NEXT
GENERATION
INTERNET

Open Call 4
ATLANTIC.EU
www.ngiatlantic.eu

Title of the Experiment:

BONSAI: cross-Border experiments for OpeN data testbedS interconnection for Atlantic Interoperability

Deliverable 3: Experiment Results (Exploitation) and Final Report

Authors	**Martin Serrano** National University of Ireland – Data Science Institute Insight SFI research Centre for Data Analytics **David Wollman** National Institute of Standards and Technology Deputy Division Chief, Smart Connected Systems Division, Communications Technology Laboratory (CTL) **Michael Dunaway** National Institute of Standards and Technology Associate Director for Innovation IoT Devices and Infrastructure Group
Due Date	30-11-2022
Submission Date	31-01-2023 (Due to Admin Extension and internal reviews)
Keywords	Data Interoperability, Experiments, Open Data, Testbeds, Data Exchange, interconnection, Interoperability

Deliverable 3: Part I

Analysis, results, and wider impact

1 Abstract

The EU-USA BONSAI project has demonstrated through the execution of cloud-based big data experiments that the interconnection of multiple open data sets is possible, the experiments focuses on testing functionalities over large, collected data sets, for data modelling, data sharing and data processing as part of data cross-Atlantic platforms interconnection. The data sets targets the interconnection of open data, including those in both sides of the Atlantic, for optimal access validating the capability to share information, that will serve as a reference providing experimental evidence on platforms interconnection and data interoperability. The NGI EU-USA BONSAI experiments promotes the use of holistic methods to classify, identify and access data and utilizes its designed service data access applications towards facilitating information frameworks and data testbeds can support data portability and interoperability.

2 Project Vision

The EU-USA BONSAI project addressed the challenge for enabling experimentation over testing data frameworks and Interconnect data platforms on both sides of the Atlantic facilitating data sharing and interoperability, for this purpose the use of open data sets is crucial, not only by the accessibility for the data sets but also for the use of open data standards. The use of open data is becoming more popular and particularly in applications where information needs to be exchange and share. The EU-USA BONSAI project aim to practically execute data management services and operations and thus experiment stream processing methods on collected open data and at the same time use the Holistic KPI methodology making data sharing and interoperability more scalable across data testbeds and frameworks as depicted in Figure 1. We focus in a particular domain where open data is very popular (i.e. smart city data systems) and take the opportunity that this paradigm exist on both sides of the Atlantic.

2.1 The Data Mashup Builder

The EU-USA BONSAI proposes that the use of data mashups is relevant in the objective of interconnecting data and thus together with the NIST CPS framework, the BONSAI project making use of both framework principles and using mashups facilitate the necessary validated platform interconnection of data frameworks for real data and information system(s).

Co-funded by the Horizon 2020 Framework Programme of the European Union

The EU-USA BONSAI approach provides an ultimate multi-source data environment for running the proposed data mashup focusing on running experiments over real data and testing that streams are generated from deployed data systems (i.e. smart city data systems) that can be shared, accessed and make them interoperable in both sides of the Atlantic.

Figure 1. EU-USA BONSAI Project Vision

The NGI Atlantic initiative, together with the NIST U.S. Department of Commerce program and NUI Galway main scope is to demonstrate and shape based on the experiments that the Next Generation Internet, is composed by an innovative frameworks ecosystem, utilising the full capacity of information systems to support large scale data, enable data sharing capabilities and interoperability to strengthen the capacity of the Internet for understanding and defining new ways for data management. The EU-USA BONSAI project focuses on the use of Key Performance Indicators (KPIs) methodology a spart of the published EU-NIST H-KPI Framework.

The Holistic KPI method is a formal method for enabling data sharing and data interoperability, across data and information systems designed and implemented on the NGI Explorers programme and published by NIST-US Department of Commerce under the category of Technical Report. The data mashups builder (SSC) is a data processing system build as a data management system that facilitate the processing of data as continuous information flow, using linked data principles and holistic methods providing a way for demonstration that data frameworks can be used in a sustainable manner for the Next Generation Internet of services/applications.

2.2 The Holistic KPIs Lifecycle

The Holistic KPIs measuring method is a 5 steps process, the overall considerations and the design of this methodology is motivated by the current misalignments between cities when KPIs are being used. By having a holistic approach identified KPIs are yet useful and valid but in a wider application it is compensated according to specific conditions.

Co-funded by the Horizon 2020 Framework Programme of the European Union

This research approach is based on data being generated by the city and by using current KPIs values thus the method can be optimal for several cities that has a corresponding set of identified data sources and KPIs.

To reduce the misalignments generated using common KPIs, relationships are introduced as the way to establishing the connections and/or dependencies (physical or virtual) within data sources and digital/physical technologies in the H-KPIs methodology. We are defining two types of relationships that will be explained in detail in the next section of this document, by now there is no need to differentiate them, but it is worth to say that if there is an increasing number of relationships in general, it represents a higher capacity for a smart city to have larger community impact. The relationships establish the capacity to the H-KPI model to account capabilities in a statistical formal manner and thus ensure the smart city capacity to address community needs more adequately. The Figure 2 summarizes the Measuring Method for Holistic Key Performance Indicators.

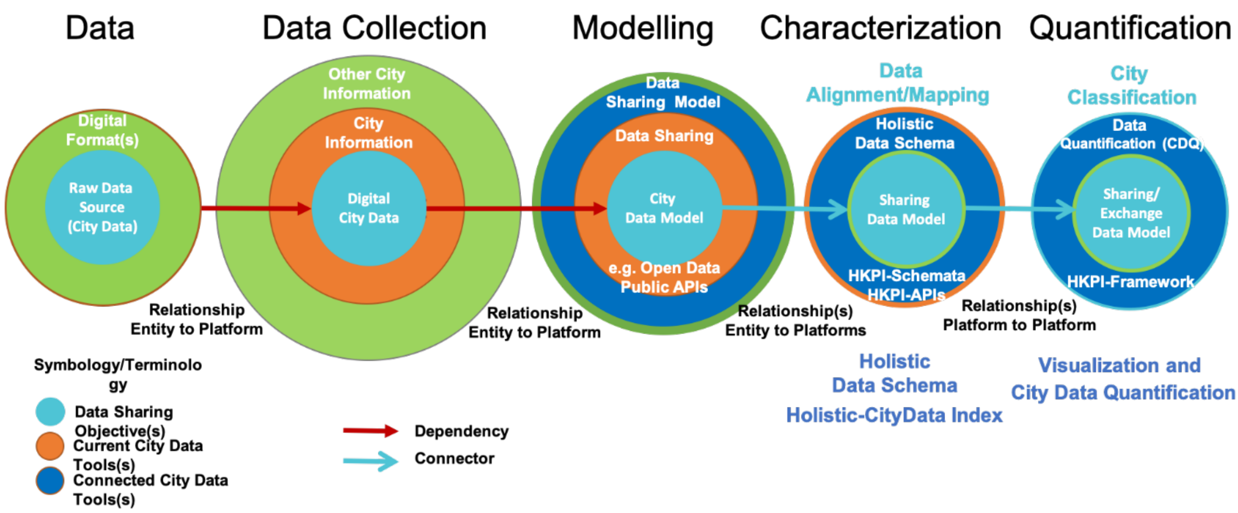

Figure 2 – Measuring Model for Holistic Key Performance Indicators Method

2.2.1 Step One Initial Data Source Selection: Find City Data Services

City KPIs are dependent of each city plans and goals, in order to identify data sources, look closely to the different available city services, look not only to the technological ones but also the platform and systems and the operation services and most important look on those that are involved directly with community impact. Once this identification is clear create your shortlisted city data services and classify them in the three layers as described in the H-KPI measuring methodology. Most likely this step is an easy one as data sources in cities are well identified because they provide the metrics for KPIs accounting.

Co-funded by the Horizon 2020 Framework Programme of the European Union

Figure 3 represent the city data as a concentric Raw Data Source. Raw data source (city data) will be associated to a digital format in order to be processed by a platform or a system e.g. number of particles from a sensors in an air quality monitoring platform for a city environmental monitoring system.

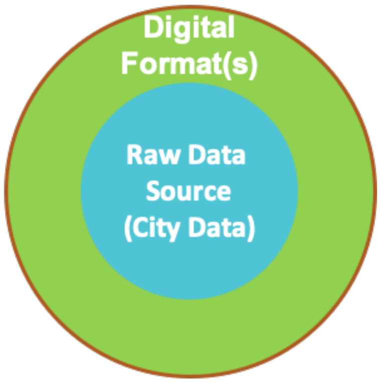

Figure 3 – Data Source Selection: Finding City Data Services to produce "Digital City Data"

2.2.2 Step Two Data Collection: Verify Available Data based on provided City Services

Once this identification in step 1 is clear, you can create your list of city data services and classify them in the three layers as described in the H-KPI measuring methodology. Most likely this step is an easy one as data sources in cities are well identified because they provide the metrics for KPIs accounting. Figure 4 represent the data with a digital format (.xls, .csv .json for example) from a city, collectively defined as "Digital City Data." City information is necessary and is represented as surrounding of the information that will classify as unique data type for a particular city. Because this Digital City Data may be also useful to other cities there are additional annotations that can be included, in the picture this is represented as another surrounding circle labelled as Other City Information.

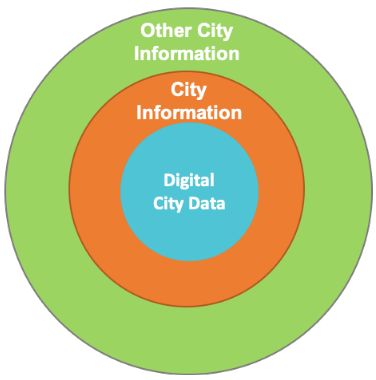

Figure 4 – Data Collection: Digital City Data Classification(s) to produce "City Data Model"

Co-funded by the Horizon 2020
Framework Programme of the European Union

2.2.3 Step Three Modelling: Selection of Data Sharing Model for City Data

At the city level, different digital formats associated to the city data sources could be found. On this step it is necessary to identify those digital formats and make use of the ones are created or prepared according to the sharing model. i.e. Open Data formats. If it is necessary, the city data needs to be transformed into one of those sharing data models in order to be usable and then H-KPI modelling can be functional. Figure 5 represent how Sharing Data Model is produced from City Data Model in order to make city Data available for the H-KPIs measuring method.

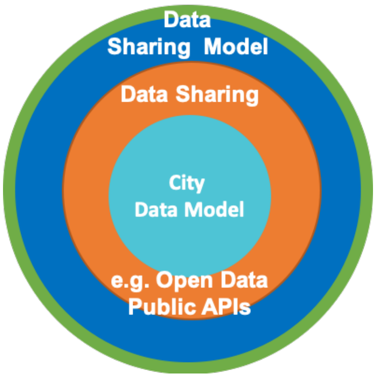

Figure 5 – Data Modelling: Data Sharing for City Data to produce "Sharing Data Model"

2.2.4 Step Four Characterization: Using the Holistic City Data Alignment/Mapping Index

City Data sources need to be identified according to the proposed layers in the H-KPIs modelling layers, but there is no restriction on how many alignments or mapping can be identified. The more alignments more accurate the measuring process will be carried out. Figure 6 represent how the Holistic Method is implemented using a data schema which allow the characterization of the City Data by means of the identified KPIs.

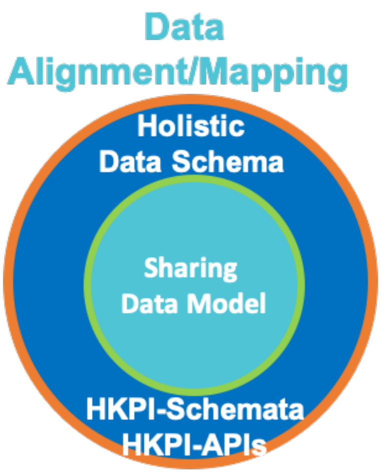

Figure 6 – Data Alignments: Mapping Data across Different Layers to produce "Characterization Matrix"

Co-funded by the Horizon 2020 Framework Programme of the European Union

The result of the step four in the H-KPIs process is a characterization matrix where all the alignments are identified and reflected for further processing and then measuring. This matrix is composed of original KPIs and weighting factors and will be used to calculate an index called Holistic City Data (HCD) Index in Step 5 based on an arithmetical function operating all the obtained values associated to the city data.

2.2.5 Step Five H-KPIs Quantification: Other considerations

The H-KPIs quantification process is a statistical analysis using the number of alignments and the city data index to compensate any city data misalignments by using a weighting model that every city would define according to their particular city plans and priorities. Figure 7 represent how the Holistic KPIs model allows the data quantification using the H-KPIs Index. The H-KPIs Index is a figure obtained and is a numerical representation of a city by following the H-KPIs alignments to simplify and reduce the misaligns between KPIs in a city or between cities as it will be explained later.

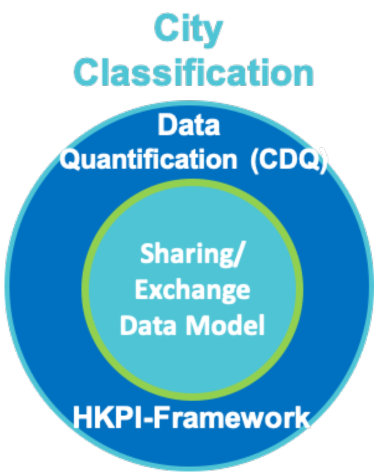

Figure 7 – H-KPIs Quantification: KPIs Measuring to produce "HCD Index"

Co-funded by the Horizon 2020 Framework Programme of the European Union

3 Details on participants (both EU and US)

1.1 EU:NUIG - National University of Ireland Galway, University Road, Lower Dangan - IDA Business Park, Insight Building

The EU-USA BONSAI project is coordinated by NUI Galway through the Insight-SFI Research Centre for Data Analytics, a lead institute on Data Analytics and Intelligent systems (previously Digital Enterprise Research Institute), recognised as one of the lead research institute on Semantic Web in the world that exhibits research excellence.

The project coordination, design and reporting activities are led by Dr. Martin Serrano, in the role of principal investigator and research unit head at the Internet of Things, Stream processing and Intelligent Systems research unit (UIoT) who is assisted by Rishabh Chandaliya, technical leader for the development and the execution of Super Stream Collider experiments. The team involved in the implementation and deployment tasks is also acting as support team for the NIST CPS Framework in matter of deployment and running tests.

Dr. Martin Serrano is a Senior Research Fellow in the Insight Centre for Data Analytics at the University of Galway. He is Guest Lecturer at SCU in Silicon Valley, Santa Clara Ca. USA. He also has Lectures at California State University, SLO, Ca, USA and TUB, Berlin and FBK-CREATE Net in Trento Italy. He holds a M.Sc. and a Ph.D. from the Technical University of Catalonia (UPC Tech), Spain. He was a Research member at the PI Team at the Telecommunications Software and Systems Group (TSSG) of the Waterford Institute of Technology in Waterford, Ireland (WIT) and Research Assistant at the Technical University of Catalonia (UPC). Dr. Serrano worked at industry as Senior Engineer Supervisor at Panasonic-AKME at the Product Design Engineering department and was Research Intern at National/Panasonic-KME (Japan).

Dr. Serrano is a recognised expert on Semantic Interoperability for Distributed Systems by his scientific contribution(s) for using Liked Data and Semantic Web Formalisms like Ontology Web Language for the Internet of Things and thus store the collected sensor's data in the Cloud. He has also contributed to define the Data Interplay in Edge Computing using the Linked Data paradigm, in those works he has got awards for his Scientific Contributions/Publications. Dr. Serrano has advanced the state of the art on Pervasive Computing using Semantic Data Modelling and Context Awareness methods to extend the "Autonomics" paradigm for networking systems. He also has contributed to enrich the Information and Knowledge Engineering Area using Semantic Annotation and Ontologies for describing data and services relations in Autonomic Computing.

Rishabh Chandaliya holds a master in Artificial intelligence form university of Cork in Ireland and is a PhD Student at the University of Galway based at the Insight SFI Research Centre for Data Analytics. Mr Chandaliya has extensive experience designing and developing advanced applications for the Android Platform, unit-testing code for robustness, including edge cases, usability, and general reliability. He used to work in industry at SME level with bug fixing and improving applications UX, gather and evaluate user feedback, assisting with facility setups, content, and configurations and assisting on Web and backend as required.

Co-funded by the Horizon 2020 Framework Programme of the European Union

He is currently investigating the optimal development and use of AI/ML algorithms, being an active and thoughtful member of the software engineering development team at the IoT, stream processing and intelligent systems contributing to the development of the prototypes.

Vinoop Sanil holds a master's in Computer Science – Data Analytics by the University of Galway, he has extensive experience in software development working for PAC Apply and EVRY India, he is currently a data engineer at the Internet of Things, stream processing and intelligent systems of the Insight SFI research Centre for Data Analytics

1.2 US:NIST - National Institute for Standards and Technology (NIST) – US. Department of Commerce, 100 Bureau Dr, Gaithersburg, MD 20899, United States

The National Institute of Standards and Technology (NIST) is an agency of the United States Government. NIST's mission is to promote innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve our quality of life.

NIST Smart Grid and Cyber-Physical Systems Program includes more than 20 scientists and engineers supported by advanced physical and computational infrastructure in pursuit of a broad portfolio of smart cities and Internet of Things research that is well aligned with the goals of this proposal. Although the creation of this report and the EU-USA BONSAI project proposal was submitted under the Smart Grid and Cyber-Physical Systems Program and started in March 2022, NIST U.S Department of Commerce went through an internal re-organisation process, and thus the Smart Grid and Cyber-Physical Systems Programme was re-assigned and integrated to the new Smart Connected Systems Division and the infrastructures and technologies were integrated in the Communications Technology Laboratory (CTL). This process is aligned with the budget assignation from the U.S Congress to NIST and the commitment and execution of acquired project is guaranteed with same budget.

NIST interest participating in EU-USA BONSAI resides in pursuing further research and produce high impact scientific results. Currently NIST is developing a mid-long term vision and a metrology-focus strategy that can act as an opportunity where the results from the EU-USA BONSAI project can pioneer further collaboration activities. EU-USA BONSAI is considered an excellent opportunity, in the context of the NGI Atlantic program, to consolidate and continue collaboration and reinforce the innovation path between US and EU organisations.

Dr. David Wollman serves as Deputy Division Chief, Smart Connected Systems Division, in the Communications Technology Laboratory (CTL) at the National Institute of Standards and Technology (NIST). Before joining CTL, he served as Deputy Director, Smart Grid and Cyber-Physical Systems Program Office in the NIST Engineering Laboratory. Prior to that, he managed efforts within the NIST Physical Measurement Laboratory to maintain and advance the Nation's electrical standards and metrology supporting the electric power industry, and he served in several other positions at NIST, including Scientific Advisor for the former Electronics and Electrical Engineering Laboratory, Program Analyst in NIST Director's Program.

Dr. Thomas Roth leads development of the technology behind the cyber-physical test bed at the National Institute of Standards and Technology as a member of its Smart Grid and Cyber-Physical Systems program office. His research interests are in formal methods for the composition of cyber-physical systems, and the detection of compromised cyber-physical devices through comparison of their reported behaviour against the constraints of the physical system.

Michael Dunaway joined NIST on June 7, 2021 as Associate Director of Innovation in the Smart Grid and Cyber-Physical Systems Program Office and leader of NIST's Global City Teams Challenge (GCTC). He has a long record of service, which includes a career in the U.S. Navy; Program Manager at the U.S. Department of Homeland Security's Science & Technology Directorate; Director of the University of Louisiana's National Incident Management Systems and Advanced Technologies Institute; and Executive Director of the University of Cincinnati's Digital Futures Resilience Program.

Dunaway's service also includes co-chair of the GCTC's Public Safety SuperCluster, working closely with his predecessor, Sokwoo Rhee – who Dunaway says he is honored to follow – and the leadership of the other eight SuperClusters. Over the last year, this SuperCluster leadership informally met almost weekly via teleconference, said Dunaway, addressing communities' emergent needs in the pandemic and discussing the way forward for the GCTC. There is a lot of enthusiasm for building on the GCTC program’s foundation, he said. Communities and cities seek more than technology integration; they seek data to better inform policy and decision making and ultimately improve their citizens' quality of life.

Dr. Edward Griffor is the Associate Director for Cyber-Physical Systems of the Smart Grid and Cyber-Physical Systems Program Office of National Institute of Standards and Technology. Dr. Griffor was, until July 2015 when he joined NIST, one of the three original Walter P. Chrysler Technical Fellows, one of the highest technical positions in industry and one that represents technical excellence throughout global industry, from the automotive to the aerospace, medical and computing industries. He is the Chairman of the Chrysler Technology Council and of The MIT Alliance, a professional association of scientists, engineers, and business experts trained at the Massachusetts Institute of Technology. Dr. Griffor completed his doctoral studies in Mathematics and Electrical Engineering at MIT in 1980 and was NSF/NATO Postdoctoral Fellow at the University of Oslo in Norway from 1980-83. He later held faculty positions at the University of Oslo in Norway, Uppsala University in Sweden, the Catholic University of Santiago in Chile and other foreign institutions as well as having taught at Harvard, MIT and Tufts University in the U.S.

4 Results

Experiments Description:

In EU-USA BONSAI we focus on framework interconnection and testing on new holistic methods for improving capabilities for distributed data characterization, quantification, processing and sharing multi-domain information, enabling portable data services; for this purpose we aim to deploy the mashup builder framework called Super Stream Collider (SSC) a middleware tool developed and maintained by NUIG Ireland. The NUIG SSC mashup builder utilises Linked data to exchange information in an interoperable and reusable way over the Internet for data exchange. Holistic Methods for data characterisation and measuring is a novel approach for data management systems developed in the context of EU-USA NGI-Explorers collaboration program a completed and awarded project in 2020. The EU-USA BONSAI aspires to expands capabilities for supporting further testbed interconnection and data sharing. The experiments to test data sharing and the data processing functions over large, collected data as part of the interconnection of multiple open data testbeds, including those in both sides of the Atlantic, that serves as a reference providing experimental evidence on platforms interconnection, data sharing and interoperability that can contribute to influence the NGI.

EU-USA BONSAI starts with a SSC Data Mashup Builder configuration running on live virtual cores (we currently tested SSC with 32 virtual cores of Amazon EC2), and EU-USA BONSAI is enabling us to test and run experiments over the real deployed infrastructure and proof that usability scenario scaling-up data streams connecting with NIST CPS framework expecting to have running data applications (as mean of web services) with 1000-10.000 entries from feeding data from large Open Data sets (e.g. smart city data) concurrently. The experiments that are performed focus on injecting large amounts of Open Data, then apply holistic transformation for enabling data processing optimisation, deploying distributed web services as data consuming services (i.e. access, search and discovery) and verify the reduction on data processing time and experimental evidence.

The focus on proposing to use NIST Cyber Physical System (CPS) Framework is to experiment over an analytical tool and established methodology. It includes the identification of CPS domains, facets, aspects, concerns, activities, and artifacts. NIST CPS framework take into consideration the CPS frequently systems of systems (SoS) approach, and thus the architectural constructs should be able to be applied recursively or iteratively to support this nested nature of experiments using CPS principles. The NIST Cyber Physical System (CPS) framework supports the provision of accurate data descriptions to allow for flexibility in virtual system creation and adaptions. As a response on this demand, EU-USA BONSAI is a collaborative cross-Border testing framework that works over large amount of open data for optimal access, using holistic methods to classify, identify and access data and utilizes its designed service data access applications towards facilitating that information frameworks and data testbeds can provide data portability and interoperability. An Holistic-KPI method was successfully developed in the context of the NGI Explorers collaboration which received the Best NGI Explorer Award for its impact and results.

Co-funded by the Horizon 2020
Framework Programme of the European Union

Implementation:

The EU-USA BONSAI implementation plan is driven by the opportunity of having the data mashup building framework (SSC) with enhanced capacities as optimized Data Provider System using the CPS Framework. It is planned to use large open data sets (i.e. from Smart Cities and validate the experiments). The Smart Cities and their data generation is one of the most suitable environments to test and validate our experiments. The US and EU cities and regions (Atlantic Cities) in terms of IoT technology adoption are continuously deploying technology and services which fits the scope of EU-USA BONSAI in terms of using real open produced data. Cities operate already IoT infrastructures and services that generate large amounts of data, which are however not exploited or use efficiently supporting strategic development goals.

EU-USA BONSAI experiments look at hoe to enable data frameworks for interconnection and interoperability experiments to expand and upscale their deployments towards facilitating and optimising data quantification experiments to demonstrate data sharing and Interoperability for the next generation Internet can be done in an effective way.

EU-USA BONSAI experiments fosters the data sharing, data interoperability and the deployment of Cloud Data, from smart cities for example and by using verified Open Data Sources for optimal frameworks interconnection our experiments can be done. Accessing and classification and its application over large heterogeneous data sources (Objective 1), with particular emphasis on the deployment of scalable, innovative, interoperable CPS Framework and SSC Middleware solution that leverage multi- source data and services from existing IoT Cloud Data infrastructures and applications in Smart Cities (Objective 2), Among the main objectives of the EU-USA BONSAI is able to design and deploy validation experiments over the experimental platforms (Objective 3) and evaluate objectively and holistically the data mashup builder performance. Figure 2 shows the proposed tasks and overall workplan.

The EU-USA BONSAI project builds a testbed-like deployment architecture, it usually takes a huge amount of effort to deploy a robust large scale deployment to show full scalable functionalities. Therefore, NIST CPS framework specifications facilities that our experiments and efforts considerably reduce and thus guarantee execute our experiments as proposed (see Figure 1 representing EU-USA BONSAI experiments Vision). The highly current demand on open data architectures and specifications focus our process of designing and building our experiments as scope and objectives are clear. It is EU-USA BONSAI clear responsibility define a simple, efficient and cost-effective experiment lifecycle management for our NIST CPS framework-based experiments.

Co-funded by the Horizon 2020
Framework Programme of the European Union

Architecture:

EU-USA BONSAI proposed experiments looks closely to advance the state of the art in platform interconnection, data sharing and interoperability by means of multi-stream data processing and by enabling the accessibility, over large heterogeneous data sources. The increased level of expertise in both participant partners will broaden the scope of the NIST CPS framework and the NUI Galway SSC Data Mashup Builder framework, as well as speed up developments in the existing Holistic framework in relation to data sharing and interoperability, in which NUI Galway will work to validate the deployed testbeds processing features and functionality. The SSC Data Mashup framework developed by NUI Galway will expand its application in other domains as Stream Processing platform, we aspire that EU-USSA BONSAI can act as an accelerator and have a direct impact on the following particular aspects:

The Figure 8 focused on representing the general design principles useful to all cyber-physical systems testbed developers and the co-simulation community; to develop specific BONSAI design concepts to guide the development, operation, and evolution of the collaboration between NUIG and SSC NIST's IoT/CPS testbed; and to establish a cross-sector IoT/CPS testbed that enables remote federation with other NIST labs and external testbeds to support NIST CPS measurement science work. Such a testbed must be integrative, reconfigurable, reproducible, scalable and usable across multi-domain CPS and IoT. It must support hardware in the loop, hardware emulation, and pure simulation in any and all permutations (the term 'co-simulation' is used throughout to refer to this combination of supported capabilities).

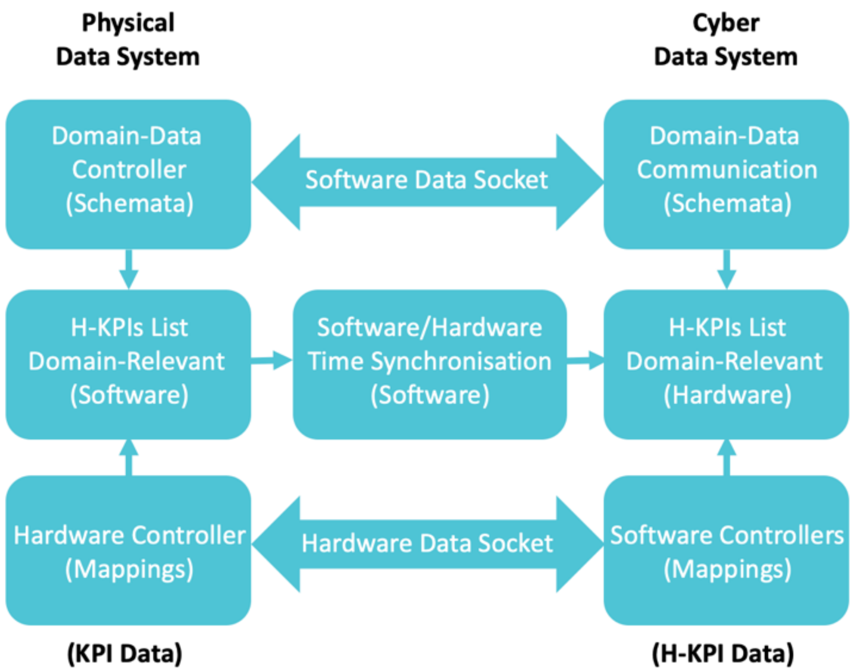

Figure 8. CPS Framework High-Level Architecture design within BONSAI

Co-funded by the Horizon 2020
Framework Programme of the European Union

The technical approach relies on three key ideas:
(1) to integrate “best-of-breed” tools from multiple domains;
(2) to do so using well-established standards for federated communications; and
(3) to define the components of an experiment in a granular form that allows experiments to be composed of well-defined and tested parts.

First, the testbed design is based on the NIST CPS Framework methodology and reference architecture concepts that promote convergence and synergy across all IoT/CPS domains. Second, the testbed’s integrative, reconfigurable, reproducible, scalable and usable design requirements makes it a powerful platform for testing concepts for composability and standards, protocols, and test methods for interoperability. Third, the testbed’s interchangeable modules allow agile reconfiguration of the testbed for varying experimental and domain-specific applications. Access to a range of physical testbeds is a unique strength at NIST, and NIST’s smart grid, SCADA, robotics, net-zero energy, building control systems, and other test facilities provide a range of candidates for remote federated experiments that allow testing of CPS/IoT concepts in a spectrum of realistic, domain-specific settings. Finally, the testbed is a resource to other NIST researchers that is both flexible and easy to use, providing documentation and usable software for many experiments that require co-simulation.

4.1 Discussion and Analysis on Results

Analysis

EU-USA BONSAI proposed experiments looks closely to advance the state of the art in platform interconnection, data sharing and interoperability by means of multi-stream data processing and by enabling the accessibility, over large heterogeneous data sources. The increased level of expertise in both participant partners will broaden the scope of the NIST CPS framework and the NUI Galway SSC Data Mashup Builder framework, as well as speed up developments in the existing Holistic framework in relation to data sharing and interoperability, in which NUI Galway will work to validate the deployed testbeds processing features and functionality. The SSC Data Mashup framework developed by NUI Galway will expand its application in other domains as Stream Processing platform, we aspire that EU-USSA BONSAI can act as an accelerator and have a direct impact on the following particular aspects:

The EU-USA BONSAI project will be used to demonstrate the envisaged value for data sharing and interoperability, directly from the experiments related to the use of Large Data i.e. from Smart Cities; currently there is a lack of a similar infrastructure not only at EU but at the international level, thus EU-USA BONSAI offers a unique environment for us to test SSC scalability and usability for Open Data from smart cities systems. The EU-USA BONSAI project will produce one of a kind specification and experimental interconnected framework using CPS and holistic methods for data sharing and interoperability design principles and most importantly it will demonstrate its application over large heterogeneous Open data sources.

Co-funded by the Horizon 2020 Framework Programme of the European Union

4.1.1. Dataset Type I – Experiments Analysis

Vehicle Traffic Dataset: There is traffic data in this dataset. The city's government has placed 449 sensor pairs along the city's main thoroughfares. Traffic data is gathered by keeping track of the number of vehicles traveling between two sites over a set period. Observations are produced every five minutes. The deployment of each traffic sensor, the distance between each pair of sensors, and the type of road are all described in a meta-data collection that is also offered. Between the two stations set over a section of road, each pair of traffic sensors reports the average vehicle speed, vehicle count, and projected travel time.

4.1.2. Dataset Type II – Experiments Analysis

Cultural Event Dataset This quasi-static dataset consists of Aarhus municipality-sponsored cultural events. The dataset is updated from time to time to include the most recent information about upcoming cultural events. The dataset is updated from time to time to include the latest information about upcoming cultural events. Updates are available as a data stream. If there are any changes to the dataset, a notification service notify this to let you know. Due to how rarely it is updated, we think of this dataset as background information and use it to show how static data and data streams can be combined.

4.1.3. Dataset Type III – Experiments Analysis

There are several datasets that contain cough audio samples related to COVID-19. We have selected here COVID-19 Data Sets for runing the experiments in the context of BONSAI:
The COVID-19 Sounds (Cambrige Dataset) is a collection of over 53,449 cough audio samples from individuals with confirmed or suspected COVID-19. The dataset includes audio recordings from a variety of devices, including smartphones and laptops, and includes both male and female subjects of various ages. The dataset was created by researchers at the University of Cambridge and University of Southampton, United Kingdom and is intended for use in the development of algorithms for the detection of COVID-19 using cough,breath audio.

The COUGHVID dataset is a cough and speech audio samples from individuals with confirmed or suspected COVID-19. The dataset includes over 25,000 audio of cough recordings. The cough audio samples were recorded using a variety of devices, including smartphones and laptops, and include both male and female subjects. The speech audio samples were recorded using a professional microphone and include male and female subjects reading a standardized script. The dataset was created in intent for use in the development of algorithms for the detection of COVID-19 using audio.

Co-funded by the Horizon 2020
Framework Programme of the European Union

The Coswara dataset is a collection of over 2,030 cough audio samples from individuals with confirmed or suspected COVID-19. The data in the dataset includes audio recordings from individuals of various ages, genders, and geographical locations. The recordings are collected in a variety of languages and dialects, and are annotated with metadata such as the age, gender, and location of the individual who provided the recording. The Coswara dataset is a collection of audio recordings that have been contributed to the Coswara project by individuals around the world. These recordings include a variety of sounds, such as breathing, cough, and speech, and are intended to be used for the development of machine learning models for COVID-19 detection.

It's important to note that these datasets may have different levels of annotation, including information about the individual's COVID-19 status, age, gender, and other characteristics. They may also have different terms of use and access restrictions, so it's important to carefully review the terms of use for each dataset before using it.

4.1.4 Data Model

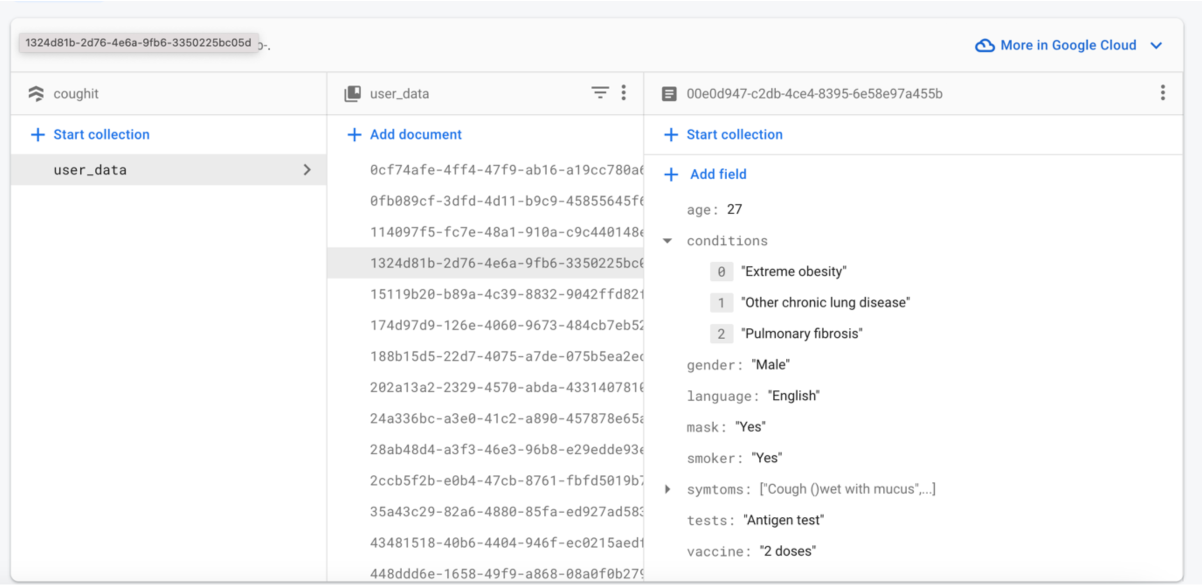

Figure 9. BONSAI Data Model

The data was saved on Firebase, and a stream was made that can be accessed at any time using the API key that Firebase gave. The cough and the breath raw audio files are stored in Firebase Storage, and below are the raw audio file conversions into respected features and are stored in the following below format.

Co-funded by the Horizon 2020 Framework Programme of the European Union

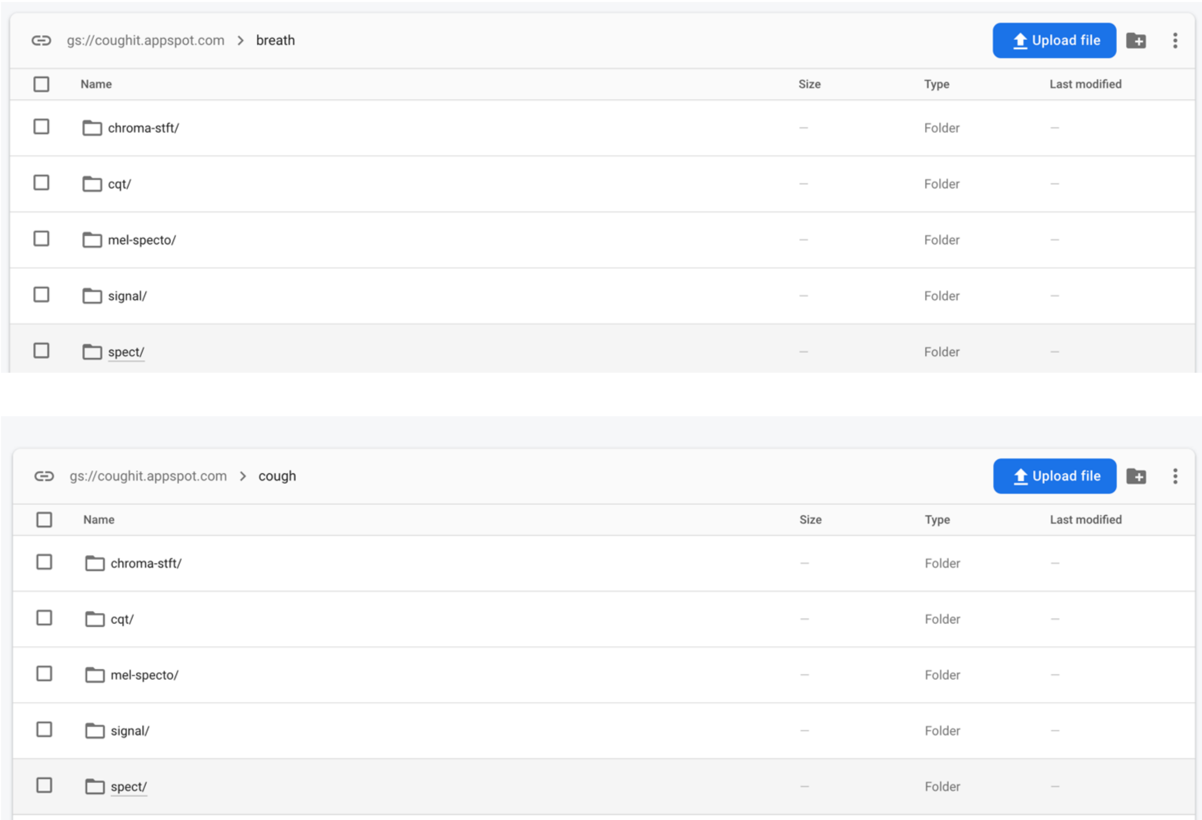

Figure 10. BONSAI Data Base Structure

4.1.5. Data Transformation:

Data is aggregated in order to improve the organization. Transformed information is easier to use on computers and by humans. In addition to protecting programmes from potential dangers like null values, unexpected duplicates, improper indexing, and incompatible formats, correctly formatted and verified data improves the quality of the data.

We ordered them chronologically to make the dataset simpler for both people and machines to understand. Inside each date document, the fields are arranged by time. The time is also translated into the report id, average measured time (average time taken to travel between two points), median measured time (median time taken to travel between two points), average speed (distance between two points over the average time taken to travel between them), vehicle count, and other data.

Co-funded by the Horizon 2020
Framework Programme of the European Union

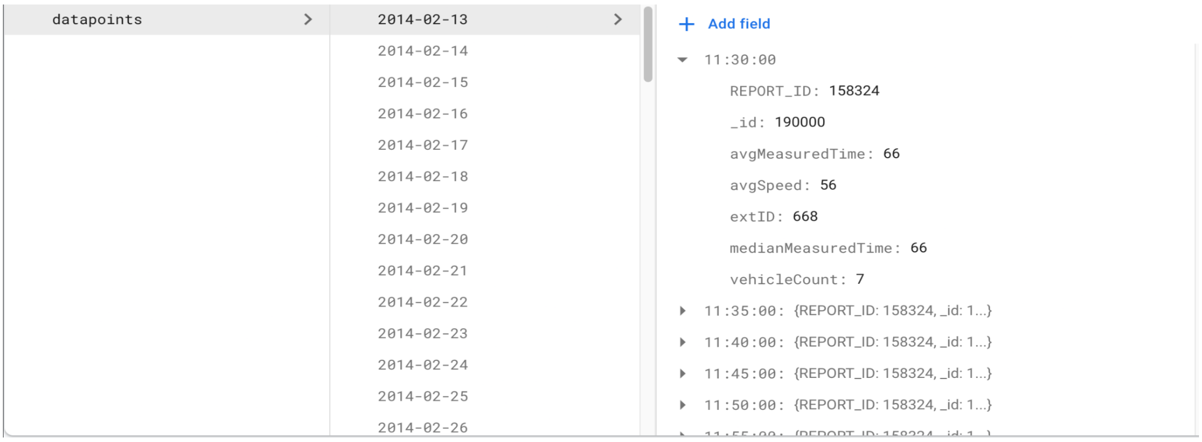

Figure 11. BONSAI Data Transformation Example

4.1.6. Script Code:

1. Read the CSV file
```
df = pd.read_csv('traffic_feb_june/trafficData158324.csv')
```
2. Split the TIMESTAMP between Date and Time(Minute)
```
df[['Date','Minute']] = df.TIMESTAMP.str.split("T",expand=True,)
```
3. Drop the TIMESTAMP column, as it is no longer needed
```
df = df.drop(['TIMESTAMP'], axis=1)
```
4. Select the required data
```
data = df.iloc[0,1:8]
data.to_dict()
```
5. Transform and add to the collection
```
collection = db.collection(u'datapoints')
for i in range(0,len(df["Date"])):
res = collection.document(str(df["Date"][i]))
      .set({df["Minute"][i]:df.iloc[i,1:8].to_dict()}, merge=True)
```

Results

The BONSAI testbed (powered by SSC platform) is designed as a builder and explorer with the capacity to organise and process dataflows/workflows under a data-driven execution environment connecting multi-owner and cross-domain processing input/outputs through pipelines of information for creating data streams known and open data mashups.

The execution of the BONSAI's experiments are executed on an controlled environment (container). The execution container is a cloud-based repository where continuous query processing engine is deployed and it is used for stream processing operators. The execution containers are running on networked virtual machines with core processing capacity to run the processed dynamically allocated based on the processing load registered to SSC platform and the CPS testbed. Figure 12 shows an informal high-level view of this architecture.

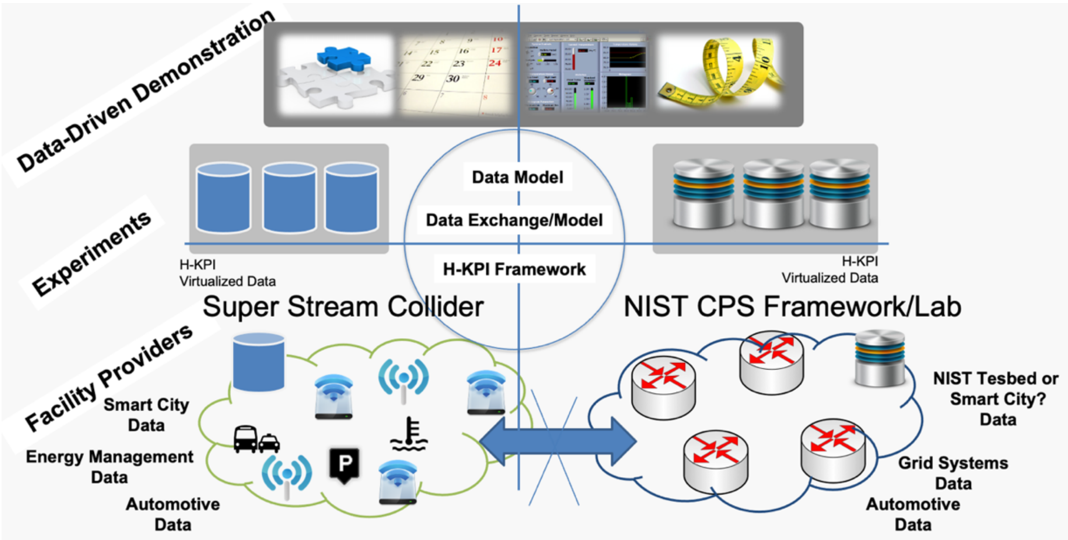

Figure 12. BONSAI Experiments Architecture within BONSAI.

The number of registered data resources can grow SSC requests additional data fetching engines from its underlying cloud computing infrastructure and allocate them to this task. he SSC can flexibly answer to dynamic load-profiles which are common in stream-based applications. In a concrete workflow, two connected operators can be executed in different execution containers. For instance, the data acquisition operator for collecting Tweets can stream data via the network to the stream processing engine. The external computing services such as SPARQL endpoints or web services are called external execution containers. To support the easy and intuitive definition of data processing workflows in a "box-and-arrows" fashion, the SSC platform offers a visual programming environment. The interactive process of creating a mashup with SSC features context-aware discovery services for data sources. This process enables the user to incrementally build a workflow in a step-by-step fashion by dragging & dropping the required building blocks and connecting and parametrizing them. Also, this supports visually debugging the workflow of the mashup. When the user finishes a mashup, it can be deployed to the SSC cloud to be re-used as a data source or an operator.

4.1.7 Testbed/Platform Builder Setup

A new deployment of the SSC with new adaptations and extension towards using the H-KPI framework data model was developed within the context of the BONSAI project, which provides a user friendly interface called SSC visual editor. This is a light-weight Web-based workflow editor for composing mashup data through drag & drop. Using the SSC visual editor, we aim at providing a programmable Web environment suitable not only for expert users but also for non-expert programmers. Figure 13 below provides an overview screenshot of SSC with the numbers feature explained in the text below.

Co-funded by the Horizon 2020 Framework Programme of the European Union

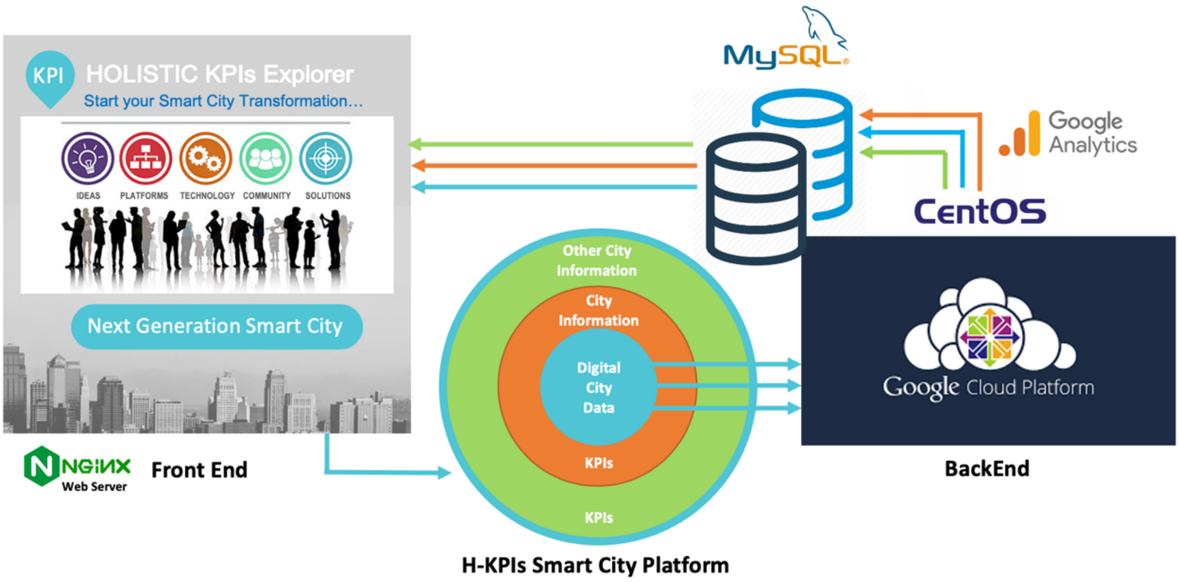

Figure 13. BONSAI Experiments BONSAI H-KPI Builder/Explorer.

Front End

- A responsive GUI based on HTML for ease of interaction (i.e. Apache Server)
- Presentation of the Holistic KPI concept and its use in Smart Cities
- Description of the Benefits of Holistic KPIs

Platform

- An implemented H-KPIs measuring algorithm using JAVA-script or Angular JS Language.
- An API to collect City Data and define Smart city KPIs weighing for a city.
- A Relational Data Base for local storage.
- A security tool for ensuring security of data and accessibility i.e. OpenAM/OpenID.

Backend (VM Server – CentOS Using Linux Distribution)

- Cloud Infrastructure
- Analytics Tool

4.2.2 Monitoring and Data Collection Explorer

The H-KPI explorer is part of the BONSAI Dashboard, which is implemented as a prototype following H-KPI measuring methods. Figure 14 shows the proof of concept where Derivative and Integrative functions are represented. On left upper side and right upper side, the sigma for quality of the data has been introduced to control the total of data points that are used for plotting the functions, while at the same time a percentage, sigma inputs and data set formats are selected to define the mapping methods and use the aggregated information form the created H-KPI data model.

Co-funded by the Horizon 2020 Framework Programme of the European Union

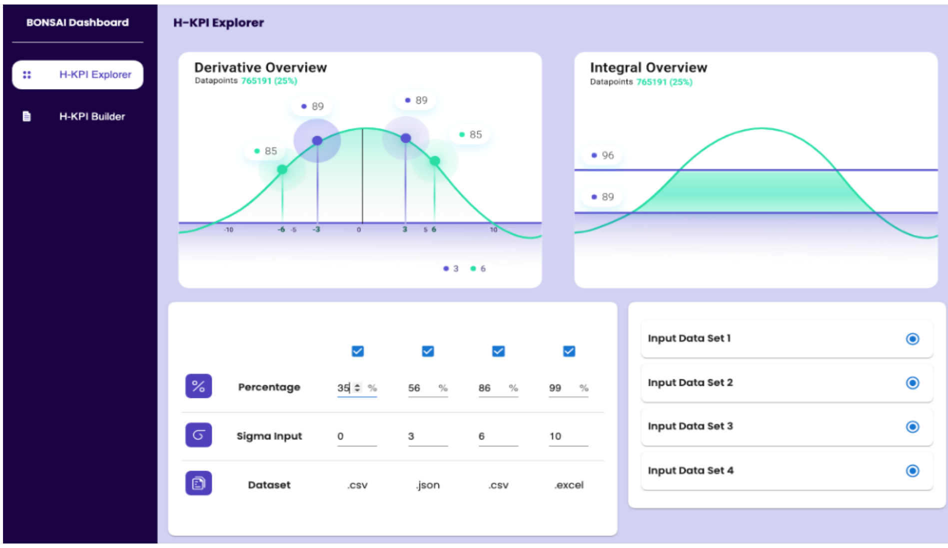

Figure 14. BONSAI H-KPI Explorer Experiments Dashboard.

4.2.3 Visualisation and Analytics Tool

The H-KPI builder is part of the BONSAI Dashboard, the current implementation relies in the capability to identify and define the maximum number of relationships as potential best-effort approach while constructing and identifying the total number of relationships. Figure 6 shows the proof of concept implementation where the different levels i.e. Level 1 to 3, as defined in the H-KPU framework specification, the number of data points in each section represent or is aligned with the defined KPIS for each of the identified domains and the type of data sets that are produced. The right part show the graph representation and the interlinks created and identified following the establishment of relationships following the random function but controlled by the total number of relationships as a percentage defined in the lower bottom part as a function of the total relationships.

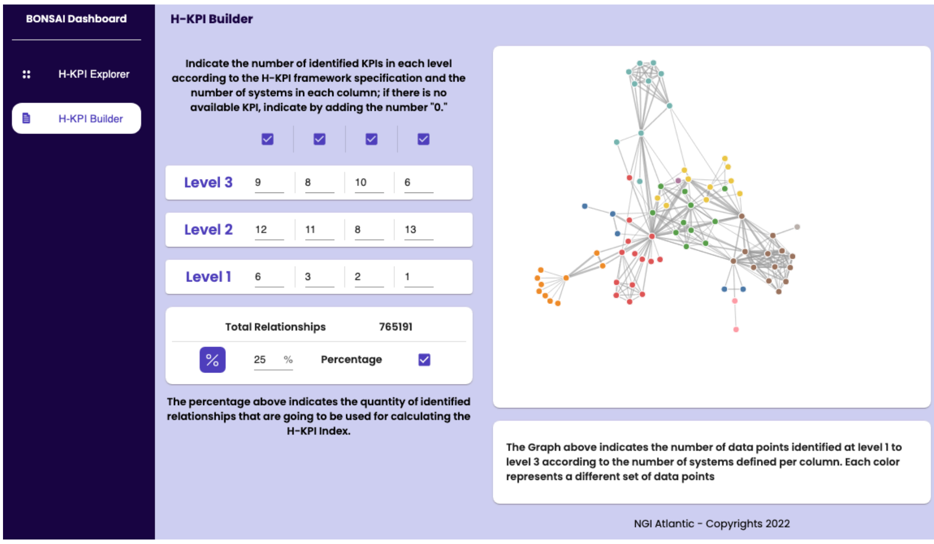

Figure 15. BONSAI H-KPI Builder Experiments Dashboard.

Co-funded by the Horizon 2020 Framework Programme of the European Union

5 Present and Foreseen TRL

The EU-USA BONSAI project will advance the SoTA providing an analysis to validate best technologies, platforms and frameworks related to the main activity proposed in addition to the NIST CPS Framework and NUIG Mashups Builder frameworks. Testing over large amount of Open Data generated from NIST CPS framework and/or other sources (e.g. smart cities data-systems) that can be processed and used by the NUIG Mashup Builder. The EU-USA BONSAI project will provide with a unified access interfaces and tools that will enable the experiments execution by using the NIST Cyber-Physical Systems (CPS) framework and the mashup builder framework called super stream collider increasing the TRL from 4 to 6 in both cases and level up the capability of the frameworks interconnection and acceptance of Data Sharing methods for the data services and information systems.

The EU-USA BONSAI project, by using the recent investigated and released Holistic method for characterising data sets, will validate the capability to offer data facilities for experimenters and application developers on the top of the interconnected frameworks and particularly for the use of data stream feeds for large-scale consuming applications (as experiments for optimal data access, portability and interoperability.

6 Exploitation, Dissemination and Communication Status

The EU-USA BONSAI is executing innovative experiments that demonstrate technological expertise, scientific novelty and quality results in a relevant area which is already identified as crucial part of the Next Generation Internet, thus a series of activities have been conducted and it is expected their results can be exploited, disseminate and communicate:

6.1 Exploitation

The EU-USA BONSAI main part for exploitation is to progress the state-of-the-art in H-KPI platform interconnection, today there is no standard way to define the way two or more data platforms can interconnect if it is not using large efforts on understanding data model and infrastructure requirements, furthermore there is no yet demonstrated framework that offers multi-stream data optimisation by means of using holistic methods.

The use of H-KPI as an holistic method using statistical analysis as baseline for data exchange and generally leads to a higher levels of data interoperability data managing and high processing capabilities expanding the capability of the NIST CPS framework.

The EU-USA BONSAI is a pioneer project in using statistical methods for characterizing the data used in the experiments and thus it is possible to demonstrate that by using a well-defined data model the CPS framework and the SSC data mashup builder components can expand data processing capacity. The use of this capacity is clearly an improvement in data identification and characterisation and leads to some level of system intelligence that can be further develop and used in intelligent-lead data-driven applications.

Co-funded by the Horizon 2020
Framework Programme of the European Union

6.2 Dissemination

The EU-USA BONSAI main dissemination activities focuses on two folds, first one to describe and demonstrate the scientific value of the statistical methods and their use in experimental testbed demonstrations and the second one to use this methods with real Open Data (ie. Open Data sets from Smart Cities), the dissemination for those two are planned as follow:

NIST at the Smart Connected Systems Division, Communications Technology Laboratory (CTL) organises public workshops (online during post-covid time) where the ongoing projects and current progression and results are showcased. The value in this dissemination activity is to present the progress in relation to the use of statistical methods for data interoperability and most important the use of open data for testbed interconnection and experimentation, today there is no standard way to define the way two or more data platforms are interconnected and with the state-of-the-art in platform interconnection, as part of the EU-BONSAI project, this is a step forward standardising this process.

A bi-lateral collaboration workshop with John Hopkins University (JHU) academic and research staff was organised by NIST in order to present the different projects and research initiatives, EU-USA BONSAI project had a slot to present the objectives, the current progress and the deployed experiments which integrated a series of COVID-19 data sets using the Stream Collider mashup builder approach and the interactive user Interface for the BONSAI experiments.

At the European side The NGI Atlantic Showcase, an event co-located at the IoT WEEK 2022 in Dublin, Ireland where the EU-USA BONSAI project progress can be presented, the objective is for enlarging the Next Generation Internet community that can benefit from innovations like the EU-BONSAI framework results using the US NIST CPS Framework and EU data mashup builder, and thus enabling knowledge transfer of EU-USA BONSAI framework validation to other stakeholders in industry, SME`s and other interested communities; i.e. academia and smart cities, establishing new cooperation links and increasing the current level of cooperation of EU and USA Partners.

The AIOTI (Alliance for the Internet of Things Association) organised an event and showcase were testbed were invited to present their experiment's progress and results, EU-USA BONSAI was invited and presented the experiments and the ongoing results showcasing the mashup and integrations of multiple smart city open data sets. This event opened the opportunity to other experiments with COVID-19 data sets, vehicle traffic data and Cultural Events Data.

The EU-USA BONSAI created a synergy with academic courses in San Jose City in Silicon valley, particularly at the north-western University campus. A class on Machine learning and Analytics tools where PhD and MSc students heard about the NGI and NGI Atlantic programme and at the same time the EU-USA BONSAI experiments were presented and the current progress and results explained demonstrating the feasibility of cross-domain data interconnection.

Co-funded by the Horizon 2020
Framework Programme of the European Union

6.3 Communication

The EU-USA BONSAI project partners has defined and implemented a communication strategy following two different levels according to the following activities:

6.3.1 Internal Communication

The EU-USA BONSAI Partner participants meet in a regular basis every two weeks as part of the EU-USA BONSAY Communication Strategy, the meeting is organised and hosted by US-NIST CTL laboratory and the agenda comprises 3 main parts:

1) Updates and management activities
   The studies, design and high level discussion are held in this part of the meeting, due to the nature of remote online meeting this part is limited to 15 minutes max, if further details or discussion need to be discussed a separate meeting/online call is scheduled.
2) Progress and technical aspects
   This part is the most relevant in order to progress the project, both sides EU and USA explain the current deployments and software tools and implementations (if any) that have been conducted in the past two weeks, there is also a short section where the plan for the next two weeks are presented and agreed, both sides work separately but there are planned 3 x Trips where there are is a week to 10 days on site work at CTL remises.
3) Other related activities and events
   In this part, announcements and news are shared with the objective to look for other opportunities for EU-USA BONSAI project, particularly participation and dissemination.

6.3.2 Public Communication

The communication about the progress of the EU-US BONSAI project is done via the Open Agile and Smart Cities Ireland web Portal (i.e. www.oasc.ie) , the project is described as an ongoing project under the activities that are relevant for the community and has impact in smart city stakeholders and events and workshops are included under News & Events.

The public communication channel is also helping towards defining opportunities to test in particular domains the data sharing and Interoperability activities as the result of the EU-USA BONSAI project. It is an opportunity to offer the opportunity for testing multi-stream multi-domain and fully distributed accessible anytime anywhere features on collected open data.

Co-funded by the Horizon 2020
Framework Programme of the European Union

7 Impacts

The EU-USA BONSAI project has executed innovative experiments that demonstrates technological expertise, scientific novelty and quality results in a relevant area which is already identified as crucial part of the Next Generation Internet. The series of activities has created high results and impact from expo demonstrators to critical specification improvements in the H-KPI framework. The impacts are described briefly here as namely:

7.1 Impact 1: Enhanced EU – US cooperation in NGI, including policy cooperation.

This impact event took place at the IoT WEEK Showcase, an In-Person Event where the European Commission's NGI Initiative, coupled with the National Science Foundation's support through a dedicated session showcasing a number of the successful EU–US projects with representatives from both the EU and US participating in the panel in an interactive format (Figure 16). NUIG as coordinator of the BONSAI project, we presented results and we took take part in an interactive Q&A session together with other projects i.e. ARES Experiment, ATLANTIC-eVISION, Integrating OpenIreland and COSMOS testbeds, Vulnerability Assessment and Robust Defenses for Optimized Attacks in Dynamic SDNs Experiment and the Secure communication based on robust 3D localization Experiment.

The BONSAI Experiments presented the Stream Platform and H-KPI analysis and the extensions to the H-KPI framework to a forum of approximately 60 attendees the panel participation and the interactions made the audience aware of the H-KPI framework extensions and the undergoing experiments alike the potential of using BONSAI approach in multiple other domains i.e. industrial manufacturing, smart cities, etc with the objective of quantify impact and measure maturity based on data produced. BONSAI extensions look at addressing the specification of large data sets for cross-domain interconnection.

Figure 16. Enhanced EU-US cooperation in NGI, including policy cooperation.

7.2 Impact 2: Reinforced collaboration and increased synergies between the Next Generation Internet and the US Internet programmes.

7.2.1 The EU-USA NGI Atlantic BONSAI project has established the ways to demonstrate that Cyber-Physical Systems and H-KPI framework using large amount of data are interoperable, by using the H-KPI specification BONSAI experiments and the use of data mashups address the objective of interconnecting data sets and thus together with the NIST CPS framework, the BONSAI framework using mashups builder principles facilitates the necessary validated platform interconnection of data frameworks for real data and information system(s).

7.2.2 The EU-USA NGI Atlantic BONSAI project has demonstrated the collaboration with the NGI Explorer team by building on top of the previous work a prototype for the BONSAI Dashboard in order to reach outstanding results and being able to demonstrate as a proof of concept that data mashups and data testbed interconnections is achievable. The reinforce collaboration between NGI programs is a great success of continuity within its purpose of bridging, expanding, and sustaining a cross-continent Research and Innovation collaboration.

7.2.3 This impact event took place at the Smart City Congress Barcelona, an event organised in collaboration with EU FIWARE and the EU-USA NGI Atlantic BONSAI project where the cross-domain data interconnectivity was the main topic to be presented and discussed. In this panel (Figure 17) experts from USA and Europe presented their views and findings in relation to how smart city data have evolved and how to increase intelligence. The expert panellist (FIWARE) will debate what is the future for smart cities and the methods to measure maturity (H-KPIs) and add more intelligence to smart (BONSAI). Conclusions and next steps were discussed.

Figure 17. Enhanced EU-US cooperation in NGI, including policy cooperation.

Co-funded by the Horizon 2020 Framework Programme of the European Union

7.3 Impact 3: Developing interoperable solutions and joint demonstrators, contributions to standards.

The EU-USA BONSAI project developed a prototype for the H-KPI explored and H-KPI Builder in the form of BONSAI Dashboard with the main objective to showcase to Governments, scientists and industry communities that it is feasible to use the H-KPI framework specification everywhere and not only in academia but also industry and also consumers on the planet's and Europe achieving its so-called twin (green and digital) transition to a zero carbon, zero waste economy that leaves no person or place behind.

At the AIOTI signature event under the theme "New green technologies are already here to help tackle the biggest challenge of our time: climate change"(Figure 18) . It was discussed that the European Commission has long promoted digital transformation to enhance economic competitiveness, while also recognising that digitisation can contribute to sustainability goals and enable the changes needed for a just green transition and that the Commission's twin green and digital goals are seen to complement each other well.

Insight SFI research Centre for Data Analytics participated presenting in a demo space for the "BONSAI Experiments Interconnecting NUIG-SSC and NIST-CPS Testbeds" focusing on the interconnection of data sets form different domain area, i.e. wellbeing via a cough-cough app data and the most common respiratory diseases, in other to process the results, the demo room attracted approximately 60 participants between policy makers and technologist (industry) and experts (Academia).

Figure 18. BONSAI Experiments Interconnecting NUIG-SSC and NIST-CPS Testbeds.

Co-funded by the Horizon 2020 Framework Programme of the European Union

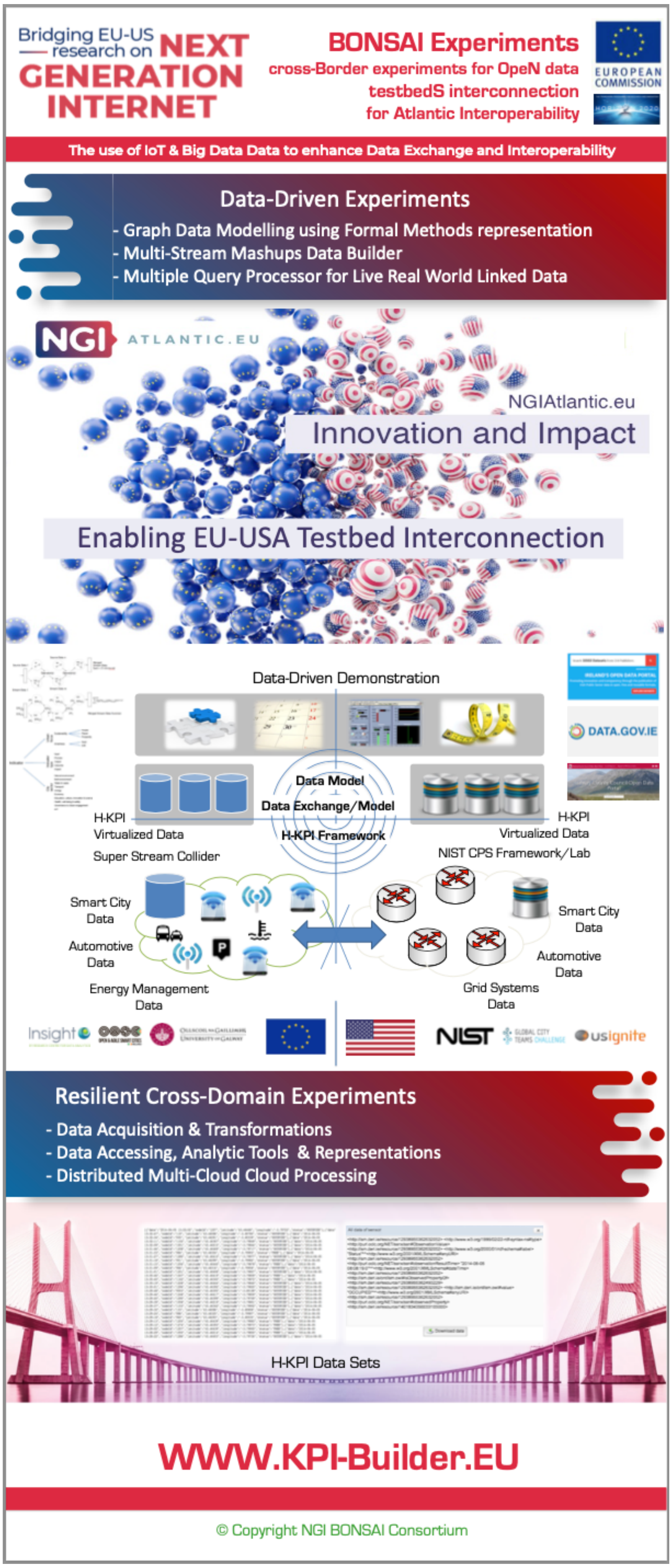

Figure 19. BONSAI Experiments – Banners Showing Data Interconnection.

Co-funded by the Horizon 2020 Framework Programme of the European Union

7.4 Impact 4: An EU-US ecosystem of top researchers, hi-tech start-ups / SMEs and Internet-related communities collaborating on the evolution of the Internet

7.4.1 The EU NGI Atlantic BONSAI addressed a community of smart city and in particular technical practitioners from SME and start-ups looking at the most trendy elements and innovation in the sector of using Smart City data for large data sets and cross-domain interoperability enabling AI solutions, The GreenCities: Urban intelligence and Sustainability is a place where the main technology industry and technology leaders meet with the purpose to define trends and align initiatives in relation to the urban management and future mobility.

The NGI BONSAI participation served as the channel to present and discuss the latest technological advances and news around H-KPI explorer tools and the advance testing performed in the context of the NGI Atlantic collaboration framework (Figure 20). A set of data-driven texted experiments with recent innovative urban data centred on people mobility was discussed. The industrial focus participation and the large engagement with smart city data ecosystem as this event on 20th-21st September attracted around 250 participants from of 2900 attendees where +300 industry +70 cities +35 countries with more than +250 experts.

Figure 20. Ecosystem of top researchers, hi-tech start-ups / SMEs collaborating

Co-funded by the Horizon 2020 Framework Programme of the European Union

7.4.2 The EU NGI Atlantic BONSAI created a direct impact in industrial education by providing a guest lecture/tutorial to approximately 25 participants about the BONSAI Experiments and the demonstration of the experiments in the topic of data modelling, mashup stream processing and cross-domain exchange. The Lecture took place at the North-eastern university, the prepared sillabubs was dedicated to instruct industry professionals sponsored for their employee companies (i.e. TESTA, Apple.) where they can learn more about latest advances in technology and the emergence of new scientific paradigms (EU - Insight SFI Research Centre Part) and standards (USA - NIST Part). The lecture also addressed other community challenges and the technical challenges for SMEs and start-ups looking at the trendiest elements in relation to building data mashups, data management systems and data processing and analytics, this activity was performed in San Jose - Silicon Valley North-Western Campus in Ca, USA. from 03rd-11th December 2022.

Figure 21. EU-BONSAI Experiments – Industrial Education in Silicon Valley University

Co-funded by the Horizon 2020 Framework Programme of the European Union

8 Conclusions and Future Work

The EU-BONSAI Project have implemented a series of experiments using open data from multiple domains i.e., Smart Cities, COVID-19, Automotive Traffic Data and Cultural Events data set as part of the testing about the implementation and deployment of the BONSAI framework. Having discussed the positive impacts of the experiments in this final report, the learnings and experiments serve as blueprint experiences that from now on will provide and serve as an overview of the current state of play on mashup building technologies.

The EU-BONSAI project includes findings from the research activities in relation to mashup building, large data sets integration, Data & Information management and data manipulation of how large amount of data from heterogeneous sources will become available in the next generation of internet services, such applications and services are not available today, characteristic that is not accessible/existing in today's Internet solutions.

The EU-USA BONSAI experiments have demonstrated that vast amounts of data can be collected from type to another type and from place to another place, taking as a reference the need to combine and merge data sets the mashup builder characteristic is very relevant in the advance of cross-domain data integration. The future of mashup services and cross domain integration is defined by a series of open specific-domain challenges as follow:

8.1 Opportunities for Data Mashup Experiments

Enhancing Data Services, many data service platforms have been around so long, propietary and open platforms, however they are considered more like non-accessible platforms mainly because the data and their life cycle are proprietary and does not correspond to actual shared problems that should be solved, but particular issues in closed environments and individual applications. Smart Cities for example, there is traffic data that changes at rush hours. Parking downtown data also represent a significant problem in cities by its diversity and sparsity in large geographical locations. Mashup technology and its application in smart city data systems represent an opportunity to address these challenges and they can certainly help. Emerging technologies can utilize the merged data keep traffic flowing, Parking spot sensors identified etc., These examples only represent the impact and connected features that the smart city data controlled and activated via data mashup technologies can do to reduce time for maintenance and provide greatly improved data services.

Co-funded by the Horizon 2020
Framework Programme of the European Union

Reduced Data Management Operating Costs, the prevention of unnecessary waste of resources used in management operations while processing data is more than just a convenience; these days with so much interest in greening the planet and save energy, having a reduced data management operating cost also saves money and reduces greenhouse gas emissions. Installing mashup builder systems reduce the amount of data systems. Multiple data formats are necessary to transform a city into a smart city, but obviously there is an initial investment in using mashup engines, but this is compensated in how the data processing capacity can be added up quickly mitigate the initial investment. Data Management operations reflect also the health of the management system, the numbers veer outside the expected range of data processing capabilities reflecting unused resources, but the mashup builder system can automatically provide real-time data to back up their decisions and compensate the use of the resources and in long term it uses will be more optimal.

Improved Data Economy and Commerce, EU-USA BONSAI project encourages cooperation between multiple data stakeholders i.e. public and private organizations to collect, analyze, and process data. Data Mashup can trigger a disruptive businesses opportunity using data collected through smart city systems and improve their offered services like understand pain points, provide mashup services and applications and better target potential customers. Other opportunities can emerge as part of the Smart Applications where data is collected automatically. In this scenario, data collection at installed downtown street corners can provide data services to help people wherever they're trying to get information from any available sources where they different formats.

Data Equitable Access. This means many residents and visitors miss out on the benefits that the data connection and interoperability can bring, by including innovation and free-flowing communication, the economic development and educational opportunity including democratic access exist. Data access and construction of mashups can be a literal lifesaver when it is necessary to provide information to people. Making a commitment to smart technology also means creating an opportunity to ensure internet access. Other alternate and innovative technologies exist to help integrate data access in data equitable manner and because it is or difficult to reach, such feature is purely technology dependent. However, the data sharing aspect and the interoperability element creates that date become a great motivator to start providing much-needed internet access to a community.

Everything Else, the possibilities with smart data technology are nearly limitless. Sensors Data is a tremendous source from tracking a car in densely populated city to know which streets has the most pollution levels and which others have been plowed after a blizzard, from sensors to bots. Data can allow the integrity of bridges and buildings to prevent full or partial collapses. the use of smart data can free up resources and make them focus on strategic activities.

Co-funded by the Horizon 2020 Framework Programme of the European Union

8.2 Potential use of H-KPI and EU-USA BONSAI Frameworks

The EU-USA BONSAI project have incorporated the use of the H-KPI framework as the methodology for modelling and for making a more efficient use of the information. There are many potential opportunities to use H-KPIs Framework, first because it is designed as a generalization for data integration and interoperability which in the context of EU-USA BONSAI project is applied to the integration of smart city data sets, however it also can be extended to other data applications where modelling of data is required and/or in other circumstances where multi-factors to achieve overall data quantification includes contextual data. The following are examples on how H-KPI framework can be used for Smart Cities Data:

8.2.1 Maturity Assessment (by means of Smart City Data Evaluation)

Smart cities have been evolving over the last decade from a perspective of increasingly starting to use emerging technology (under the umbrella of smart city innovation) for addressing people demanding more city services i.e. permanent connectivity in the city, Information on the move, digital services in all municipality processes are just some of the examples of those demands. In cities today nothing is disconnected from global economies and societies. Information travels so fast that there is no city that does not know what the latest developments are of one or another city around the globe. However, in this fast moving exchange of information and growing demand for services there is no a single method that define the best way to identify demands and map them with available services in the city. There is no a standard way that can define the full behaviour of a city and what is most relevant, there is no a single validated method to assess activities all together from the city.

Smart cities use KPIs to have them as a reference from one city to another, for example if a city has implemented a smart lighting system and this represents a good solution to the demands of that city and additionally the system has been successfully recognized, the other cities will consider implementing it in their cities using similar metrics or adapting them to the conditions of the other city. This is the effect of sharing experiences and best practices amongst cities. In the scenario, where the city conditions are not the same and possibly the demands and geographical conditions and public interest vary, there is no guarantee the smart city systems will impact equally in the other city. It is true that the cities have their own way to assess KPIs but, if similar KPIs appear in the plans of two cities and one of the two succeeds relative to those KPIs, then chances are greater than the other city will also succeed.

Co-funded by the Horizon 2020
Framework Programme of the European Union

In Maturity Assessments for smart cities there is no intention to evaluate the impact by introducing new KPIs, but rather an intent to assess the impact and measure the corresponding impact on the community. Maturity is the result of a smart city transformation process and it involves all the activities for the community and relies on the technology and platforms that have being deployed. The H-KPI Measuring model addresses this need and involves the overall activities, introducing a methodology to agnostically account for all those actions and benefits.

8.2.2 Scale Up KPIs (City Strategic Planning by means of Smart City Data)

Cities are generators of data, each interconnected system is a massive source of information to understand cities' behavior, citizens and their activities – it is part of a city fingerprint in terms of the data generated. This behavior in the city is evolutive and changes dynamically not only over the time but also across the different conditions in the city e.g. seasonal, periodical, due to events and festivals, etc. The dynamic changes in a city demand the adaptation of technology and the expansion and contraction of services based on public demands and the evolutive process. All cities aspire to having more digital services while deploying services and technology.

In this particular scenario, the H-KPI Measuring Method is applicable in the third part of this flow which is used not only to evaluate the capacity of a city to satisfy smart citizens demands, but also to provide guidance and identify the parts where a city needs to act in order to scale up solutions. Strategically cities define plans and implement actions that will drive them towards achieving those objectives by using the Scale UP KPIs approach to the assessment to identify the gaps by means of a numerical value that can be referenced to a plan or to a defined evolutive scale. Cities are the creators of KPIs and producers of the data that will have metrics for achievement, however traditionally KPIs are assigned to technical aspects or ones that impact the condition separately.

The H-KPIs Smart City vision is to use the H-KPIs measuring method to assess a city according to the overall conditions and thus realize numerically whether the KPIs under analysis are below or above the expected limits. This is an estimation to assist in changing the conditions in the model and may be performed before the actions to modify the KPI conditions are applicable to understand the impacts of change beforehand, in other words it will work very well as a strategic tool for smart city planning. Scale-UP KPI explorer method will help to optimize efforts (Maximum Benefit - Optimal Efforts) to produce the expected results at the time of implementing services or deploying technology in a smart city.

Co-funded by the Horizon 2020
Framework Programme of the European Union

9 References

[1] Griffor, E, Greer, C, Wollman, D, Burns, M (2017) Framework for Cyber-Physical Systems, Volume 1 Overview (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 1500-201 https://doi.org/10.6028/NIST.SP.1500-201

[2] Greer, C, Burns, M, Wollman, D, Griffor, E (2019) Cyber-Physical Systems and Internet of Things (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 1900-202 https://doi.org/10.6028/NIST.SP.1900-202

[3] Serrano, M. Nguyen Mau Quoc, H, Le Phuoc, D. Hauswirth, M. et al. (2015) Defining the Stack for Service Delivery Models and Interoperability in the Internet of Things: A Practical Case With OpenIoT-VDK. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7012040

[4] Kazmi, A, Zappa, A, Serrano, M. (2016) Overcoming the heterogeneity in the internet of things for smart cities International workshop on interoperability and open source solutions, https://www.researchgate.net/profile/Zeeshan-Jan/publication/314205793_Overcoming_the_Heterogeneity_in_the_Internet_of_Things_for_Smart_Cities/links/58b99ab4aca27261e51fdfa7/Overcoming-the-Heterogeneity-in-the-Internet-of-Things-for-Smart-Cities.pdf

[5] Bosch, P, Jongeneel, VR, Neumann, H-M, Airaksinen, M, Huovila, A (2017) CITYkeys indicators for smart city projects and smart cities. Available at https://nws.eurocities.eu/MediaShell/media/CITYkeystheindicators.pdf

[6] Huovila, A, Bosch, P, Airaksinen (2019) Comparative analysis of standardized indicators for Smart sustainable cities: What indicators and standards to use and when? *Cities* 89:141-153. https://doi.org/10.1016/j.cities.2019.01.029

[7] Hara M, Nagao T, Hannoe S, Nakamura J. (2016) New Key Performance Indicators for a Smart Sustainable City. *Sustainability* 8(3):206. https://doi.org/10.3390/su8030206

[8] Personal, E, Guerrero, JI, Garcia, A, Peña, M, Leon, C (2014) Key Performance indicators: A useful tool to assess Smart Grid goals. *Energy* 76:976-988. https://doi.org/10.1016/j.energy.2014.09.015

[9] Rhee, S (2020) 2019 Global City Teams Challenge: Smart and Secure Cities and Communities Challenge Expo (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 1900-204 https://doi.org/10.6028/NIST.SP.1900-204

[10] Serrano, M. , Griffor, E. , Wollman, D. , Dunaway, M. , Burns, M. , Rhee, S. and Greer, C. (2022), Smart Cities and Communities: A Key Performance Indicators Framework, Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, https://doi.org/10.6028/NIST.SP.1900-206,

10 Glossary

AI	Artificial Intelligence
API	Application Program Interface
AV	Autonomous Vehicle
CPAC	Cybersecurity and Privacy Advisory Committee
CPS	Cyber-Physical System
CS/P	Cybersecurity/Privacy
DHS	U.S. Department of Homeland Security
DHS S&T	DHS Science and Technology Directorate
DOC	U.S. Department of Commerce
GCTC	Global City Teams Challenge
GPS	Global Positioning System
H-KPI	Holistic Key Performance Indicator
H-KPIs	Holistic Key Performance Indicators
ICT	Information and Communications Technology
IES	IoT-Enabled Smart City
IoT	Internet of Things
KPI	Key Performance Indicator
KPIs	Key Performance Indicators
NGI	Next Generation Internet
NIST	National Institute of Standards and Technology
NTIA	National Telecommunications and Information Administration
R&D	Research and Development
SBAC	Smart Building Action Cluster
SDN	Software Defined Networking
SSC	Super Stream Collider
SC3	Smart and Secure Cities and Communities Challenge
VNF	Virtual Network Function
WIT	Waterford Institute of Technology

Co-funded by the Horizon 2020
Framework Programme of the European Union