Theory and practice of linked open statistical data

Tambouris, Efthimios; Kalampokis, Evangelos; Janssen, Marijn; Matheus, Ricardo; Hermans, Paul; Kalvet, Tarmo

DOI
10.1145/3209281.3209341

Publication date
2018

Document Version
Final published version

Published in
Proceedings of the 19th Annual International Conference on Digital Government Research

Theory and Practice of Linked Open Statistical Data

Efthimios Tambouris
University of Macedonia and ITI-CERTH
Thessaloniki, Greece
tambouris@uom.gr

Evangelos Kalampokis
University of Macedonia and ITI-CERTH
Thessaloniki, Greece
ekal@uom.edu.gr

Marijn Janssen
Delft University of Technology
Delft, The Netherlands
m.f.w.h.a.janssen@tudelft.nl

Ricardo Matheus
Delft University of Technology
Delft, The Netherlands
r.matheus@tudelft.nl

Paul Hermans
ProXML bvba
Keerbergen, Belgium
paul@proxml.be

Tarmo Kalvet
Tallinn University of Technology
Tallinn, Estonia
tarmo.kalvet@ttu.ee

ABSTRACT

The number of Open Statistical Data available for reuse is rapidly increasing. Linked open data technology enables easy reuse and linking of data residing in different locations in a simple and straightforward manner. Yet, many people are not familiar with the technology standards and tools for making use of open statistical data. In this tutorial, we will introduce Linked Open Statistical Data (LOSD) and demonstrate the use of LOSD technologies and tools to visualize open data obtained from various European Countries. We will also give the participants the opportunity to use these tools thus obtaining a personal experience on their capabilities.

CCS CONCEPTS

• **Information systems** → **Information systems applications**;

KEYWORDS

Linked Open Data, Linked Open Statistics, ICT Tools

ACM Reference Format:
Efthimios Tambouris, Evangelos Kalampokis, Marijn Janssen, Ricardo Matheus, Paul Hermans, and Tarmo Kalvet. 2018. Theory and Practice of Linked Open Statistical Data. In *dg.o '18: dg.o 2018: Proceedings of the 19th Annual International Conference on Digital Government Research, May 30-June 1, 2018, Delft, Netherlands,* Anneke Zuiderwijk and Charles C. Hinnant (Eds.). ACM, New York, NY, USA, Article 4, 2 pages. https://doi.org/10.1145/3209281.3209341

1 INTRODUCTION

During the last years, large volumes of highly structured numerical data have been released online for others to reuse [2]. These Open Statistical Data (OSD) describe financial, social, and political aspects of the world, and thus play crucial role for being a major element of not only the government economic and social but also business decision-making process. Publishing OSD however does not automatically yield benefits [6]. This difficulty of efficiently re-using OSD is explained, among others, by the fragmented environment of OSD [5]. Linked data technologies facilitate the discovery and integration of data on the Web because they semantically annotate data and uniquely define a piece of data using URIs [1] Linked data have been recently recommended by W3C as the most effective way to open up data on the Web (W3C, 2017). However, as this technology is emerging we are lacking methods and tools to manage Linked Open Statistical Data (LOSD). OpenGovIntelligence is a project funded to bridge this gap. In particular, OpenGovIntelligence has developed a number of tools that facilitate the exploitation of LOSD in a simple and straightforward manner. This tutorial could help interested stakeholders to:

- Understand the nature and characteristics of LOSD
- Become familiar with the OpenGovIntelligence tools and their functionalities
- Practice on exploiting real LOSD using OpenGovIntelligence tools
- Become aware of practical experiences and recommendations for using the tools shared by the OpenGovIntelligence team.

2 AUDIENCE, PRIOR KNOWLEDGE AND COURSE MATERIAL

The target audience of this tutorial includes different types of participants of DGO. More specifically, the tutorial should be particularly interesting for:

- Government professionals, since it proposes a new way for publishing Statistical Data exploiting Linked Data technologies. The authors will also share the experience they gained during the application of LOSD in real PAs.
- Software developers, because they will familiarize themselves with the developed tools and their functionalities. In addition, they will have the opportunity to evaluate them from a technical point of view.
- Researchers and anyone interested in LOSD.

dg.o '18, May 30-June 1, 2018, Delft, Netherlands

ACM ISBN 978-1-4503-6526-0/18/05.
https://doi.org/10.1145/3209281.3209341

No specific knowledge is required to follow the tutorial. However, basic knowledge of linked data technologies (e.g. RDF) as well as the multidimensional (cube) data model might facilitate following the tutorial. The material to be handed out to the attendees will be the tutorial slide-set.

3 THE OPENGOVINTELLIGENCE PROJECT

The work presented in this tutorial is conducted within the EU funded project OpenGovIntelligence (OpenGovIntelligence.eu) [3]. This project continues work that started within the EU funded project OpenCube (opencube-project.eu) [4].

The OpenGovIntelligence project aims to modernize PA by connecting it to Civil Society through the innovative application of LOSD. The publication of high quality LOSD can transform society, services and enterprises throughout Europe. Towards this end, the project seeks to support approaches for service co-creation; whilst providing software tools to aid decision making and better manage the complexities and precise nature of high quality LOSD.

The development and adoption of new business processes, policies, and tools will enable the active participation of society and enterprise in data sharing and in the co-production of innovative data-driven public services. The objectives of the project objectives include:

- Identifying the challenges involved in opening, managing, and representing LOSD for the co-production of innovative data-driven public services.
- Creating a framework of business processes, policies, and data infrastructure.
- Delivering an ICT toolkit comprising user centred tools, documentation and specifications that can help realise an eco-system of connected data and services.
- Proving the effectiveness of the approach through pilots in six countries. The pilots will develop services at national and local levels to tackle challenges within society and PA such as enabling better decision-making, enhancing e-services provided by Points of Single Contact, and improving policy-making.

4 TUTORIAL OUTLINE

This is a 3-hour tutorial. The tutorial consists of presentations, a brief hands-on session and time for discussions. The tutorial programme follows.

(1) Welcome and tour de table of participants.
(2) Presentation of the OpenGovIntelligence project.
(3) Presentation of LOSD.
(4) Presentation of LOSD tools and practical experiences.
(5) Hands-on session on using LOSD tools.
(6) Discussion and feedback.

5 PRESENTERS BIOGRAPHIES

Ass. Prof. Efthimios Tambouris is Associate Professor of Information Systems and eGovernment at the Applied Informatics Department of the University of Macedonia, Greece. He has served as an expert in standardisation activities at CEN, ERC and the European Commission.

Dr. Evangelos Kalampokis is an Adjunct Lecturer at the Business Administration Dept. of the University of Macedonia, Greece and a research fellow with Information Systems Lab at the same university. He is a co-chair at the SemStats workshop, co-located with the International Semantic Web Conference.

Prof. Dr. Marijn Janssen is a full Professor in ICT & Governance and chair of the ICT section of the Technology, Policy and Management Faculty of Delft University of Technology. He is Co-Editor-in-Chief of Government Information Quarterly, conference chair of IFIP EGOV series and is chairing mini-tracks at the DG.o, ICEGOV, HICCS and AMCIS conferences.

Ricardo Matheus MSc is a researcher and PhD Candidate at Delft University of Technology, The Netherlands. Currently, he is working in the OpenGovIntelligence, where he leads the Pilots Operation and Evaluation WP and contributes to dissemination, capacity building and exploitation of the project.

Paul Hermans is the lead architect of ProXML, a company offering consultancy and implementation services in the field of (linked) open data. ProXML has been involved in open data and semantic projects at the EC, Dutch and Flemish government, Provinces of Limbourg and Flemish-Brabant, City of Louvain.

Dr. Tarmo Kalvet is a Senior Research Fellow in the field of technology governance at Ragnar Nurkse Department of Innovation and Governance. He has published more than 70 research and policy-analysis papers in the fields of eGovernment, research and innovation, business and economics.

ACKNOWLEDGMENTS

The work reported this paper is funded by the European Commission within the H2020 Programme in the context of the project OpenGovIntelligence (http://OpenGovIntelligence.eu) under grand agreement No. 693849.

REFERENCES

[1] Bizer C, Heath T, and Berners-Lee T. 2009. Linked data-the story so far. *Semantic Services, Interoperability and Web Applications: Emerging Concepts* (2009), 205–227.
[2] Kalampokis E, Tambouris E, Karamanou A, and Tarabanis K. 2016. Open Statistics: The Rise of a New Era for Open Data?. In *15th Annual International IFIP Electronic Government Conference (EGOV)*. Springer, 31–43.
[3] Kalampokis E, Tambouris E, and Tarabanis K. 2016. Linked Open Cube Analytics Systems: Potential and Challenges. *IEEE Intelligent Systems* 31, 5 (2016), 89–92.
[4] Kalampokis E, Tambouris E, and Tarabanis K. 2017. ICT Tools for Creating, Expanding, and Exploiting Statistical Linked Open Data. *Statistical Journal of the IAOS* 33, 2 (2017), 503–514.
[5] Hassani H, Saporta G, and Silva E S. 2014. Benefits, adoption barriers and myths of open data and open government. *Data Mining and Official Statistics: The Past, the Present, and the Future* 2, 1 (2014), 31–43.
[6] Janssen M, Charalabidis Y, and Zuiderwijk A. 2012. Benefits, adoption barriers and myths of open data and open government. *Inf. Syst. Manag* 29, 4 (2012), 258–268.