Paper ID # EU-TP0997EU-TP0997

Big data platform and big data analytics for ITS applications - pilot project proposal

Yaroslav Domaratsky, PhD
CTO, Head of Engineering, Sreda Software Solutions,
Sedova street, bld. 12, St. Petersburg, 192019, Russia
+7 931 203 0230, yaroslav@sredasolutions.com

Abstract
The full potential of the big data for ITS applications is yet to be investigated, but there the diversity of solutions now being developed demonstrates the potential. In the paper we propose high level architecture to integrate big data platform and big data analytics with the information coming from service providers and discuss how to utilize the above data to provide new services to citizens, smart (digital) cities and government agencies. We also discuss how to utilize information coming from V2x ground infrastructure and Proof-of-Concept demo we developed in our company.

KEYWORDS:
Data analytics, ITS, V2x

1. High level system architecture
Currently ITS data is fragmented and available through different ITS service providers and data aggregators. Data fragmentation and data diversity grows when we integrate ITS data sources with smart city data sources. To effectively utilize the above data sources we propose to deploy big data platform which accommodates data coming from ITS service providers, OEM service providers, data aggregators, traffic management systems, government agencies and other actors as described in the figure below.

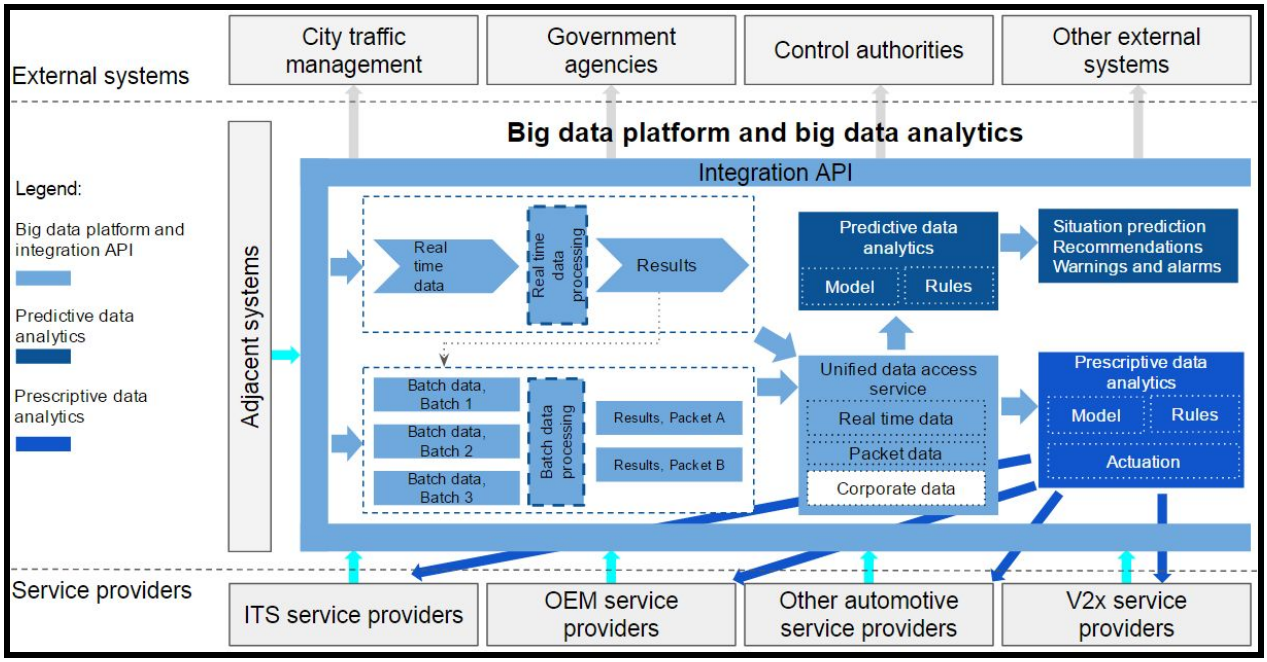


Figure 1: High level system architecture

Big data platform implements Lambda architecture pattern [1] and provides the following features:

- Collect, process, organize and preserve data coming from different sources

- Real-time and batch data analytics
- High reliability and availability
- Unlimited horizontal scalability
- Software as a Service (SaaS) and Platform as a Service (PaaS) capabilities
- Shortened software development and delivery life cycle through DevOps process
- Decrease in software development and maintenance efforts by using microservices.

Based on our experience we propose to implement the big data platform using proven Open Source Software (OSS) components such as Hadoop, Kafka, ELK, Redis, HBase, MongoDB, Docker, etc.

This provides the following benefits:

- High system reliability and availability
- The solution is not tightened to a particular vendor
- Software quality audits could be organised effectively
- There are no license fees for the base platform software
- Long term technical roadmap is in place, simple migration to new software versions
- OSS experts community is ready to help with system design and applications development.

2. Predictive data analytics for ITS applications

Predictive data analytics subsystem shown in the Figure 1 provides the following features based on predicted situation such as environmental conditions, road conditions, traffic conditions, infrastructure state conditions, etc:

- For citizens
  - Effective prediction of the emergency situations
  - Traffic flows optimization
    - to increase citizen safety and to minimize traffic jams
    - to decrease fuel consumption and to minimize environmental impact
    - to increase cargo transportation efficiency to (from) seaports, airports, cargo hubs
    - to provide priority journey for emergency vehicles and public transport.
- For the transport (city) infrastructure owners
  - Predict transport (city) infrastructure life cycle
  - Decrease negative impacts for the transport (city) infrastructure
  - Optimize maintenance schedule for the transport (city) infrastructure
  - Predict and prevent failures of the transport (city) infrastructure
  - Network Operations Center (NOC) analytics
  - Predict and prevent regulation policy violations
- For government agencies
  - Predict government systems and infrastructure life cycle
  - Optimize maintenance schedule for the government systems and infrastructure
  - Predict and prevent failures in the government systems (infrastructure)
  - Automatically control suppliers performance, predict and prevent regulation policy violations
  - NOC analytics, predict and prevent emergency situations.

Sreda Software Solutions company already implemented similar prescriptive data analytics features for wireless communications NOCs [2].

3. Prescriptive data analytics for ITS applications

Prescriptive data analytics subsystem shown in the Figure 1 drives actions based on the situation prediction to:

- Minimize risks (increase safety level) for
  - Citizens
  - Transport (city) infrastructure owners
  - Vehicle OEMs and insurance companies
- Decrease the cost of ownership and maximize the return of investments for
  - Vehicle owners and fleet managers
  - Transport (city) infrastructure owners
  - Government agencies.

The proposed system also could be used to launch new business models in the multimodal cargo transportation, insurance telematics, car sharing, autonomous driving and other markets.

The shift from context aware systems to the systems with situation awareness and continuous insight based on prescriptive data analytics is shown in the Figure 2 [3].

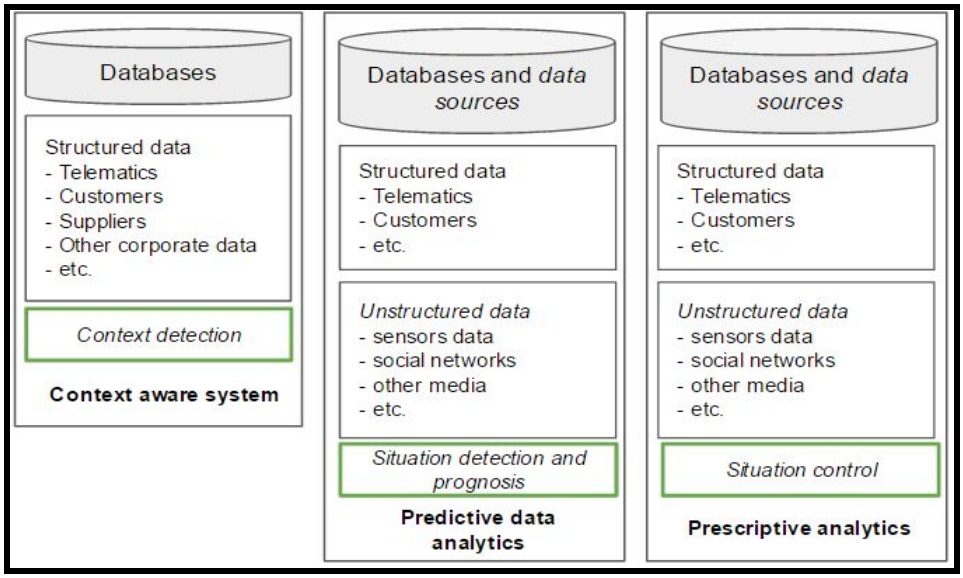


Figure 2: Context aware and situation aware systems

4. Using V2x systems as the data source

We believe V2x systems are the ideal data source for automotive big data analytics because we could collect a lot of up to date environmental and vehicle data over V2x interface. Therefore we propose to combine the big data platform and big data analytics subsystems described in the above sections with V2x prototype platform as shown in the Figure 3.

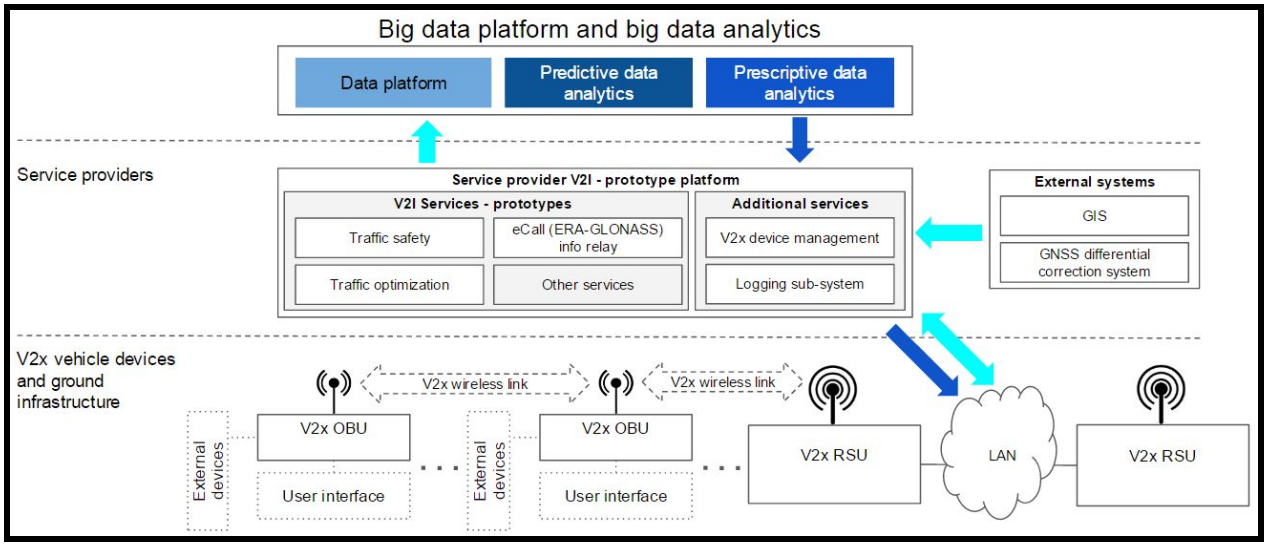


Figure 3: Integration with V2x prototype platform

If the above integration completed, then we could use real-time data coming from V2x system to validate prescriptive data analytics algorithms. Static data available in RDE database [4] and traffic simulators such as SUMO [5] could be used at initial stages if real V2x data is not available.

There is no final decision in Russia about V2x device architecture and standards to be used. To deploy prototype V2x system in Russia we should make decision on network, facilities and application layers per V2x device software structure shown in Figure 4.

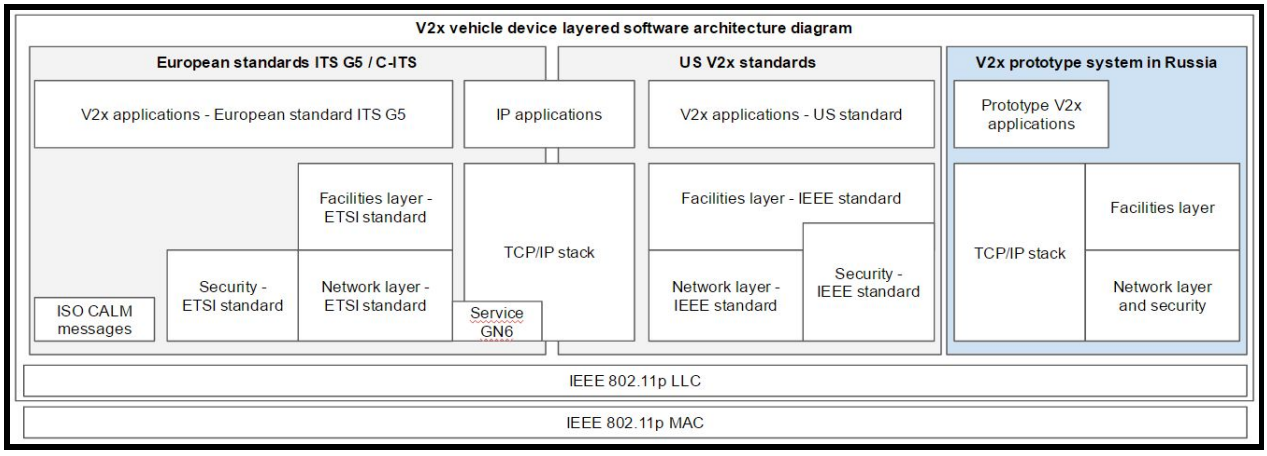


Figure 4: V2x OBU software architecture diagram

5. Implementation details

In our company we developed Proof-of-Concept (PoC) demo system based on publicly available OSS components. PoC system includes mobile clients, data platform, data analytics and machine learning sub-system and services sub-system. The general architecture diagram of the data platform is shown in the figure below.

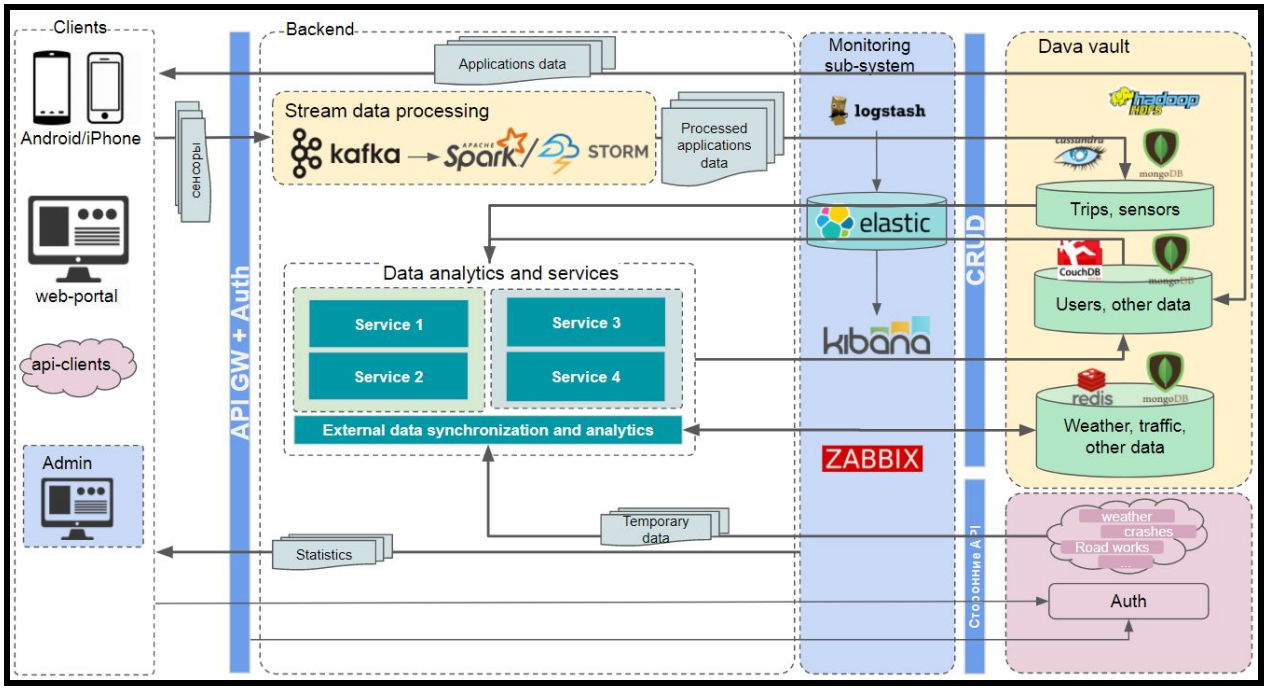


Figure 5: PoC demo system architecture

We also developed PoC for mobile applications. The general architecture diagram of the mobile application software is shown in the figure below.

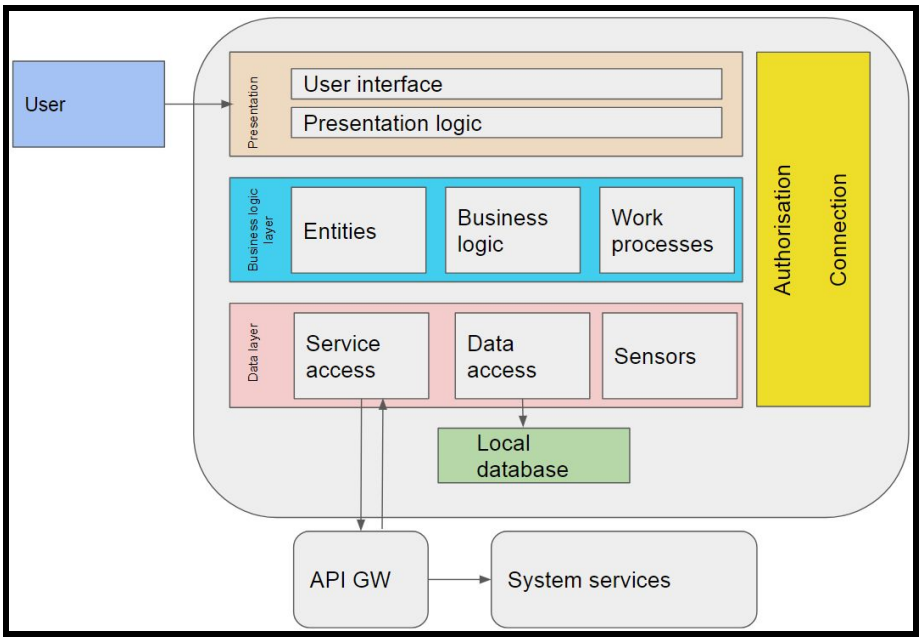
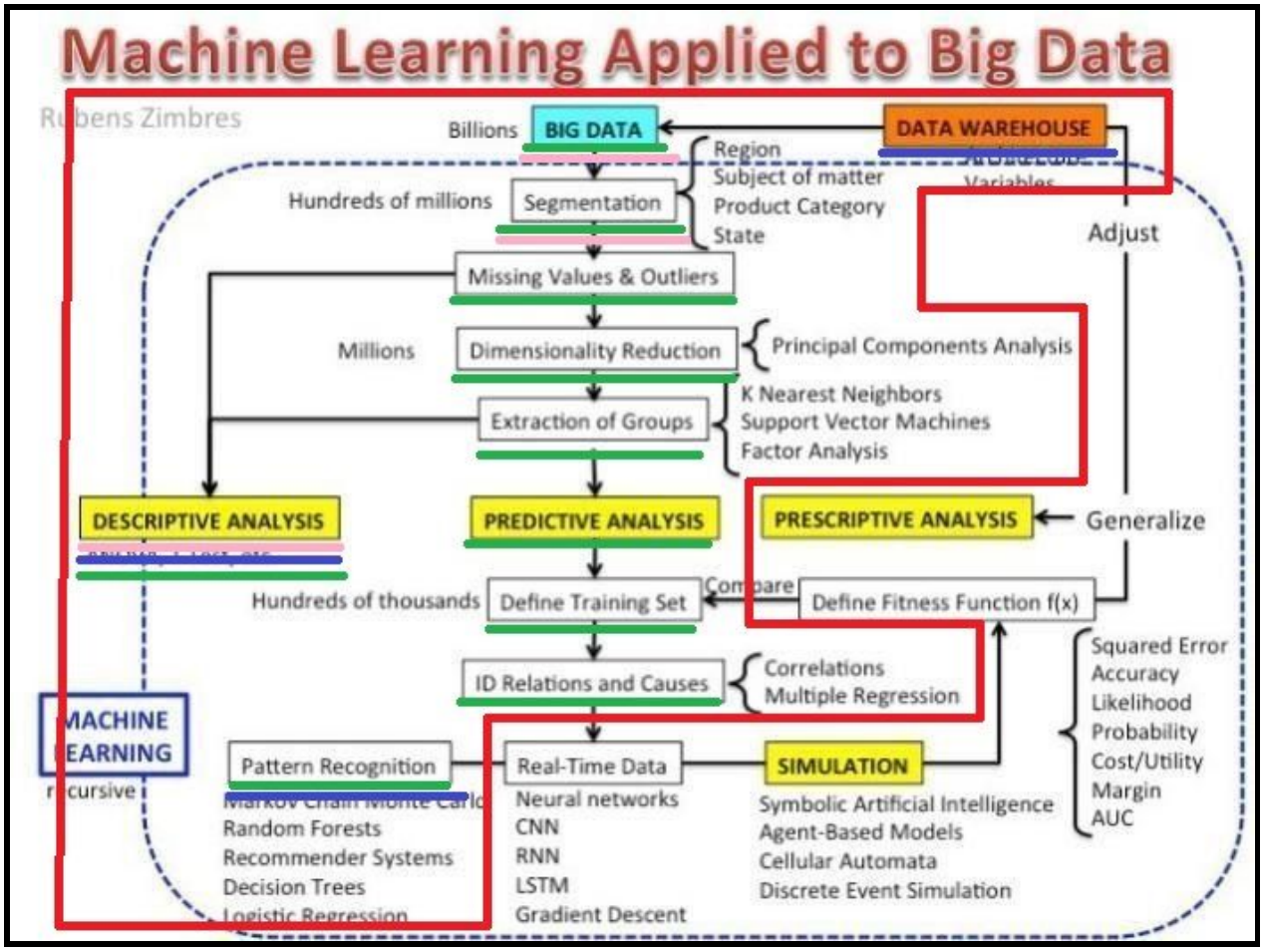


Figure 6: PoC demo mobile application architecture

Big data platform and big data analytics for ITS applications

Currently the data analytics logic implemented in perl programming language using NumPy and SciPi libraries. To train and verify data analytics algorithms we intensively use data collected in USDOT safety pilot and available through RDE database (https://www.its-rde.net/). Currently we concentrate on descriptive and predictive analytics as shown in the diagram below.



Information source: Mike Tamir post at Linkedin (https://www.linkedin.com/in/miketamir/)

Figure 7: Data analytics - current scope

In the next phase we plan to move focus to the prescriptive analytics.

You may refer to the below screen shots for the demo system UI.

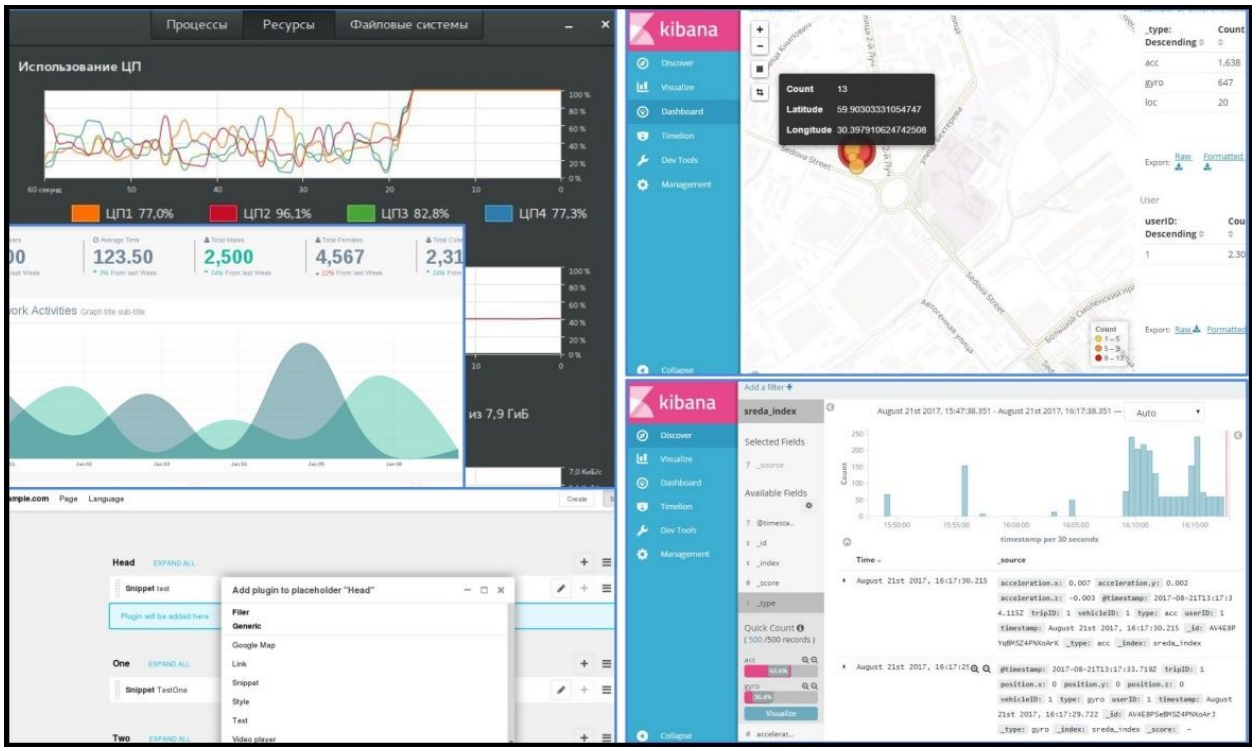


Figure 8: PoC demo screenshots

The demo system runs well and is ready for demonstration.

6. Pilot deployment plan

Sreda Software Solutions established contacts with more than 6 leading V2x technology providers. Based on the information we got from them we propose the following pilot development plan.

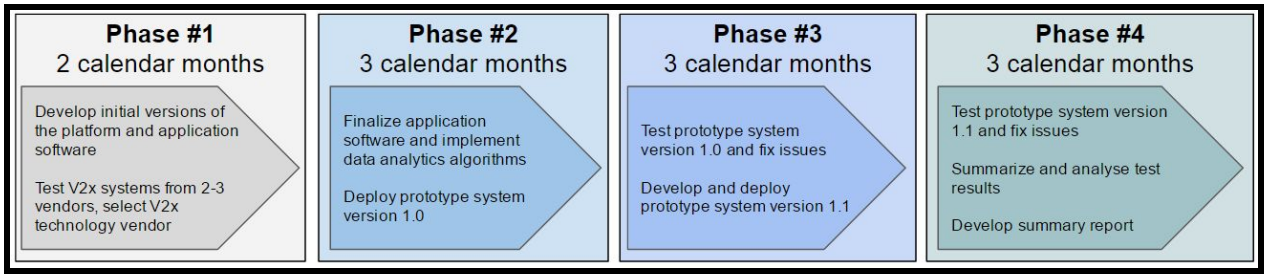


Figure 9: Pilot development plan

The main goals of the pilot formulated as follows:

- Develop and verify OSS based big data platform integrated with ITS service providers and data aggregators
- Develop and verify predictive data analytics algorithms including situation detection and situation prediction
- Develop and verify prescriptive data analytics algorithms including capability to minimize risks (increase safety level) for
  - Citizens and transport (city) infrastructure owners
  - Vehicle OEMs and insurance companies

capability to decrease the cost of ownership and maximize the return of investments for

- Vehicle owners and fleet managers
- Transport (city) infrastructure owners
- Government agencies.

7. Conclusion and next steps

We like to productize PoC system described in the Section 6. Please contact me at yaroslav@sredasolutions.com if you like to get more information.

At ITS World Congress 2017 we plan to discuss how the approach described in the paper could be aligned with ITS technology roadmaps. We also intend to continue discussion with V2x technology vendors and ITS service providers about early V2x (IEEE 802.11p, Cellular-V2x) technology trials in Russia.

References

1. Kinley, James. (2014). *The Lambda architecture: principles for architecting realtime Big Data systems*, http://jameskinley.tumblr.com/post/37398560534/the-lambda-architecture-principles-for
2. *Sreda Software Solutions company Web site,* http://sredasolutions.com
3. *Systems of Insight for Digital Transformation. Using IBM Operational Decision Manager Advanced and Predictive Analytics,* IBM Redbooks
4. *USDOT Research Data Exchange,* https://www.its-rde.net/
5. *SUMO – Simulation of Urban MObility*, http://www.dlr.de/ts/en/desktopdefault.aspx/tabid-9883/16931_read-41000/