CROSS-CPP

Creating access to sensor data from various industrial sectors to enable new dimension of innovative business ideas

NEWSLETTER, ISSUE #3

ECOSYSTEM FOR SERVICES BASED ON INTEGRATED CROSS-SECTORIAL DATA STREAMS FROM MULTIPLE CYBER-PHYSICAL PRODUCTS AND OPEN DATA SOURCES

Letter from the Coordinator

Welcome to the third edition of our project newsletter.

From the very beginning, we were absolutely convinced that Data Markets have to become more attractive for its key stakeholders (data owners/providers and consumers) to overcome existing obstacles, such as the limited access to multiple data streams, or privacy concerns. Thus, we have built up an Open Ecosystem to empower Data Owners to exploit and control their most valuable data assets from smart products and to give Data Customers access to this great spectrum of sensor data. All along, we have followed the maxim to think about the needs of Data Owners and Data Customers, but also to win smart product manufacturers (e.g. car makers) to open up their (cyber physical) products, by designing a secure and trustworthy Ecosystem.

Recently we have finalized the implementation and integration of the Cross-CPP Ecosystem and have started its validation and assessment by data providers and data customers. A first public presentation of our Data Marketplace solution and Analytic Toolbox was presented a short time ago in the BDVe Webinar 'How to monetize your data in an open data Marketplace', which is still accessible via the BDVe website. In this newsletter you will find some more details about the Data Analytics Toolbox that has been developed, with a special focus on a tool for building advanced machine learning (ML) models from smart product data, as well as on the Cross-CPP big data marketplace.

In the near future we would like to come back to you with some validation and assessment actions planned in cooperation with the ReachOut project (https://www.reachout-project.eu/), to validate and optimise the solution based on your expertise. We hope to continue the conversation with you—our readers—as the Cross-CPP ecosystem evolves.

Enjoy the read, and please contact us with your feedback or questions!

Christian Wolff

In this issue

POLITÉCNICA "Ingeniamos el futuro" | CAMPUS DE EXCELENCIA INTERNACIONAL

Data Analytics Toolbox

In line with the milestones set in the previous newsletters, the Data Analytics Toolbox has been developed to provide valuable information in a wide range of use scenarios.

Meteologix, Siemens and Volkswagen have collaborated in this project providing data and expertise from their respective areas of business (weather forecasting, as well as building and vehicle automation, respectively). This has resulted in greatly informed and feasible use cases, which guarantees the applicability of this R&D project's outcomes.

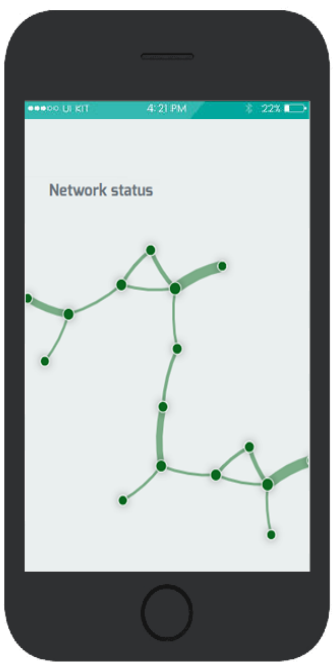
Amongst the elicited scenarios, we can find:

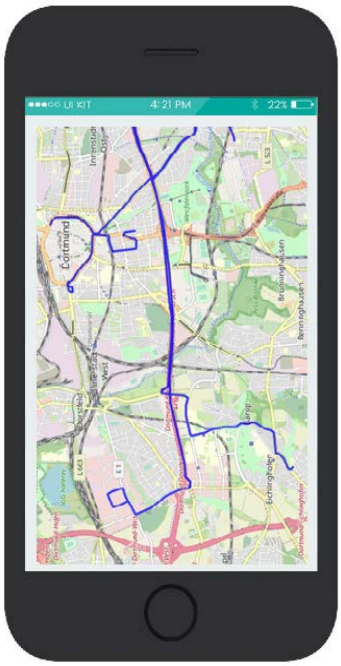
1) The identification of **malfunctioning sensors** on both buildings and vehicles.

2) The grouping of **similar trajectories**, enabling more precise analysis of each group's characteristics.

3) The ability to determine how the **weather is influencing the measurements** made by your car's sensors.

This approach is based on the **combination of data** from all available sources (buildings or vehicles), instead of exclusively relying in local data.





In the end, this means more and better information to **improve the services** you use! By providing your available information to the model, you are simultaneously contributing to the refinement of these services and benefiting from their improvement.

By capitalizing on the wide variety of available signals (temperature, speed, rain intensity…) and the unified format used for their storage, the possibilities that arise from the **joint usage** of the available analysis are endless.

From obtaining an evolving **comparison of the available signals** in the system to being able to automatically **fixing anomalies** from a vehicle's GPS coordinates (as depicted in the previous graphs) service providers will be able to continuously develop **added value** to your device's functionalities.

What's even more, the Analytics Toolbox functionalities support **real-time processing** of data, meaning that a continuous flow of fresh information can help you get the most out of your favourite services.

BRNO UNIVERSITY OF TECHNOLOGY

Building ML Models from CPP Data

In addition to specialised functions of the Data Analytics Toolbox, Cross-CPP also offers a generic interface for building advanced **machine learning** (ML) models from large CPP data available through the Data Marketplace. It interconnects the system with existing **popular frameworks** – Scikit-learn, Tensorflow, and PyTorch.

As expected, the **processing pipeline** differs from that of other analytics components – the **training phase** can take a significant time so that it is run in an **asynchronous** manner. Two **modes** of ML model building and applying are distinguished – **batch** and **streaming**.

The **Batch mode** employs Marketplace **Data Views** as an input and provides outputs directly in the **response** of the endpoint calls.

ML model - batch

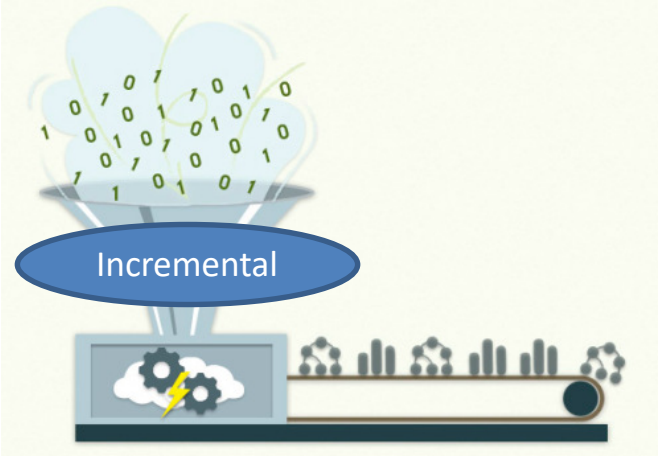
POST	/ml_model/batch/build
POST	/ml_model/batch/apply

There are also **common functions** relevant for both the processing modes – providing operation **status**, **export**, and **import** of the models, etc.

POST	/ml_model/status	POST	/ml_model/export
POST	/ml_model/drop	POST	/ml_model/import

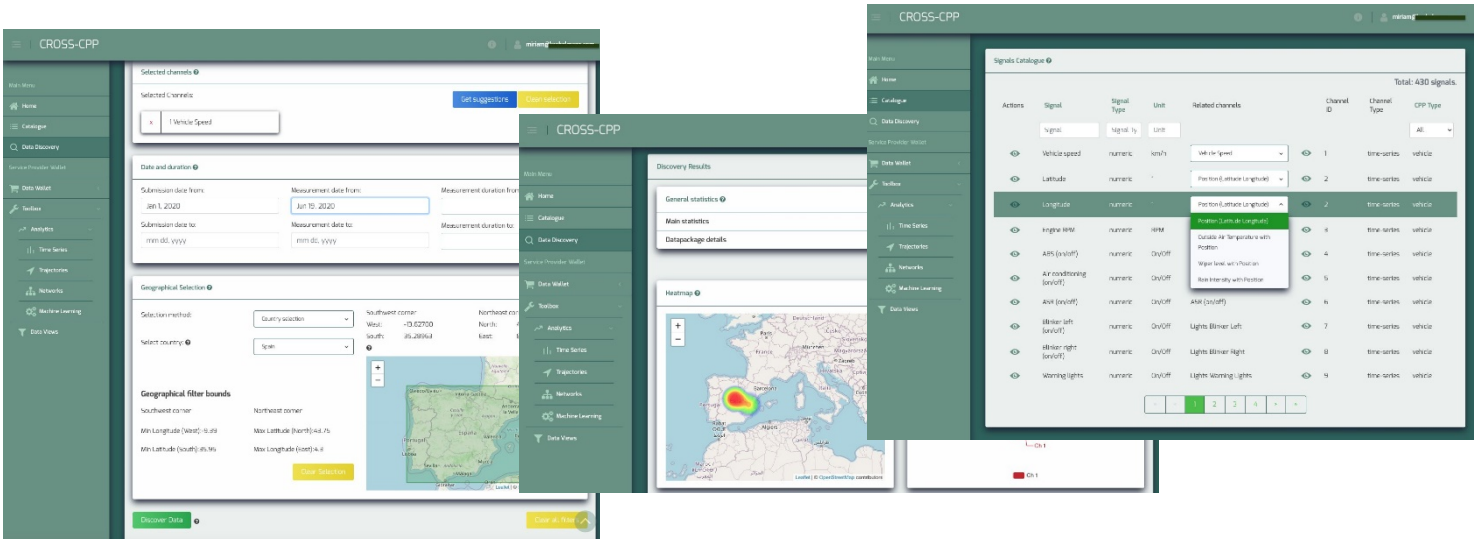
The streaming ML services take input from a **data feed** (**Aeon** channels or **Kafka** topics) and can produce results in the same type of the data stream

Incremental ML methods are used to create a model from data streams. The model building then corresponds to a **continuous process** – the model can be updated every time a new data point appears.



The ML services have been used to build a model predicting the use of **Siemens eChargers** by electric vehicles and to create a classification model for **driving behaviour** relevant to **car insurance** (smart) pricing.

What Cross-CPP Data Marketplace offers?



Agora

Brokering of brand-independent **data** from the different CPP Data Providers

API-based: unlocking **new** B2B and B2C data-based **products** and **services.**

Cross-CPP data marketplace (aka AGORA) connects Data Providers and Data Consumers for selling and acquiring Connected Vehicle and Home Building data under the standardized data model (CIDM, Common Industrial Data Model). It offers a secure and privacy preserving experience when selling or buying sharing big data, by having the full control over your data shared, to whom and for what purposes.

Catalog | Offering | Contracts | Exchange | New solutions

- Brokering of **brand-independent data** from the different IoT devices allowing Services Providers to create **new** B2B and B2C data-based **products** and **services in the new Data Economy era**.
- Management of **permissions and consent(s)** (future smart contract provision service).
- Data exchange **compliant** with current **standards** and **regulations** (GDPR, ISO20028, EXVE standard).
- **Secure** and **reliable** end-to-end communication (from the cloud to the Service Providers).
- **Flexibility to incorporate data coming from various industrial sectors** (interoperability) and to extend it to future upcoming needs by both, Data Providers and Data Customers **thanks to CIDM**.
- Easy to use **data discovery functionalities** for Service Providers, enabling an easy access, search/detection and selection of required data. Search for more than 200 sensor signals, and retrieve those datasets in a seamless experience
- A wide **data catalogue** and delivery of statistics for Service Providers. Advanced **visualization representations** (Histograms, Geo-Histograms, Time Series)

Cross-CPP dissemination

The Cross-CPP development work and final prototype components have been presented in several relevant events and for various audiences:

- Summit of **BDV in Porto** (February 2020)
- The 8th International Conference on Cyber-Physical Systems and Internet-of-Things – **CPS&IoT** (June 2020)
- **BDVe Webinar** How to monetize your data in an open data Marketplace (June 2020)

twitter

Linkedin

Final steps of the project

We entered the last part of the project in June 2020. Final prototypes of key Cross-CPP ecosystem components have been delivered and made available for their evaluation in project demonstrators. There are still many relevant milestones in the coming months related to the validation, but also updates to our joint and individual exploitation plans, standardization of results, and further engagement of UIG (User Interest Group) members and other early adopters of our results. Reports to be finished also include testing and final assessment protocols, as well as the report on the Cross-CPP methodology for using integrated data streams from cross –sectorial CPPs.

CROSS-CPP

ATB

SIEMENS Ingenuity for life

MX Meteologix AG

Atos

THE Open GROUP

POLITÉCNICA "Ingeniamos el futuro" | CAMPUS DE EXCELENCIA INTERNACIONAL

BRNO UNIVERSITY OF TECHNOLOGY

More information in https://cross-cpp.eu/

The Cross-CPP project receives funding under the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 780167.

Co-funded by the Horizon 2020 programme of the European Union