

## **SUBMISSION TO EUROPEAN DATA FORUM 2015**

Contributor names (name, surname)	Sebastian Grabowski Jarosław Legierski Maciej Grzenda
Short CVs	<p>Sebastian Grabowski graduated from the Faculty of Electronics and Telecommunications, Koszalin University of Technology, Faculty of Management at the University of Warsaw and the Warsaw School of Social Psychology. His research focuses on Open Data, Open API, interactions between users of telecommunications networks and the Internet as well as issues related to Open Government. For the past 12 years he has been involved in the telecommunications industry, currently he works in Orange Poland telecom, a part of global Orange telecom group. He is the Director of Orange Poland Research and Development Center. He is a co-founder of the Open Middleware 2.0 community – a program bringing together different communities of developers, researchers, technology suppliers and IT companies in the field of telecommunications and IT.</p> <p>Dr. Jarosław Legierski received his M.Sc. in electronics and telecommunication and Ph.D. degree in electronics from Technical University of Lodz. Since 1998 he has worked in the telecommunications industry, between 1999 and 2011 as an Integration Expert in Siemens in B2B telecommunication area. Currently he is an R&amp;D Expert in Research and Development Center of Orange Poland telecom. He is the co-creator of Open Middleware 2.0 Community (<a href="http://www.openmiddleware.pl">www.openmiddleware.pl</a>). His research interests include: next-generation Service Delivery Platforms, Unified Communications, Open Application Programming Interfaces (APIs) and Open Data. He is the main architect and system integrator of the platform exposing Open Data of the City of Warsaw (<a href="http://api.um.warszawa.pl">http://api.um.warszawa.pl</a>). Author of publications in the area of exposure of Open API and Open Data.</p> <p>Maciej Grzenda, PhD is an Assistant Professor at the Warsaw University of Technology, Faculty of Mathematics and Information Science, as well as a Senior Project Manager in the Research and Development Center of Orange Poland. He received his MSc and PhD from the Warsaw University of Technology in 1997 and 2001, respectively. He is a researcher in the field of machine learning (ML), with particular emphasis on industrial applications of ML techniques. In 2014, he received his habilitation in Computer Science. He proposed new data preprocessing methods to be used</p>

	when modelling technical processes with ML techniques. His works have been published in journals in this field including Applied Soft Computing and Journal of Intelligent Manufacturing and concentrate on the processing of heterogeneous data including relational, time series, geospatial, and semi-structured data.
Type of the presentation	Impact contribution
Title of the presentation	<b>Developer in the Loop – Towards Community-Driven Open Data and Open API Exposure</b>
Summary of the presentation	An important challenge is involving wider communities in the development and use of Open Data portfolios. The presentation summarizes social aspects of selected Open Data initiatives in Poland. A fundamental aspect of these works was the constant participation of developer communities in requesting, selection, exposition and evaluation of individual data sets and APIs. Different forms of developer participation in Open Data exposition are discussed. Particular attention is paid to the project that has already opened 250 data sets of the city of Warsaw in parallel with developer contests using Open Data, but also Open API for 3Vs data sets.

### **EXTENDED ABSTRACT OF THE PRESENTATION**

Open Data initiatives are undertaken by numerous organizations. A frequently raised question is the way Open Data sets should be published. This involves several different aspects. One of key questions being asked is the sequence of opening individual data assets an organization has. The way large, fast changing data sets should be made available is also a major issue, especially when sensor data of large volume and velocity is planned to be exposed. Last, but not least the way the use of the open data sets can be promoted is still being investigated.

The presentation summarizes the experience collected in the community of the largest Polish cities, software developers, NGOs, and business angels. Such a community has been organized and co-developed by Research and Development Center of Orange Poland telecom over last few years. During this period, community events such as international developer contests have been organized. More precisely, in the last two years, Research and Development Center of Orange Polska has organized multi-stage contests dedicated for developers, namely Open API Hackathon in 2013, and Business Intelligence Hackathon API (BIHAPI, <http://www.bihapi.pl/>) in 2014. In 2013, in the first edition only data sets from the

city of Warsaw were used. During the last edition of the contest, which was finished this year, open data assets from three major Polish cities: Warsaw, Poznań and Gdańsk were made available for participants.

The use of individual Open Data sets and requests made by developers have been monitored, which provided basis for the further development of Open Data road maps by individual cities. Also, the ways of increasing the adoption of Open Data by software developers were analysed. This included putting developers in the loop and taking into account their expertise when tuning APIs and data formats used in the exposition process. These initiatives provided basis for projects integrating academic, commercial, city and NGO partners. A key project that builds on the previous experience of community-driven development of Open Data is Apps for Warsaw (*Dane Po Warszawsku* in Polish).

Apps for Warsaw is the project oriented on the exposition of Open Data from the City of Warsaw. The main goal of project is developing a platform for aggregating and sharing public data, allowing anyone to access these data as well as undertaking activities leading to creating innovative applications based on the data. As a result of the project, a hybrid platform for open data exposure was created, which is composed of two parts: CKAN for static data exposition and Function server for the exposition of dynamic and real time open data sets. Function server was developed in Java and was made available for public use as open source. Until now expose 253 data sets have been made available via CKAN or Function server. All of them are API oriented. The APIs have been developed in cooperation with independent software developers and are available at <https://api.um.warszawa.pl>. The data sets include raster map layers, vector map layers, geo time series (such as location of trams), relational data. This includes near-real time data.

A major part of the project is also a developer contest announced at <https://konkurs.danepowarszawsku.pl/en>. Over 70 ideas of applications have been submitted to the competition. What should be emphasized is that the project is performed not by the city itself, but involves NGOs, commercial and academic partners. It is an R&D project co-founded by The National Center for Research and Development. Importantly, the project is not performed as a part of technical research programs, but within Social Innovations program.

Project partners of Apps for Warsaw are NGOs: TechSoup (leader), Digital Center: Project Poland Foundation, The Unit for Social Innovation and Research „Shipyard”, Orange R&D Center Poland, and Warsaw University of Technology. The project is executed in close cooperation with the City of Warsaw.

The presentation will summarize the experience of building community-driven Open Data exposure, show the impact of technical decisions (data formats, APIs) on the development of innovative applications, and investigate key data categories that attracted particular attention of software developers. Moreover, lessons learned from the cooperation of NGOs, academic, commercial and city partners will be summarized. A proposal of best practices when opening city data with developer participation will be made.