguous situations or those requiring intuition, creativity, emotion, judgment, and empathy, computers are better at processing large amounts of data quickly and consistently as well as carrying out tasks based on logical and explicit rules [8].

Although AI has been already defined in 1956, it only recently started becoming so popular. The key factors that contributed to this recent rise of AI include:

- Data explosion [7] [102]. The explosion of data that are being produced the recent years within businesses, organisations, governments, and societies and that are coming from various sources and being of various types (e.g., IoT data, multimedia including text, audio, and video) makes AI possible and necessary as well. Especially in the public sector, many government operations relate to the generation and maintenance of citizens' registries (e.g., births, marriages). Governments also maintain tremendous amounts of other data including geospatial and weather data from satellites, property records, and health and safety records, among many others. In recent years, governments have increasingly pursued the publication of government data in machine-readable formats through open government data (OGD)

1 https://www.gov.uk/government/publications/understanding-artificial-intelligence/a-guide-to-using-artificial-intelligence-in-the-public-sector
2 https://standards.ieee.org/industry-connections/ec/autonomous-systems/

policies, and associated portals for datasets and Application Programming Interfaces (APIs). This contributes to the availability of data for AI systems to leverage. [7]

- Maturity [7]. A significant body of knowledge has been accumulated with many different projects launched over the last few decades. Old algorithms and models have been refined and new ones have emerged. Programming languages and frameworks have been developed and refined and many new applications created as more people become familiar with AI. For example, the idea of artificial neurons has been around since the 1940s, but the development of Deep Learning AI only took off during the last decade.
- Democratisation of computers and programming [7]. While technology has improved, it has also become available to a growing number of people. New users today are also more connected and better equipped to learn and exchange information about AI. Collaborative platforms and tools supported by vibrant communities are making programming and coding possible not only for experts and companies, but also for individuals from all backgrounds. For example, GitHub and Kaggle allow people to collaborate on digital solutions. This also allows bottom-up ideas and solutions to emerge in ways that were not often possible in the past. Freely available online courses and tutorials, including those provided by the public sector, also contribute to this democratisation.
- Better technology and increased processing power (especially in the cloud) [7] [102]. In the cloud, massive amounts of computing power are available on demand. The cost to process an AI application today is a small fraction of what it would have cost a decade ago, if it were even possible. Computers today are cheaper, have more computing power and take up much less physical space. This increase in processing power allows devices to run larger and more complex programmes, and process more data faster. Data storage costs have also decreased dramatically.
- New algorithm types including neural networks and deep learning that have come to maturity in the last decade or so [102]. These new algorithm types have been a primary driver behind such AI use cases as image, video and voice recognition, speech recognition and natural language processing, and natural language generation.
- Other, socio-political factors [102] like the growth of technology capabilities in China, the relative absence of data privacy standards in the United States, the massive acceptance of digital business offerings in many countries around the world, and spending by government bodies.

AI can be applied to various industries using a plethora of technology types. Figure 2 presents a heatmap for the volume of use cases per industry and technology type. The colors used indicate the number of related use cases; the darker the color the largest the number of cases. According to the Figure, most cases are related to:

- Using knowledge management in professional cases and the public and social sector
- Using Vision in the sectors of consumer goods and services, and public and social sector
- Using Speech recognition in the consumer goods and services sector
- Using Natural Language Processing (NLP) in the consumer goods and services sector
- Using analysis, optimization and prediction in the sectors: basic materials; financial services; technology and social sector; telco; transportation; utilities
- Using robotics and sensors in the sectors: basic materials; energy; industries; transportation; utilities

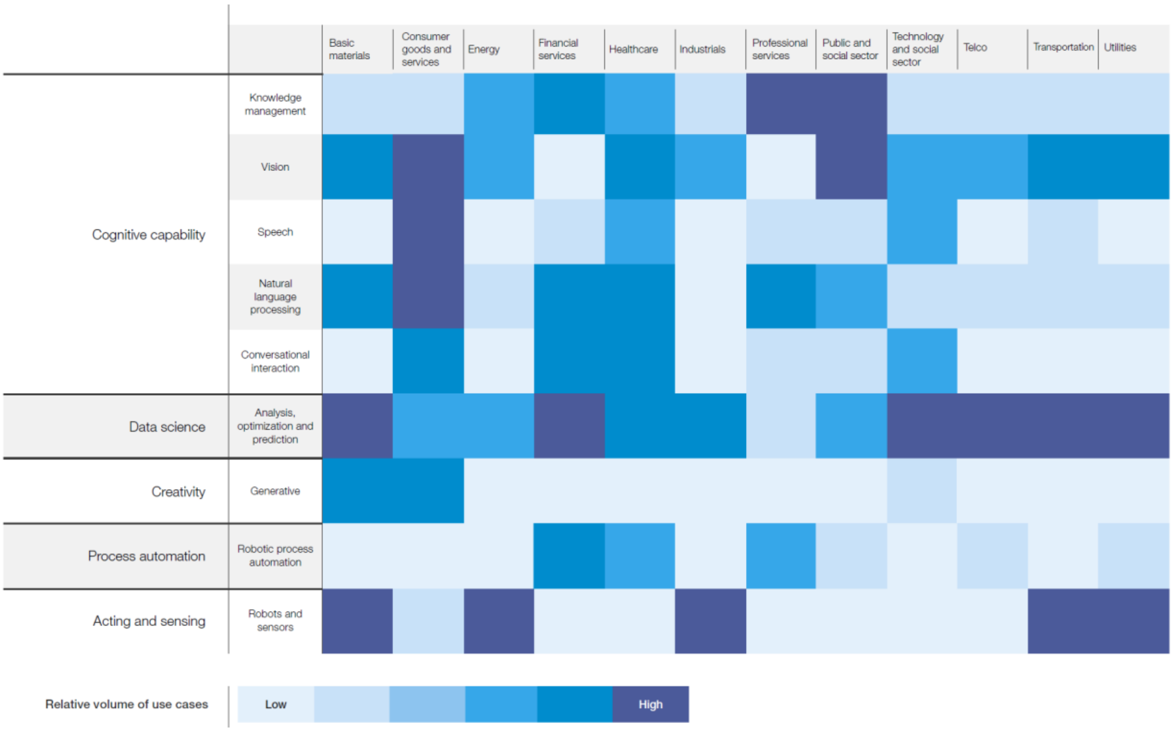

Figure 2 Common AI use cases by industry and technology type [102]

In addition, according to a recent McKinsey's report [61], the most commonly adopted AI case uses service-operations optimization followed by AI-based enhancement of products, contact-center automation, product-feature optimization, predictive service and intervention, customer-service analytics, creation of new AI-based products, customer segmentation, risk modelling and analytics, and fraud and debt analytics.

However, AI is not a general purpose solution which can solve every problem [94]. Current applications of AI focus on performing narrowly defined tasks. AI generally cannot: be imaginative; perform well without a large quantity of relevant, high quality data; infer additional context if the information is not present in the data.

3.2 Machine Learning

The most widely used form of AI is Machine learning (ML), a field at the intersection of computer science, mathematics and statistics. ML mainly focuses on developing algorithms based on large volumes of data and of various more or less complex structures of data, in order to make predictions, recommendations or decisions. It has contributed to various innovations including navigating autonomous vehicles, predicting infrastructure failures, or recommending products online.

The main types of ML are:

1. supervised learning. In supervised learning the AI model is trained based on past, labeled observations for a particular output. A typical example of supervised learning problem is the prediction of the prices of houses in a region based on labeled data that describe, for example, structural aspects of the houses (e.g., number of rooms etc.). This type of learning is generally useful in clearly defined problems where the content and structure of the data are sufficiently described. Supervised learning problems include regression and classification. The first is used to predict the value of a variable, while the latter to predict the group to which a new data point will be classified. The house prediction mentioned above is, for example, a regression problem, while predicting if a customer will churn is a classification problem.
2. unsupervised learning. In this type of ML the AI model is trained based on unlabeled and unclassified data in order to gain new insights on the data. A typical unsupervised learning problem is market segmentation that aims to cluster customers (e.g., of a superstore) based on their previously purchased products and consuming behavior (e.g., how frequent they order a product). This type of learning is generally used for unguided pattern discovery and is based on identifying common characteristics between different data points. A common approach to do so is using clustering.
3. reinforcement learning, which allows an AI model to learn as it performs a task in an environment. For example, a reinforcement learning model could be trained to predict the next move of an agent in a chess game. Specifically, reinforcement learning is based on getting feedback from the environment about an agent’s previous interactions. This helps the agent learn through the trial and error method, since error is “punished” while success is “rewarded: by the environment. The agent then automatically adjusts its behavior over time producing more refined actions. Reinforcement learning has recently grown popularity because of the advances in computing capabilities.
4. semi-supervised learning, which combines supervised and unsupervised learning. Specifically, semi-supervised learning trains the AI model using both labeled and unlabeled data.

The final area of Machine Learning is deep learning. Deep learning is based on neural networks that are algorithms designed to mimic the way the human brain processes information. Deep learning refers to deep neural network, which is a specific configuration where neurons are organized in multiple successive layers [44]. The main distinction with the classical ML algorithms lies in the design of deep learning algorithms, which is inspired by the biology of human brains. Indeed, deep learning is often discussed in conjunction with Artificial Neural Networks (ANN). The “depth” of an Artificial Neural Network relates to its number of hidden layers. Deep learning algorithms use ANNs which have two or more hidden layers.

3.3 Stakeholders

There are two sets of stakeholders for building and using AI [102]:

- Stakeholders that are required to deliver and manage AI projects, and
- Stakeholders whose trust is required to ensure the successful adoption of AI.

In the first instance, AI project and operational teams need to be built. As with many digital projects, it is important to balance technology and business skills with domain knowledge. Broadly speaking, these roles include:

- Data engineers: responsible for creating the nuts and bolts of data infrastructure and pipelines
- Business intelligence, insight and analytics professionals: responsible for ad hoc reporting and dashboards
- Data scientists: responsible for building predictive algorithms and ensuring the recommendations of the data team are statistically robust
- ML engineers: responsible for scaling data science models and putting them into production
- Data product managers: responsible for coordinating all the above resources as well as aligning with business stakeholders and engineering to deliver value
- Data governance specialists: responsible for managing and documenting data assets and ensuring they are ethical and compliant.
- Depending on the desirable outcome, it is critical to have experienced product managers in the team to ensure the successful delivery of solutions that satisfies the end users’ needs.

4 The Case of Dynamic Traffic Data

4.1 Data Collection

In order to create an ecosystem for deploying Artificial Intelligence (AI) in public administration, we present a case study that focuses on the usage of dynamic government data and, specifically, traffic data to create AI applications for the public sector. The case study uses three open traffic datasets from (i) the city of Thessaloniki in Greece, (ii) the region of Attica in Greece, and (iii) Switzerland. An overview of the three OGD datasets is presented in Table 2.

	Dataset 1	Dataset 2	Dataset 3
Location	Thessaloniki, GR	Attica region, GR	Switzerland
Access	open	open	open
Accessibility medium	files	files/API	files/API
Historical data	Yes	Yes	No
Aggregation level	minute	hour	minute
Anonymization	Not required	Not required	Not required
Data format	JSON, XML, CSV, KML, MAP	JSON	XML

Table 2 An overview of the three datasets with traffic data

4.1.1 Traffic Data from the City of Thessaloniki in Greece

The Smart Mobility Living Lab, one of the largest European mobility labs located in Thessaloniki, Greece is the data analysis and modelling laboratory of the Hellenic Institute of Transport (HIT). The lab hosts transportation and mobility related datasets generated by various both conventional and innovative data sources.

Among them are the datasets with open data from taxis and Bluetooth detectors in the urban area of Thessaloniki [8]. Specifically, two types of measurements are provided. The first one refers to floating car data including the speed measured by the GPS of over one thousand vehicles that operate in the city of Thessaloniki. This dataset is being updated almost real-time providing about 2,000 new records per minute. The second one includes aggregated vehicle detections of over 43 Bluetooth devices located in main road junctions of the city of Thessaloniki at a specific timeframe. Additional data included in the same dataset are trip trajectories with the sequence of locations or the origin and destination. The latest datasets are updated every 15 minute. All data can be acquired via proprietary APIs. Historical data are also available as text files.

4.1.2 Traffic Data from the Attica Region in Greece

Data.gov.gr serves as Greece's official data portal for Open Government Data. The portal comprises 49 datasets that span ten thematic areas, including the environment, economy, and transportation. Notably, the new version of the portal incorporates a free Application Programming Interface (API), enabling users to retrieve and access data via a graphical interface or code. Acquiring a token is needed to use the API through completing a registration process. The traffic data for the Attica region in Greece is sourced from traffic sensors that periodically transmit information regarding the number of vehicles and their speed on specific roads in Attica. To mitigate privacy concerns, the data is aggregated hourly and is updated every hour with a one-hour delay. Provided data measurements include the absolute number of the vehicles detected by the sensor during the hour of measurement along with their average speed in km per hour.

4.1.3 Traffic Data from Switzerland

The Open Transport Data Portal of Switzerland (ODPCH)[3] provides access to more than 40 datasets. Among them are datasets with real-time traffic data generated and collected by traffic sensors positioned in road segments throughout Switzerland. Historical data are not available. Data include the number of vehicles passing from specific locations, along with their average speed. The data are minutely aggregated and updated every minute and specifically, 20 seconds after the minute in Coordinated Universal Time (UTC) 0. Data are described using the DATEX II[4] standard for exchanging road traffic data. An access token is required to get limited access to these data through the corresponding API for six months. The API allows for submitting in total 260,000 requests in the six-month period, which actually, corresponds to the update interval of the data (one update per min).

4.2 Construction

Traffic data have the potential to aid policy-makers and public authorities in designing and managing transportation systems that are efficient, safe, environmentally friendly, and cost-effective. One way to achieve this is through predicting future traffic conditions. There are several methods commonly used for forecasting future traffic conditions, including traditional parametric methods such as Autoregressive Integrated Moving Average (ARIMA) [44], machine learning techniques such as Support Vector Machine (SVM) [3], and deep learning [12] [109].

The three datasets analyzed in this study contain sensor-generated traffic data, which often exhibit quality issues [90]. For example, the Attica dataset initially suffered from a high number of missing observations and anomalous values, although most of these issues have been resolved [12] [44] [52]. Various methods, such as time series analysis (e.g., Seasonal – Trend decomposition using Loess - STL),

[3] https://opentransportdata.swiss/en
[4] https://www.datex2.eu/

machine learning (e.g., Isolation Forest), and deep learning (e.g., Generative Adversarial Networks - GANs), can be used to identify anomalous values in the datasets [51]. However, handling missing and anomalous values requires making decisions on a case-by-case basis, depending on factors such as the level of aggregation for the temporal dimension. For example, missing values in the Thessaloniki and Switzerland datasets, which are minute-level aggregated, can be imputed using the average of the previous and next observations, while synthetic data can be used for the Attica dataset, which is hourly aggregated.

Integrating traffic data with other datasets, such as weather data, car accident data, can provide valuable insights and enhance the accuracy of AI models [57]. Using explainable AI techniques can also help understand the reasons behind anomalous values or model decisions [44].

Real-time data access is critical for applications that rely on dynamic data such as traffic intelligent systems. All three datasets are available in Open Government Data (OGD) portals and two of them can be accessed programmatically using an API. The Thessaloniki and Attica datasets are hourly updated, while the Zurich dataset is minutely updated.

Finally, selecting the appropriate AI algorithm for each dataset is crucial. In literature, machine learning approaches such as K-nearest Neighbour [78], and Bayesian models [86], and XGBoost (eXtreme Gradient Boosting) [97] have been used to predict traffic. Recently, the emerging development of deep learning and Graph Neural Networks have achieved state-of-the-art performance in traffic forecasting tasks [3] [12]. Data may also play a significant role for the selection of the AI algorithm. For instance, the level of granularity and other dataset-specific factors directly affect the quality of the AI model. For instance, the minute-level granularity of the Zurich dataset allows for more accurate traffic flow predictions in the near future, while the hourly-level granularity of the Attica dataset is better suited for predicting traffic flow at the hourly level.

4.3 Evaluation

To assess the performance of the AI model, it is crucial to conduct a performance-based evaluation. Depending on the selected algorithm, various metrics can be employed. For instance, previous studies have utilized metrics such as RMSE, MAPE, and MAE to evaluate the performance of Graph Neural Network models in traffic flow prediction tasks [12]. For the traffic datasets and Graph Neural Networks, the accuracy of the AI model can be affected by the density of traffic sensor locations. Specifically, the Thessaloniki and Attica datasets contain sensor measurements from urban areas, which can lead to the creation of denser graphs when analyzed using Graph Neural Networks. This, in turn, can result in more accurate deep learning algorithms. On the other hand, the Zurich dataset comprises sensor measurements primarily from highways, resulting in sparser graphs [12].

Moreover, explainability can be employed to interpret the decisions made by the AI model. For instance, SHAP has been used to explain the decisions of a neural network that predicts traffic for traffic light control [97] [79]. In this case, integrating external data such as weather and vehicle accident data can facilitate better understanding of the decisions made by the model.

4.4 Translation

Traffic forecasts can be used to anticipate future needs and allocate resources accordingly, such as managing traffic lights [68], opening or closing lanes, estimating travel time [79], and mitigating traffic congestion [4]. In order to understand the potential of creating AI applications for the public sector using traffic data, we interviewed a public servant of the Region of Central Macedonia, the second largest region in Greece. The primary objective of the interview was to generate user stories that effectively describe potential AI applications based on the traffic data. These applications should have the potential to streamline the region's operations and enhance the efficiency of its employees.

The interviews resulted in the three user stories described below, namely (i) Management of vehicle traffic in the wider urban area of Thessaloniki through traffic lights, (ii) Optimal route for scheduled checkpoints, and (iii) Optimal use of GPS Data.

4.4.1 Management of vehicle traffic in the wider urban area of Thessaloniki through traffic lights

The Department of Maintenance of Transport Projects, which belongs to the Technical Works Directorate of the region of Central Macedonia in Greece, is responsible for the operation and planning of traffic regulation in the urban web through traffic lights. An application useful for the region of Central Macedonia would be the utilization of traffic/mobility data, traffic load/free flow of vehicles, and real-time reprogramming of traffic regulation and vehicle emptying times in the central vehicle flows, with the aim of immediately relieving traffic congestion and relieving overloaded areas within 10 minutes by selecting and opening the appropriate vehicle flows with the goal of optimal traffic management, quality of life, more rational resource management, and reduction of vehicle emissions/pollution.

4.4.2 Optimal route for scheduled checkpoints

Many services of the Region of Central Macedonia/Greece (e.g.,Technical Works Department, Health Department, Veterinary Department, etc.), whose headquarters are located at the city of Thessaloniki carry out scheduled inspections/checks with teams of competent employees. In these scheduled inspections/checks, the teams of employees visit from 5 to 15 different points (depending on available time). The inspections/checks are carried out within the urban fabric of Thessaloniki and its surroundings, but often also within the wider region of Thessaloniki (as well as in neighboring regions, within the Municipality). An AI application useful for the traffic office and responsible drivers would be to provide them with a proposed route in order to visit the predetermined points. The proposed AI route will be the best possible suggestion in terms of distance/traffic loads/avoidance of bottlenecks, in relation to the evolution of traffic data over time, as theoretically some traffic flows are more congested at specific times. The aim of this application is to save human resources, reduce vehicle emissions/pollution, and provide better working conditions for employees.

4.4.3 Optimal use of GPS Data

A GPS system has been installed in all vehicles in all regions (Pieria, Pella, Imathia, Serres, Kilkis, Halkidiki, and Thessaloniki) of Central Macedonia. A useful AI application for the supervisors and responsible parties of these vehicles would be to create a notification/report system utilizing the GPS data of each vehicle, indicating, for example, if one of the vehicles is moving at a speed greater than the permitted speed limits, if it goes beyond the allowed routes and movement limits, if it is stationary for a long time so that it can be allocated to another service that needs a vehicle, etc. The goal is the rational management of the resources of the region of Central Macedonia.

5 Significance and Potential Usage of Artificial Intelligence (AI) in the public sector: Questionnaire Data Analysis

This Section analyses the responses to the questionnaire for the survey on the potential usage of Artificial Intelligence (AI) in the public sector. The questionnaire was disseminated to 16 postgraduate students of the Master in Public Management program at the University of Macedonia in Greece. An introduction to Artificial Intelligence (AI) and its potential uses in the public sector was done before disseminating the online questionnaire. The aim of this introduction was to brief explain the concept of AI and its main technologies and applications in the public sector. Examples were also presented.

The questionnaire consists of four questions pertaining to demographics, as well as six questions concerning the use and significance of AI in the public sector. All questions of the questionnaire can be found in (APPENDIX A).

The survey on the usage of Artificial Intelligence (AI) in the public sector yielded 14 responses from Greek students that attend the Master programme. The majority of them are women (71.4%) and the rest of them men (28.6%). Their age was either in the range 36-49 (57.1%), or in the range 50-64 (57.1%). Four of the respondents are occupied in a local authority (e.g., municipality or region), two in Public Employment Service (DYPA), two in a judicial public service, one in the Greek ministry of interior, one in a Citizens' Service Centre (KEP), one in a regional Social Welfare Centre, one in the Greek Ministry of Health, one in the Greek Social Security Organisation (EFKA), and one in another public service. Moreover, among the respondents, there are ten public employees, which include secretaries, three supervisors, and one chairman of a board. The following section gives a brief overview of the answers given to each survey question.

5.1 Analysis of the Responses to the Questionnaire

The participants believe that AI is in general important for the public sector (Figure 3). Specifically, ten of them (72%) believe that it is extremely important, two that it is very important, and two that is important. None of the participants reported that AI is of little importance or that is not important at all.

In addition in the question “Do you think Artificial Intelligence can affect the Public Sector in the future?”, the vast majority of the respondents (13 out of the 14 respondents or 92.9%) believe that AI will definitely affect the public sector in the forthcoming years. The rest of them are not in the situation to know if AI will affect the public sector in the future.

The next question is interested in knowing if the participants rely on their experience while making decisions or if they solely rely on specific rules and guidelines in their daily job routines. (“In your job, are you called upon to make decisions based on experience rather than solely applying specific rules or guidelines?”). Based on the responds, the majority of the participants (eight out of fourteen or

57.1%) confirm that their everyday job requires making such decisions. The use of AI technologies can greatly enhance the effectiveness of such decisions. Hence, the next question asks for providing examples of such decisions. The responds can be found in Table 3.

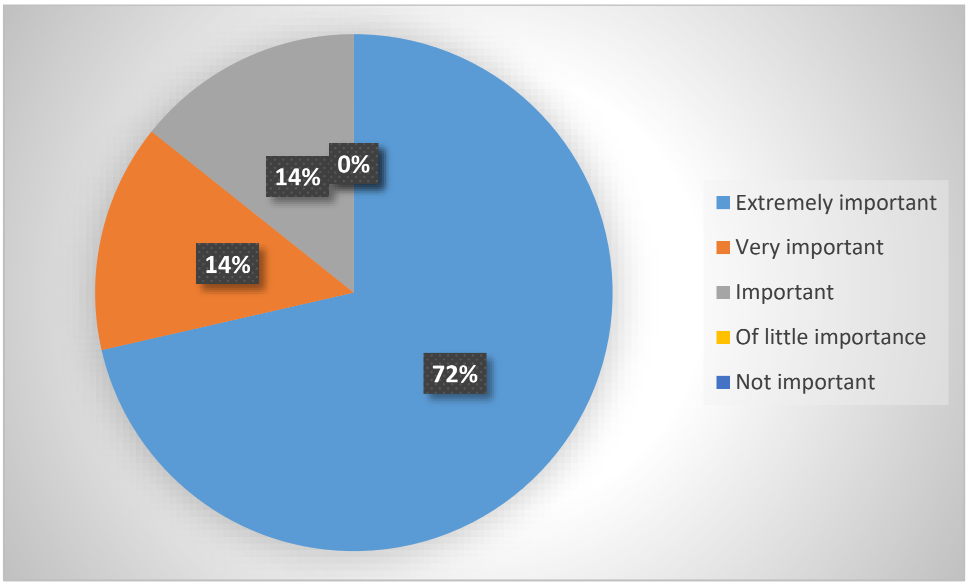

Figure 3 "How important do you think Artificial Intelligence is for the Public Sector?"

Responds
"While executing processes in the field of social provisions."
"When I am asked to collect justified funds for request, I have to rely on my experience as to which official holds them, although I am aware of the competent department."
"Pension distribution decisions/administrative issues administration for service boards."
"Predict how many of the requests will be accepted."
"Admission of a child in a structure for children of typical development with mild characteristics of disability. Experience says that you help these children instead of immediately sending them to a structure."
"Miscellaneous assignment decisions."
"Decisions on granting unemployment benefits and other financial assistance."
"Search for information on the issuance of fake driver's licenses."

Table 3 "Can you describe a situation where you are called upon at work to make a decision based on your experience?"

Following the preceding question, the participants were asked whether the public service they work for maintains historical data for these decisions. Most respondents (eleven out of fourteen or 78.6%) confirmed that there are related data available in their public service.

Finally, the last question asks for the opinion of the respondents on the major challenges and obstacles of AI when applied in the public sector. The following table (Table 4) presents the results. In summary, the participants are concerned about insufficient data, time to process data, software, funds, and proper education. In addition, one is concerned about the limitation in the number of available jobs. Someone also believes that the ethical values are not taken into account in AI applications, while another one believes that the diversity of the requests does not help the development of AI applications for the public sector.

Responds
“The lack of suitable software.”
“Getting data from the services.”
“Lack of information and education.”
“Ethical values are not taken into account/ lack of emotional criteria.”
“That we usually follow a well-trodden path. Nevertheless, I think that with proper information we could find fields where artificial intelligence could be applied.”
“Collect sufficient data.”
“Diversity of requests.”
“Especially in my area, human action or autopsy is NECESSARY, face-to-face assessment per case, etc.”
“Lack of funds.”
“I think it takes a long time for so much data to process.”
“Less jobs”
“Lack of familiarity with the technology and of necessary equipment.”

Table 4 "What do you think are the main challenges/obstacles to using Artificial Intelligence in your work?"

6 The AIGOV Ecosystem

In the previous section, a case study was presented, and its generalization has helped in the development of an ecosystem for deploying Artificial Intelligence (AI) in public administration.

6.1 The Key Stakeholders of Artificial Intelligence in the Public Sector

The key stakeholders of Artificial Intelligence (AI) in the public sector are those roles that either use the outputs of the AI systems in the public sector or are impacted by the AI systems and can be *internal* or *external*. Internal stakeholders include:

- public authorities and public organizations [14] [17] [69]
- public servants [18]
- regulators [18] [94]
- policy makers [32]
- IT stakeholders including agency's AI experts and developers [32] [20]
- public service designers [20]

External stakeholders include:

- citizens and residents [8] [17] [32] [20]
- customer advocacy groups [20]
- academic & scientific community [8] [32] [20]
- third parties [20]
- businesses [8] [17]
- organizations [8]
- practitioners [32]

6.2 Definition

The AIGOV Ecosystem, as shown in Figure 4, provides a comprehensive framework for collecting data from the public sector, using it to develop AI models with various AI technologies, and integrating different techniques for data pre-processing, federated learning, transfer learning, data augmentation, evaluation, explanation, and translation, and, evaluating and explaining the models and, finally, incorporating them in AI applications. The ecosystem is built upon four pillars, namely collection, construction, evaluation, and translation. Additionally, it comprises four key components, namely Data, AI Algorithms, AI Models, and AI Applications.

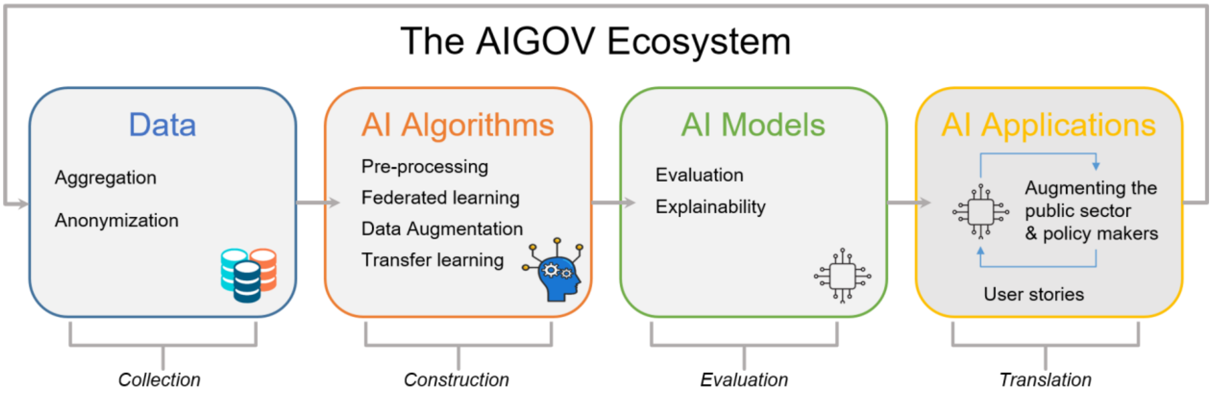

Figure 4 The AIGOV Ecosystem

6.3 Data

Governments must use data to improve public sector intelligence and, as a result, expand their capacity to create policies and services that are long-lasting, inclusive, and trustworthy [70]. There are several classifications of government data. For example government data can be classified based on (i) their type of access as open or closed, (ii) their structure as structured, unstructured, or semi-structured, (iii) the time of publication as real-time or batch data, (iv) the source of the data as internal and external [9] [19] [86].

Public sector data can be *open* or *closed*. Open Government Data (OGD) can be freely used and reused by the public. Currently, a large number of governments, public authorities and organizations and, statistical authorities worldwide provide Open Government Data (OGD) in their official data portals. For example, data.gov is the official open data portal in the U.S., https://data.europa.eu/ is the official data portal in Europe, while data.gov.gr is the Green open data portal. These portals provide to the public a plethora of datasets, usually organized in thematic categories. For example, the Greek open data portal classifies 49 datasets in nine topics namely Business and Economy, Crime and Justice, Education, Environment, Health, Society, Technology, Telecommunication, and Transport. At the same time, the European data portal currently hosts 1,604,307 European public sector datasets from 36 countries grouped in 14 categories. OGD are traditionally offered as downloadable files (e.g., PDF, CSV, XLSX). Some data portals also provide OGD as linked data [46] to facilitate their combination with other datasets and, hence, can be accessed by submitting SPARQL queries to a SPARQL endpoint. Recently, however, an increasing number of OGD portals also provide an Application Programming Interface (API) to allow the public programmatically and real-time access to OGD. On the contrary, closed government data include, for example, employee service records, employee performance assessment reports, confidential and secret government data. Closed government data Closed data access is limited to the data owner(s) and groups due to security limitations and relevant public policies.

In addition, public sector data may be *structured*, *unstructured* or *semi-structured*. Structured data have a well-defined structure so to be easily stored in databases. Towards this end, structured data require both a well-defined data model and a reliable data repository, typically in the form of a database. The data model serves as a blueprint for organizing various data elements and establishing their interrelationships, thereby ensuring consistency and coherence in the data. Structured data are stored in databases and data warehouses. Public health records organized and stored in columns and rows is a simple example of such data. In contrast to structured data, semi-structured data are characterized by a lack of rigid formal structure, yet it contains "tags" that facilitate the separation of data records or fields. Semi-structured data include documents in specific formats like XML and JSON documents, web logs and content from NoSQL databases. Finally, unstructured data has no discernible structure, as the name suggests. This includes various types of data such as text messages (e.g., emails, log files), click streams, photos, videos, and audio files, information about share rates (stock ticker), financial transactions, raw data from scientific research etc. [19]. With the growing complexity of datasets - ranging from structured to unstructured - the need for processing and analytical capabilities to effectively collect, manage, and analyze the data also increases significantly.

Government data can be *real-time* or *batch data*. Real-time data are commonly dynamic OGD produced by sensors (e.g., traffic data). The provision of this type of data as OGD has only recently begun. Dynamic Open Government Data (OGD) has the potential to create value-added services, applications, and generate high-quality job opportunities. Therefore, it is classified as high-value data (HVD) due to its significant societal, environmental, and financial benefits. Notwithstanding, the collection and reuse of this type of data entail overcoming several challenges. Dynamic data are characterized by their high variability and rapid obsolescence, necessitating prompt availability and regular updates to create value-added services and applications. Moreover, a recent study [44] demonstrated that dynamic traffic data face significant quality challenges, frequently attributed to sensor malfunctions resulting from adverse weather conditions [90]. On the contrary, batch data are historical data that are not immediately provided but provided some time after their collection.

Finally, data of the public sector can be *internal* or *external.* Internal data refers to the data that already have been produced by a public administration prior to the development of an AI system and exist inside the public organization's structure. These can be, for example, master or transactional data. On the contrary, external government data exist outside an organization's structure or is incorporated, specifically for the purpose of developing and training an AI system.

6.4 Collection

6.4.1 Data Aggregation

Government data may be individual-level or aggregated. The level of aggregation varies depending on the dataset. For example, apart from being aggregated (e.g., averaged) geographically (e.g., in the country level), data can also be aggregated based on time (e.g., in the hour level), demographic factors, (e.g., gender, race), and other relevant fields that are specific to a given dataset. In the realm of Open Government data and especially when it pertains to individual-level data, it is crucial to ensure that data is properly aggregated before publication. The initial level of aggregation may change depending on the requirements of problem.

6.4.2 Data Anonymization

Part of government data contains personal information protected by regulation (e.g., by the European Data Protection Regulation – GDPR [3636]) and cannot be used in its primary form. In order to protect privacy without compromising the usefulness of the data, it is essential to first anonymize this type of information prior to its analysis with AI. Towards this end, data can be either anonymized or pseudo-anonymized. Anonymization involves either encrypting or eliminating personally identifiable information from datasets to ensure that the individual's identity cannot be ascertained directly or indirectly. Whereas, pseudo-anonymization involves replacing personal information with a pseudonym or a unique identifier in a way that it is possible to re-identify individuals by combining the pseudonym with other available information provided that such supplementary information is separately maintained [57].

Traditional data anonymization operations include generalization, suppression, permutation, perturbation, and anatomization [59]. In addition, open source software for data anonymization are also available, such as AMNESIA[5] and ARX Data Anonymization[6], that use various anonymization techniques including k-Anonymity, k-Map, l-Diversity, t-Closeness, δ-Disclosure privacy, β-Likeness, δ-Presence, and (ε, δ)-differential privacy [91]. Since, however, literature has shown that using these operations does not ensure re-identification [21] [24], advanced synthetic data generation services have been proposed as an alternative such as synthetic data creation with Generative Adversarial Networks (GAN) [104]. AI-generated synthetic data are artificial data that have been generated based on an original dataset to mimic real-world observations and are an accurate representation of the original data.

6.5 Construction

Once data has been collected, it is essential to handle it appropriately to ensure that it can be effectively utilized in creating an AI model.

[5] https://amnesia.openaire.eu/index.html
[6] https://arx.deidentifier.org/

6.5.1 Artificial Intelligence Algorithms

Various types of Artificial Intelligence (AI) algorithms have been developed that can be used to analyse data and develop intelligent applications. Although the field of AI encompasses numerous subtypes, some of the most widely recognized and frequently utilized types of algorithms include [11]:

- Computer vision focuses on recognizing, tracking, and interpreting patterns and objects in visual data (e.g., images, videos). Applications include image and video analysis, object detection, and autonomous vehicles.
- Natural Language Processing (NLP) that enables understanding, interpreting, and generating human language. It is used in a wide range of operations, including text classification, sentiment analysis, language translation, chatbot development, and speech recognition.
- Speech Recognition converts spoken language into text or other machine-readable formats. Applications include virtual assistants, voice-enabled devices, and speech-to-text transcription.
- Knowledge-based systems are able to make decisions based on expert knowledge and domain-specific rules. They typically consist of a knowledge base with domain-specific knowledge and rules, and an inference engine, which uses this knowledge to make decisions and solve problems.
- Automated Planning allows generating plans or sequences of actions to achieve particular goals. It enables reasoning about the problem domain, generating plans, and executing them in a dynamic environment. Applications include robotics, manufacturing, logistics, and scheduling.

The selection of the proper AI algorithm from each sub-field of depends on the available data as well as on the requirements of the application.

6.5.2 Pre-processing

Data preprocessing plays a crucial role in converting raw data into a format that is compatible with Artificial Intelligence (AI) algorithms. This essential step involves various techniques and procedures that help to clean, transform, and organize data, making it easier for the AI system to extract meaningful insights and patterns. Data preprocessing methods include:

- Data cleaning: Data cleaning is a critical step in data preprocessing, as it ensures that the data is accurate, consistent, and reliable, making it suitable for use in AI models and other applications. Data cleaning involves various tasks including identifying missing values, i.e., cells or fields in the dataset that are empty or null and deciding how to handle them, e.g., impute or remove them, identifying and removing any duplicate rows or records. Data imputation can be performed by employing advanced AI methods including Generative Adversarial Nets (GAN) [53] [105]. In addition, data cleaning is also responsible for ensuring that data is consistent and follows a standard format or structure (e.g., for dates or units of measurement), identifying and correcting any errors or inconsistencies in the data (e.g., misspelled values), and identifying and handling (e.g., remove, impute) outliers or anomalous values, i.e., values that are significantly different from the rest of the dataset. Statistical analysis, machine learning, including synthetic data [108], is used for anomaly detection. Effective data cleaning is particularly critical for real-time data like traffic data generated by sensors, which must be promptly accessible without extensive pre-processing before publication.

- Data integration Government datasets can be leveraged by integrating them with other datasets, internal or external, to increase the value and effectiveness of AI applications [57]. This integration process can enhance the value and effectiveness of the applications. Data integration refers to combining data from multiple sources, formats, and structures into a single, consistent, accurate, and comprehensive view. Data integration requires extracting data from the various sources and transforming them into a common format. Some works have already explored the integration and exploitation of government data (e.g., [46], [107]). However, the heterogeneity of the original data from different sources presents various challenges, including legal, structural, or other issues [57]. Even if government data are available in formats that facilitate integration (such as linked data), addressing structural challenges is still necessary [46].
- Data augmentation: This technique involves generating new data by applying various transformations, such as rotation, translation, or scaling, to the existing data to increase the size of the dataset. Data augmentation is a technique used in machine learning to increase the size of a dataset by generating new, synthetic data from the existing data. The aim of data augmentation is to improve the performance and robustness of the machine learning model by exposing it to a larger and more diverse set of training examples [106]. Data augmentation techniques involve applying a range of transformations to the existing data, such as flipping, rotating, scaling, cropping, or adding noise to images. For example, in image classification, data augmentation can be used to generate new images by flipping, rotating, or cropping the existing images. This creates a larger and more diverse dataset, which can improve the accuracy and generalization of the model. Data augmentation can also be used in other types of data, such as text, audio, or video. For example, in natural language processing, data augmentation can be used to generate new sentences by replacing words with synonyms or shuffling the word order. Data augmentation is particularly useful in scenarios where the dataset is small, or when the model is prone to overfitting. By generating new, synthetic data, data augmentation helps to expose the model to a more diverse set of training examples, making it more robust and generalizable.
- Federated learning. Federated learning is a recent AI technique that has been introduced by Google and has been successfully used in many research areas and industries [54]. Federating

learning allows training the model on decentralized data sources without the need for centralized data storage. In federated learning, the data remains on the user's device or edge servers, and the model is trained locally on each device. The updated model parameters are then sent back to a central server where they are aggregated to create a global model. The main advantage of federated learning is that it training machine learning models without requiring users to upload their data to a central location, thereby protecting their privacy. This is particularly useful in scenarios where data privacy is critical, such as in healthcare or financial applications as well as in scenarios where the data is too large or too sensitive to be stored in a central location. Within the scope of the public sector, federated learning has been employed, for example, to facilitate the development of smart cities services [46], in healthcare to develop predictive models using sensitive patient data [80], in transportation to improve traffic management by combining data from multiple sources such as traffic sensors, cameras, and GPS data [9], and in education to collect and analyze students' behavior data and discover how students learn [40].

- Transfer learning is a machine learning technique that involves leveraging knowledge gained from one task to improve the performance of another related task. Specifically, in transfer learning, instead of building a model from scratch, a model that has been pre-trained in a dataset is used to train a new model on a different (but related) dataset. The pre-trained model has already learned a set of features from a large dataset, and these features can be used as a starting point for learning new features on a smaller dataset. Examples of already existing pre-trained models include You Only Loo Once (YOLO) [80], a pre-trained model for object detection, and Bidirectional Encoder Representations from Transformers (BERT) [21], a family of pre-trained NLP models. Transfer learning is a process that saves time and computational resources since the new model does not have to learn everything from scratch and is particularly useful in scenarios where the size of the dataset is small or when there is a scarcity of labeled data. The "Hugging Face"[7] provides a library of a wide range of pre-trained NLP models including BERT.

These techniques are often combined in various ways depending on the specific requirements of the problem.

6.6 Evaluation of Artificial Intelligence Models

Once the AI algorithm is selected and data have been successfully pre-processed, the AI model will be created by training the algorithm on the data. This process involves feeding the model with input data and adjusting its parameters (hyperparameter tuning) to minimize the error between its predictions

[7] https://huggingface.co/

and the actual outcomes. Hyperparameter tuning can be done with methods such as include grid search, Random search, bayesian optimization, gradient-based optimization, and Ensemble-based methods as well as with cross-validation. Cross-validation partitions the data into training and validation sets and iteratively evaluates the model's performance on different subsets of the data and with different values for the hyperparameters. Cross-validation can be also employed in training to detect and avoid overfitting of the model.

The model needs to be evaluated on a separate validation dataset to ensure its generalizability and performance and ensure that they are effective and reliable in their intended application. Towards this end, several evaluation metrics can be used depending on the type of algorithm. The performance of the model, for example, can be assessed through various metrics such as precision, recall, F1-score, logarithmic loss, the Area Under Curve (AUC), Mean Absolute Error (MAE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE). These metrics can help assess the model's performance in detecting true positives and minimizing false positives and false negatives.

The efficiency of the AI model is also important and involves measuring its speed and computational requirements. This is important for real-time applications where the model needs to make quick predictions (e.g., in cases where dynamic data generated by sensors are used to create the model). The robustness of the model could also be evaluated in order to assess its ability to perform well under various conditions, such as changes in input data or noisy environments.

In recent years, the importance of explainability in AI models has grown significantly since it helps to improve the understanding of the model's decisions, promoting transparency and trust in the results. This is crucial for the public sector where transparency is a requirement. To achieve this goal, various methods have been developed to explain the decisions of both supervised and unsupervised AI models. There are various methods that can be used for explaining AI models. For example, LIME [81] and SHAP [58], which stands for SHapley Additive exPlanation, can be used to explain machine learning predictions. In addition, Class Activation Maps (CAMs) [110] can be utilized to explain Convolutional Neural Network decisions. For neural networks, Global Attribution Mappings (GAMs) [44] can be also employed. These techniques can provide insight into the factors that the AI model is considering when making decisions, allowing users to better understand and interpret the results.

6.7 Translation

Artificial Intelligence (AI) models can be deployed in real-world, intelligent applications for the public sector. AI applications, for example, have the potential to streamline decision-making processes in the public sector. This can be achieved through either fully autonomous decision-making or by providing decision-makers with AI-driven insights and recommendations to aid in the decision-making process. In this sense, humans and machines, rather than competing, could benefit from mutual collaboration and potentially solve problems and achieve better outcomes than each could on their own [15]. Figure 5 shows the different levels or stages of automation that can be reached using AI applications. AI applications that are built based on government data are able to enhance the efficacy of public services [46], such as the implementation of Integrated Public Services (IPS) [88] that are co-created and continuously evaluated through feedback loops by the public sector.

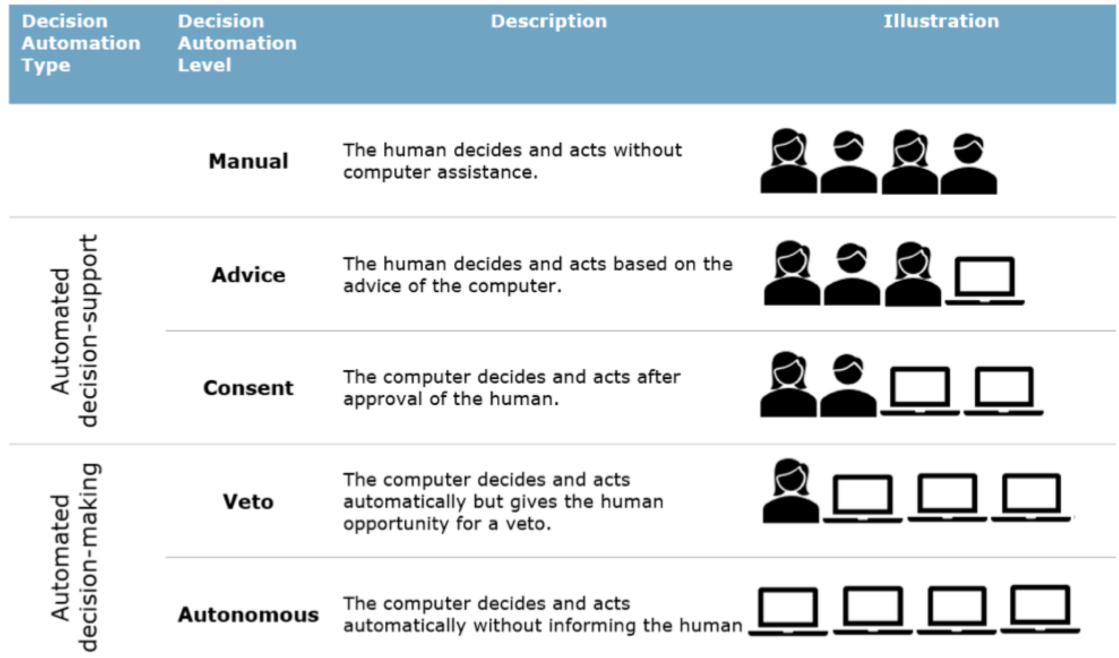

Figure 5 Levels and types of decision automation processes of the public sector [4] [76]

AI applications in the public sector can be classified based on the type of public services in [41]:

- Democratic process
- Tailored solutions
- Process optimization
- Maintenance
- Inspection and enforcement
- (Crime) investigation
- Knowledge and archive
- Forecasting and policy development

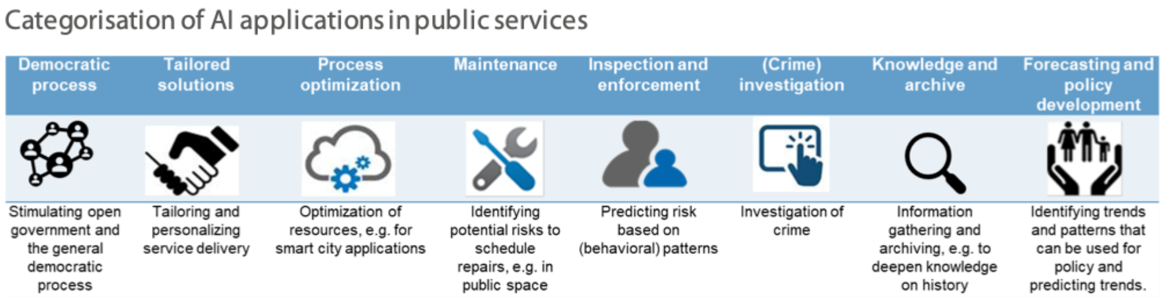

Figure 6 Categorization of AI applications in public services

User stories are essential in designing AI applications, as they ensure that end-users, such as public servants and policymakers, have their needs and preferences fully understood and incorporated into the development process. They serve as a link between data and applications, allowing for the identification of new AI-based applications that can benefit the public sector. User stories can be pulled by the end user who discovers the need for an AI application that could potentially improve their work, or they can be pushed by available public sector data. However, public sector data may not always be of high quality, making them unreliable and inconsistent sources of information. They may contain inaccuracies, errors, and missing values, rendering them unsuitable for developing trustworthy AI applications. Additionally, some public services may not store their data or store them in hardcopies or formats that are not machine-readable. In this case, data themselves can be used as a starting point, and high-quality data can drive the inspiration for the need for an AI application in the public sector. Regardless of whether the user story is pulled or pushed, the ultimate goal is to ensure that the AI application meets the needs of the end-user while utilizing reliable and accurate data.

7 Conclusions

This deliverable proposes an ecosystem for deploying Artificial Intelligence (AI) in public administration. The ecosystem comprises four main artefacts: Data, AI algorithms, AI models, and AI applications, built upon three pillars: collection, construction, evaluation, and translation. The data collection process should consider the nature of the source data and the requirements of the AI algorithm, and the collected data should be properly pre-processed and augmented to mitigate the detrimental effects of small sample sizes. Federated learning methods can be employed to overcome data privacy issues. The created AI model should be evaluated for its performance, efficiency, and robustness, and its explainability should be ensured by employing various explainability methods. Finally, the AI model should be integrated into an AI application co-created with the public sector employees, where user stories can be used to describe public sector scenarios that use AI applications.

The AIGOV ecosystem is a result of an exploratory case study that uses three traffic measurement datasets. The artefacts are complementary, and the selection of the AI algorithm depends not only on the available data but also on the needs of the AI application. The purpose of the AI model will define the AI application that will be used. The requirements for an AI application can be either pulled by the public sector employees or pushed by the available data, where data may create new needs for applications.

We anticipate that the ecosystem for deploying AI in public administration can serve as high-level guidelines for adopting AI in the public sector.

APPENDIX A

Χρήση της Τεχνητής Νοημοσύνης στον Δημόσιο Τομέα

Ερωτηματολόγιο στα πλαίσια του ερευνητικού έργου "AIGOV", Αριθμός Πρότασης 2412, 2η Προκήρυξη Ερευνητικών Έργων ΕΛ.ΙΔ.Ε.Κ. για την ενίσχυση Μελών ΔΕΠ και Ερευνητών/τριών.

Σημαντικές Επισημάνσεις:
1. Οι απαντήσεις σας είναι εμπιστευτικές και η ανωνυμία σας θα τηρηθεί με αυστηρότητα. Δεν διατηρείται πουθενά το όνομα σας ή οποιοδήποτε άλλο προσωπικό στοιχείο στο ερωτηματολόγιο.
2. Η συμπλήρωση και η επιστροφή του ερωτηματολογίου αποτελεί αυτομάτως την αποδοχή σας να συμμετάσχετε στην έρευνα. Μπορείτε να αποσυρθείτε από την έρευνα οποιαδήποτε στιγμή το επιθυμείτε.

* Απαιτείται

1. Ποιό φύλο σας προσδιορίζει; *

   Να επισημαίνεται μόνο μία έλλειψη.

   - Άνδρας
   - Γυναίκα
   - Προτιμώ να μην απαντήσω

2. Σε ποιό ηλικιακό εύρος ανηκετε; *

   Να επισημαίνεται μόνο μία έλλειψη.

   - 18- 35
   - 36 - 49
   - 50 - 64
   - άνω των 65

3. Σε τι είδους υπηρεσία του δημόσιου τομέα εργάζεστε; *

4. Ποιός είναι ο ρόλος σας στην εν λόγω υπηρεσία; *

5. Πόσο σημαντική πιστεύετε πως είναι η Τεχνητή Νοημοσύνη για τον Δημόσιο Τομέα; *

Να επισημαίνεται μόνο μία έλλειψη.

- Εξαιρετικά σημαντική
- Πολύ σημαντική
- Αρκετά σημαντική
- Λίγο σημαντική
- Καθόλου σημαντική

6. Πιστεύετε ότι η Τεχνητή Νοημοσύνη μπορεί να επηρεάσει τον Δημόσιο Τομέα στο μέλλον; *

Να επισημαίνεται μόνο μία έλλειψη.

- Ναι
- Όχι
- Δεν γνωρίζω

7. Στην εργασία σας καλείστε να λάβετε αποφάσεις με βάση την εμπειρία και όχι εφαρμόζοντας αποκλειστικά συγκεκριμένους κανόνες ή οδηγίες; *

Να επισημαίνεται μόνο μία έλλειψη.

- Ναι
- Όχι

8. Μπορείτε να περιγράψετε μια περίπτωση που καλείστε στην εργασία σας να λάβετε απόφαση με βάση την εμπειρία σας; *

9. Διατηρεί η υπηρεσία σας δεδομένα για την εν λόγω περίπτωση; *

Να επισημαίνεται μόνο μία έλλειψη.

- Ναι
- Όχι

10. Ποιές πιστεύετε ότι είναι οι βασικότερες προκλήσεις/εμπόδια στη χρήση της Τεχνητής Νοημοσύνης στην εργασία σας; *

References

1 Accenture (2017). *How AI Boosts Industry Profits and Innovation.* https://www.accenture.com/fr-fr/_acnmedia/36dc7f76eab444cab6a7f44017cc3997.pdf

2 Accenture (2018*). Artificial Intelligence, genuine impact: Public Services in the Era of Artificial Intelligence.* https://www.intelligentautomation.network/events-aiforgovernment/downloads/artificial-intelligence-genuine-impact-public-services-in-the-era-of-artificial-intelligence

3 Agafonov, A.: *Traffic flow prediction using graph convolution neural networks*. In: 2020 10th International Conference on Information Science and Technology (ICIST). pp. 91–95 (2020). https://doi.org/10.1109/ICIST49303.2020.9201971

4 Arciszewski, H.F.R., Greef, T.E. & Delft, J.H. van (2009, November). *Adaptive Automation in a Naval Combat Management System*. In: IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, 39(6), 1188-1199.

5 Ata, A., Khan, M.A., Abbas, S., Khan, M.S., Ahmad, G.: *Adaptive iot empowered smart road traffic congestion control system using supervised machine learning algorithm*. The Computer Journal 64(11), 1672–1679 (2021).

6 Australian Government (2019). Australia's Artificial Intelligence Ethics Framework. https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework

7 Australia Securities & Investments Commission (2022). *Corporate Plan 2022–26 Focus 2022–23. https://download.asic.gov.au/media/v3vhdqiw/asic-corporate-plan-2022-26-focus-2022-23-published-22-august-2022.pdf*

8 Ayfantopoulou, G., Salanova Grau, J.M., Tzenos, P., Tolikas, A.: Open data from taxis and bluetooth detectors to extract congestion and mobility patterns in thessaloniki. Data in Brief 47, 108899 (2023). https://doi.org/https://doi.org/10.1016/j.dib.2023.108899

9 Balbin, P.P., Barker, J.C., Leung, C.K., Tran, M., Wall, R.P., Cuzzocrea, A.: *Predictive analytics on open big data for supporting smart transportation services.* Procedia Computer Science 176, 3009–3018 (2020). https://doi.org/10.1016/j.procs.2020.09.202, knowledge-Based and Intelligent Information and Engineering Systems: Proceedings of the 24th International Conference KES2020

10 Ballester, O.: An artificial intelligence definition and classification framework for public sector applications. In: DG.O2021: The 22nd Annual International Conference on Digital Government Research. p. 67–75. DG.O'21,

Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3463677.3463709

11 Berryhill, J., et al. (2019). *Hello, World: Artificial intelligence and its use in the public sector*, OECD Working Papers on Public Governance, No. 36, OECD Publishing, Paris, *https://doi.org/10.1787/726fd39d-en.*

12 Brimos, P., Karamanou, A., Kalampokis, E., Tarabanis, K.: Graph neural networks and open government data to forecast traffic flow. Information p. [Accepted for publication] (2023)

13 Bughin, J., Song, J., Manyika, J., Chui, M. and Joshi, R. (2018). *Notes from the AI Frontier: Modeling the Impact of AI on the World Economy*. Report, McKinsey & Company, Washington, DC. https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-aifrontier-modeling-the-impact-of-ai-on-the-world-economy# .

14 CAHAI (Ad Hoc Committee on Artificial Intelligence) (2021). *Artificial Intelligence in Public Sector*. Council of Europe, Policy Development Group. CAHAIPDG(2021)03 https://rm.coe.int/cahai-pdg-2021-03-subwg2-ai-in-public-sector-final-draft-12032021-2751/1680a1c066

15 Carter, S., Nielsen, M.: Using artificial intelligence to augment human intelligence. Distill 2(12), e9 (2017)

16 Commission de Surveillance du Secteur Financier (2018). *Artificial Intelligence. Opportunities, Risks and Recommendations for the Financial Sector. https://www.cssf.lu/wp-content/uploads/files/Publications/Rapports_ponctuels/CSSF_White_Paper_Artificial_Intelligence_201218.pdf*

17 CSSF (2018), Artificial Intelligence: Opportunities, Risks and Recommendations for the Financial Sector, Luxembourg, Commission de Surveillance du Secteur Financier, https://www.cssf.lu/wp-content/uploads/files/Publications/Rapports_ponctuels/CSSF_White_Paper_Artificial_Intelligence_201218.pdf

18 Deloitte AI Institute (2021). *Scaling AI in Government. How to reach the heights of Enterprisewide adoption of AI. https://www2.deloitte.com/content/dam/insights/articles/us164831_cgi-state-of-ai-for-gov_/DI_CGI-state-of-ai-for-gov.pdf*

19 Desouza, K.C., Jacob, B.: Big data in the public sector: Lessons for prac-titioners and scholars. Administration & Society 49(7), 1043–1064 (2017). https://doi.org/10.1177/0095399714555751

20 Desouza, K. (2021). Artificial intelligence in the public sector: A maturity model. *IBM Center for The Business of Government, Washington, DC, USA*.

https://www.businessofgovernment.org/report/artificial-intelligence-public-sector-maturity-model

21 Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: *Bert: Pre-training of deep bidirectional transformers for language understanding.* ArXiv abs/1810.04805 (2019)

22 Eggers, W.D., Schatsky, D., Viechnicki, P.: Ai-augmented government. Using cognitive technologies to redesign public sector work. Deloitte insights (2017)

23 El Emam, K., Buckeridge, D., Tamblyn, R., Neisa, A., Jonker, E., Verma, A.: The re-identification risk of canadians from longitudinal demographics. BMC Medical Informatics and Decision Making 11(1), 46 (Jun 2011). https://doi.org/10.1186/1472-6947-11-46

24 Erlich, Y., & Narayanan, A. (2014). Routes for breaching and protecting genetic privacy. *Nature Reviews Genetics*, *15*(6), 409-421.

25 European Commission (2018a). *EU Member States sign up to cooperate on Artificial Intelligence. https://digital-strategy.ec.europa.eu/en/news/eu-member-states-sign-cooperate-artificial-intelligence*

26 European Commission (2018b). *Coordinated Plan on Artificial Intelligence*. https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence

27 European Commission (2019a), A Definition of AI: Main Capabilities and Disciplines, definition developed for the purpose of the AI HLEG's deliverables, Independent High-Level Group on Artificial Intelligence set up by the European Commission, Brussels, https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligencemain-capabilities-and-scientific-disciplines

28 European Commission (2019b). *Ethics guidelines for trustworthy AI. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai*

29 European Commission (2020a). *Commission Report on safety and liability implications of AI, the Internet of Things and Robotics. https://ec.europa.eu/info/publications/commission-report-safety-and-liability-implications-ai-internet-things-and-robotics-0_en*

30 European Commission (2020b). *White Paper on Artificial Intelligence: a European approach to excellence and trust. https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en*

31 European Commission (2021). *Proposal for a Regulation laying down harmonised rules on artificial intelligence. https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence*

32 European Commission, Joint Research Centre, Manzoni, M., Medaglia, R., Tangi, L. (2022a). *AI Watch, road to the adoption of artificial intelligence by the public sector: a handbook for policymakers, public administrations and relevant stakeholders,* Publications Office of the European Union. https://data.europa.eu/doi/10.2760/288757

33 European Commission (2022b). *Natural Language Processing for Public Services*. https://joinup.ec.europa.eu/collection/catalogue-services/document/study-natural-language-processing-public-services

34 European Parliament (2017). *Civil Law Rules on Robotics. https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_EN.html*

35 European Parliament (2020). *Civil liability regime for artificial intelligence. https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_EN.html#title1*

36 European Parliament, European Council: Directive (eu) 2019/1024 of the european parliament and of the council of 20 june 2019 on open data and the re-use of public sector information (recast). Off. J. Eur. Union 172, 56–83 (2019)

37 European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council. Office Journal of the European Union. 2016. Available online: https://op.europa.eu/en/publication-detail/-/publication/3e485e15-11bd-11e6-ba9a-01aa75ed71a1

38 Fatima, S., Desouza, K. C., & Dawson, G. S. (2020). *National strategic artificial intelligence plans: A multi-dimensional analysis.* Economic Analysis and Policy, 67, 178-194. https://doi.org/10.1016/j.eap.2020.07.008

39 G20 (2019). *G20 Ministerial Statement on Trade and Digital Economy. https://www.mofa.go.jp/files/000486596.pdf*

40 Guo, S., Zeng, D.: Pedagogical data federation toward education 4.0. In: Proceedings of the 6th International Conference on Frontiers of Educational Technologies. p. 51–55. ICFET ’20, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3404709.3404751

41 Hoekstra, M, Chideock, C. & Veenstra, A.F. van (2021). Quickscan AI use in public services II. https://www.rijksoverheid.nl/documenten/rapporten/2021/05/20/quickscan-ai-in-publieke-dienstverlening-ii

42 IBM (2021). IBM Institute for Business Value—Research, reports, and insights.. https://www.ibm.com/thought-leader¬ship/institute-business-value.

43 IBM Center for The Business of Government (2021). Artificial Intelligence in the Public Sector: A Maturity Model. https://www.businessofgovernment.org/report/artificial-intelligence-public-sector-maturity-model

44 Ibrahim, M., Louie, M., Modarres, C., Paisley, J.: Global explanations of neural networks: Mapping the landscape of predictions. In: Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society. pp. 279–287 (2019)

45 Jacob, V., Song, F., Stiegler, A., Rad, B., Diao, Y., Tatbul, N.: Exathlon: A benchmark for explainable anomaly detection over time series. Proceedings of the VLDB Endowment (PVLDB) (2021)

46 Janssen, M., Charalabidis, Y., Zuiderwijk, A.: *Benefits, adoption barriers and myths of open data and open government*. Information systems management 29(4), 258–268 (2012)

47 Jiang, J.C., Kantarci, B., Oktug, S., Soyata, T.: *Federated learning in smart city sensing: Challenges and opportunities*. Sensors 20(21) (2020)

48 Kalampokis, E., Karamanou, A., Tarabanis, K.: Interoperability conflicts in linked open statistical data. Information 10(8) (2019). https://doi.org/10.3390/info10080249

49 Karamanou, A., Kalampokis, E., Tarabanis, K.: Linked open government data to predict and explain house prices: the case of scottish statistics portal. Big Data Research 30, 100355 (2022)

50 Karamanou A., Brimos P., Kalampokis E., Tarabanis K. (2022) *Exploring the Quality of Dynamic Open Government Data Using Statistical and Machine Learning Methods*. Sensors. 2022; 22(24):9684. https://doi.org/10.3390/s22249684

51 Karamanou, A., Brimos, P., Kalampokis, E., Tarabanis, K.: Exploring the quality of dynamic open government data using statistical and machine learning methods. Sensors 22(24) (2022). https://doi.org/10.3390/s22249684

52 Karamanou, A., Brimos, P., Kalampokis, E., Tarabanis, K.: Exploring the quality of dynamic open government data for developing data intelligence applications: The case of attica traffic data. In: Proceedings of the 26th Pan-Hellenic Conference on Informatics. p. 102–109. PCI '22, Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3575879.3575974

53 Kim, J., Tae, D., Seok, J.: *A survey of missing data imputation using generative adversarial networks*. In: 2020 International conference on artificial intelligence in information and communication (ICAIIC). pp. 454–456. IEEE (2020).

54 Konecny, J., McMahan, H.B., Yu, F.X., Richtarik, P., Suresh, A.T., Bacon, D.: *Federated learning: Strategies for improving communication efficiency*. In: NIPS Workshop on Private Multi-Party Machine Learning (2016)

55 Kumar, P.B., Hariharan, K., et al.: *Time series traffic flow prediction with hyper-parameter optimized arima models for intelligent transportation system*. Journal of Scientific & Industrial Research 81(04), 408–415 (2022).

56 LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning*. Nature, 521(7553), 436–444. doi:10.1038/nature14539

57 Lee, J.S., Jun, S.P.: Privacy-preserving data mining for open government data from heterogeneous sources. Government Information Quarterly 38(1), 101544 (2021)

58 Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model predictions. In: Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (eds.) Advances in Neural Information Processing Systems. vol. 30. Curran Associates, Inc. (2017)

59 Majeed, A., Lee, S.: Anonymization techniques for privacy preserving data publishing: A comprehensive survey. IEEE Access 9, 8512–8545 (2021). https://doi.org/10.1109/ACCESS.2020.3045700

60 Marcia, V., Desouza, K. C., (2021). *The EU Path Towards a Regulation on Artificial Intelligence*. Brookings. https://www.brookings.edu/blog/techtank/2021/04/26/the-eu-path-towards-regulation-on-artificial-intelligence

61 McKinsey Global Institute (2021). *The state of AI.* https://www.mckinsey.com/capabilities/quantumblack/our-insights/global-survey-the-state-of-ai-in-2021

62 McCorduck, P., & Cfe, C. (2004). *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. CRC Press.

63 Mehr, H. (2017). *Artificial Intelligence for Citizen Services and Government*. Harvard Ash Center for Democratic Governance and Innovation, 19. https://ash.harvard.edu/publications/artificial-intelligence-citizen-services-and-government.

64 Microsoft (2020). *Artificial Intelligence in the Public Sector. European Outlook for 2020 and Beyond. How 213 Public Organizations Benefit from AI. https://info.microsoft.com/rs/157-GQE-382/images/EN-CNTNT-eBook-artificial-SRGCM3835.pdf*

65 Miller, H., Stirling, R. (2019). *Government Artificial Intelligence Readiness Index 2019. https://ec.europa.eu/futurium/en/system/files/ged/ai_readiness_index_2019__0.pdf*

66 Ministry of Business, Innovation and Employment (2017). *Strategic Science Investment Fund Investment Plan 2017 – 2024.* https://www.mbie.govt.nz/dmsdocument/1435-strategic-science-investment-fund-investment-plan-pdf

67 Misuraca, G., & van Noordt, C. (2020). *Overview of the use and impact of AI in public services in the EU* (JRC EUR 30255). Publications Office of the European Union. https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/ai-watch-artificial-intelligence-public-services

68 Navarro-Espinoza, A., L´opez-Bonilla, O.R., Garc´ıa-Guerrero, E.E., Tlelo-Cuautle, E., L´opez-Mancilla, D., Hern´andez-Mej´ıa, C., Inzunza-Gonz´alez, E.: Traffic flow prediction for smart traffic lights using machine learning algorithms. Technologies 10(1) (2022). https://doi.org/10.3390/technologies10010005

69 Nili, A., Barros, A., & Tate, M. (2019). The public sector can teach us a lot about digitizing customer service. MIT Sloan Management Review, 60(2), 84-87.

70 van Ooijen, C., Ubaldi, B., Welby, B.: A data-driven public sector. OECD library (33) (2019). https://doi.org/10.1787/09ab162c-en

71 Organisation for Economic Co-operation and Development (OECD) (2018a). *Science, Technology and Innovation Outlook 2018*, OECD Publishing, Paris, www.oecd.org/sti/oecd-science-technology-and-innovation-outlook-25186167.htm

72 Organisation for Economic Co-operation and Development (OECD) (2018b), *OECD Regulatory Policy Outlook 2018*, OECD Publishing, Paris. https://doi.org/10.1787/9789264303072-en .

73 Organisation for Economic Co-operation and Development (OECD) (2019). *Recommendation of the Council on Artificial Intelligence*, OECD Legal Instruments, OECD/LEGAL/0449, 22 May 2019, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 (accessed 20 August 2022).

74 Organisation for Economic Co-operation and Development (OECD) (2021). *Unlocking the potential of crowdsourcing for public decision-making with artificial intelligence*. Observatory of Public Sector Innovation. *https://oecd-opsi.org/innovations/unlocking-the-potential-of-crowdsourcing-for-public-decision-making-with-artificial-intelligence/.*

75 PwC (2017). *The economic impact of artificial intelligence on the UK economy. https://www.pwc.co.uk/economic-services/assets/ai-uk-report-v2.pdf*

76 Parasuraman, R., Sheridan, T.B. & Wickens, C.D. (2000, May). A model for types and levels of human interaction with automation. In: IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 30(3), 286-297.

77 Partnership for Public Service/IBM Center for the Business of Government (2019*). More than Meets AI: Assessing the Impact of Artificial Intelligence on the Work of Government*, Washington, DC, www.businessofgovernment.org/sites/default/files/More%20Than%20Meets%20AI.pdf .

78 Pang, X., Wang, C., Huang, G.: *A short-term traffic flow forecasting method based on a three-layer k-nearest neighbor non-parametric regression algorithm*. Journal of Transportation Technologies 06, 200–206 (01 2016). https://doi.org/10.4236/jtts.2016.64020

79 Ran, X., Shan, Z., Fang, Y., Lin, C.: *An lstm-based method with attention mechanism for travel time prediction*. Sensors 19(4) (2019)

80 Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 779–788. IEEE Computer Society, Los Alamitos, CA, USA (jun 2016). https://doi.org/10.1109/CVPR.2016.91

81 Ribeiro, M., Singh, S., Guestrin, C.: "why should I trust you?": Explaining the predictions of any classifier. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations. pp. 97–101. Association for Computational Linguistics, San Diego, California (Jun 2016). https://doi.org/10.18653/v1/N16-3020

82 Rieke, N., Hancox, J., Li, W., Milletar`ı, F., Roth, H.R., Albarqouni, S., Bakas, S., Galtier, M.N., Landman, B.A., Maier-Hein, K., Ourselin, S., Sheller, M., Summers, R.M., Trask, A., Xu, D., Baust, M., Cardoso, M.J.: *The future of digital health with federated learning*. npj Digital Medicine 3(1), 119 (Sep 2020)

83 Rizzo, S.G., Vantini, G., Chawla, S.: Reinforcement learning with explainability for traffic signal control. In: 2019 IEEE Intelligent Transportation Systems Conference (ITSC). pp. 3567–3572 (2019). https://doi.org/10.1109/ITSC.2019.8917519

84 Savitz, S., Davenport, A., & Ziegler, M. (2020). The Marine Transportation System, Autonomous Technology, and Implications for the U.S. Coast Guard. RAND Corporation. *https://doi.org/10.7249/PE359.*

85 Stirling, R., Miller, H., & Martinho-Truswell, E. (2017). Government AI Readiness Index. *https://www.oxfordinsights.com/government-ai-readiness-index.*

86 Shah, S.I.H., Peristeras, V., Magnisalis, I.: Government big data ecosystem: Definitions, types of data, actors, and roles and the impact in public administrations. ACM Journal of Data and Information Quality 13, 1 – 25 (2021)

87 Sun, S., Zhang, C., Yu, G.: *A bayesian network approach to traffic flow forecasting*. IEEE Transactions on Intelligent Transportation Systems 7(1), 124–132 (2006). https://doi.org/10.1109/TITS.2006.869623

88 Tambouris, E., Tarabanis, K.: Towards inclusive integrated public service (ips) co-creation and provision. In: DG.O2021: The 22nd Annual International Conference on Digital Government Research. p. 458–462. DG.O'21, Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3463677.3463726

89 Tangi, L., Van Noordt, C., Combetto, M., Gattwinkel, D. and Pignatelli, F. (2022) *AI Watch. European landscape on the use of Artificial Intelligence by the Public Sector*, EUR 31088 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-53058-9, doi:10.2760/39336, JRC129301.

90 Teh, H.Y., Kempa-Liehr, A.W., Wang, K.I.K.: Sensor data quality: a systematic review. Journal of Big Data 7(1), 11 (Feb 2020). https://doi.org/10.1186/s40537-020-0285-1

91 Tomas, J., Rasteiro, D., Bernardino, J.: *Data anonymization: An experimental evaluation using open-source tools*. Future Internet 14(6) (2022). https://doi.org/10.3390/fi14060167

92 Tinholt, D., Carrara, W., Linden, N.: Unleashing the potential of artificial intelligence in the public sector. Capgemini Consulting (2017)

93 Ubaldi, B. et al. (2019), "State of the art in the use of emerging technologies in the public sector", OECD Working Papers on Public Governance, No. 31, OECD Publishing, Paris, https://doi.org/10.1787/932780bc-en

94 UK Office for Artificial Intelligence, Government Digital Service (2020). *A guide to using artificial intelligence in the public sector. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/964787/A_guide_to_using_AI_in_the_public_sector__Mobile_version_.pdf*

95 U.S. Defense Innovation Unit (DIU) (2020). *Responsible AI Guidelines in Practice. Lessons learned from the DIU portfolio. https://assets.ctfassets.net/3nanhbfkr0pc/acoo1Fj5uungnGNPJ3QWy/3a1dafd64f22efcf8f27380aafae9789/2021_RAI_Report-v3.pdf*

96 U.S. Department of Defense (2020). *Ethical Principles for Artificial Intelligence*. *https://www.defense.gov/Newsroom/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/*

97 Wei, Z., Das, S., Zhang, Y.: Short duration crash prediction for rural two-lane roadways: Applying explainable artificial intelligence. Transportation Research Record 2676(12), 535–549 (2022). https://doi.org/10.1177/03611981221096113

98 Wilson, H. J., Daugherty, P., & Bianzino, N. (2017). The jobs that artificial intelligence will create. *MIT Sloan Management Review*, *58*(4), 14.

99 World Bank (2020). *Artificial Intelligence in the Public Sector: Maximizing Opportunities, Managing Risks.* Equitable Growth, Finance and Institutions Insight; World Bank, Washington, DC. © World Bank. https://openknowledge.worldbank.org/handle/10986/35317

100 World Economic Forum (2018*). Harnessing Artificial Intelligence for the Earth.* https://www3.weforum.org/docs/Harnessing_Artificial_Intelligence_for_the_Earth_report_2018.pdf

101 World Economic Forum (2019*). Empowering AI Leadership An Oversight Toolkit for Boards of Directors*. https://wef-ai.s3.amazonaws.com/WEF_Empowering-AI-Leadership_Oversight-Toolkit.pdf

102 World Economic Forum (2022). *Empowering AI Leadership: AI C-Suite Toolkit.* https://www.weforum.org/reports/empowering-ai-leadership-ai-c-suite-toolkit/

103 Yin, R.K.: Case study research: Design and methods, vol. 5. sage (2009)

104 Yoon, J., Drumright, L.N., van der Schaar, M.: *Anonymization through data synthesis using generative adversarial networks (ADS-GAN)*. IEEE J Biomed Health Inform 24(8), 2378–2388 (Mar 2020).

105 Yoon, J., Jordon, J., Schaar, M.: Gain: *Missing data imputation using generative adversarial nets*. In: International conference on machine learning. pp. 5689–5698. PMLR (2018).

106 Young, M.M., Himmelreich, J., Honcharov, D., Soundarajan, S.: *Using artificial intelligence to identify administrative errors in unemployment insurance*. Government Information Quarterly 39(4), 101758 (2022)

107 Zeginis, D., Kalampokis, E., Palma, R., Atkinson, R., Tarabanis, K.: A semantic meta-model for data integration and exploitation in precision agriculture and livestock farming. Semantic Web – Interoperability, Usability, Applicability (In press)

108 Zhang, M., Alvarez, R.M., Levin, I.: Election forensics: Using machine learning and synthetic data for possible election anomaly detection. PLOS ONE 14(10), 1–14 (10 2019)

109 Zhou, Q., Chen, N., Lin, S.: *Fastnn: A deep learning approach for traffic flow prediction considering spatiotemporal features*. Sensors 22(18), 6921 (2022)

110 Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A.: *Learning deep features for discriminative localization.* In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2921–2929 (2016)