Chapter 41
Integrating Semi-Open Data in a Criminal Judicial Setting

Mortaza S. Bargh
Research and Documentation Centre (WODC), Ministry of Security and Justice, The Netherlands

Sunil Choenni
Research and Documentation Centre (WODC), Ministry of Security and Justice, The Netherlands & Research Centre Creating 010, Rotterdam University of Applied Sciences, Rotterdam, The Netherlands

Ronald F. Meijer
Research and Documentation Centre (WODC), Ministry of Security and Justice, The Netherlands

ABSTRACT

Judiciary systems comprise various partner organizations (e.g., police, public prosecutor, courts, and rehabilitation centres) that collaboratively resolve criminal cases. These partner organizations have their own data administration and management systems, which are setup/operated separately and integrated barely. This chapter explains the approach of the authors' organization for integrating the data sets of the Dutch judiciary systems, and for opening the data integration outcomes to the public and/ or to specific groups. These outcomes (e.g., data sets and reports) are meant to provide useful insights into (the performances of) the partner organizations individually and collectively. Such data opening efforts do not comply with all Open Data requirements, mainly due to the quality, (privacy) sensitivity and interoperability issues of the raw data. Nevertheless, since these initiatives aim at delivering some benefits of Open Data, the chapter introduces the new paradigm of Semi-Open Data for acknowledging such data opening initiatives.

INTRODUCTION

In recent years we have observed a surge in Open Data movements. The term Open Data refers to data that is reusable and distributable by third parties without any costs (ODef, 2015). The idea of Open Data has widely gained popularity for several good reasons. For a government, Open Data could be a means to strengthen democracy. By opening up its data, a government provides transparency into its operations,

meets regulatory compliance, and increases public participation and collaboration. Open Data is also expected to boost economic developments and innovations. Commercial companies and start-ups can exploit the data to develop new applications and services such as recommender apps (Choenni, Bargh, Roepan, & Meijer, 2015) and to create new types of professions such as data journalists and fact checkers. Furthermore, businesses and individuals can use Open Data to support their strategic and personal decisions. For example, individuals can use neighborhood statistics for buying houses and for implanting trees and plants in polluted areas. Scientists can reuse the data for reviewing scientific results as well as for their own evidence-based research.

Data opening, however, appears to be a tedious and almost infeasible process for many organizations. There are several impediments of Open Data identified in the literature, see for example (Zuiderwijk, Janssen, Choenni, Meijer, & Sheikh Alibaks, 2012; CoA report, 2015). These obstacles range from privacy concerns to technical issues. In this chapter, the authors argue that acting according to the current definition of Open Data (ODef, 2015) is too challenging for the judicial field. For example, three typical challenges are privacy, legacy and interoperability of judicial data. First, opening judicial data 'as it is' often contravenes the privacy law and regulations because such data pertain to real-life persons and encompass privacy sensitive attributes like names, birth-dates, crime types, and sentences. Even if these sensitive information items are removed or anonymized, there is a chance to infer such sensitive information by combining several junks of information. As an example of such inference, suppose that a school decides to publish the average exam marks of its male and female students. If we know that only two female students took part in an exam and if we are able to derive from social media that one of these female students got a 9, then the result of the other female student can be deduced.

Opening judicial data as raw as possible, as required in Open Data (OD101, 2015; Wonderlich, 2010), may put data analysts on the wrong track due to the legacy nature of judicial data. The semantics of judicial data may evolve over time since rules and regulations are subject to frequent changes. For example, crimes are categorized into different categories – the well-known categories of crimes include violent crime, theft, burglary, drugs offence, drunk driving, vandalism, and economic crime (Moolenaar, Choenni, & Leeuw, 2007). The crime categories today might be different from those categories of ten years ago. For example, ten years ago the Internet related crimes were assigned to the category of 'others' or 'economic crime', while today there is a call for a new category 'cybercrime'. If a new category is added or an existing category is deleted in a categorization system, then the objects need to be redistributed among the new set of categories. For a proper interpretation of the results obtained by analyzing or querying such data, it is of vital importance to know how the semantics of the data have evolved over time. Suppose that someone is interested in knowing how the category economic crime has developed over the last ten years, during which the objects among the categories have been redistributed due to the introduction of the new category of cybercrime. This redistribution may cause the occurrence of a trend break. For a proper interpretation of such a break in the trend, therefore, one should know how and when the redistribution was carried out.

The strength of Open Data lies in the ability of integration and combination of a data set with other data sets. Therefore, interoperable data sets are required. Unfortunately, however, many data sets in the judicial field are not interoperable and a significant amount of effort is required to process these data sets to make them interoperable. In the judicial field, there are many systems belonging to various partner organizations that collaboratively resolve criminal cases, e.g., police, public prosecutors, courts, and rehabilitation centres. Each organization has its own core business, specific rules, and regulations to meet. Furthermore, in the field of justice, it is important to satisfy the principle of 'trias politica' (i.e.,

Related Content

Libraries, Digitized Cultural Heritage, and Social Cohesion of Smart Cities: Model-Like LIS-Educational Implementations in Hungary

János Fodorand Péter Kiszl (2022). *Handbook of Research on the Role of Libraries, Archives, and Museums in Achieving Civic Engagement and Social Justice in Smart Cities (pp. 334-354).*

www.irma-international.org/chapter/libraries-digitized-cultural-heritage-and-social-cohesion-of-smart-cities/291407

Heritage and Museums as Objects of Education for Citizenship in the Teaching of History

Ainoa Escribano-Miralles, Pedro Miralles-Martínezand Francisca-José Serrano-Pastor (2020). *Handbook of Research on Citizenship and Heritage Education (pp. 103-126).*

www.irma-international.org/chapter/heritage-and-museums-as-objects-of-education-for-citizenship-in-the-teaching-of-history/246778

Civic Technology and Data for Good: Evolutionary Developments or Disruptive Change in E-Participation?

John G. McNuttand Lauri Goldkind (2022). *Research Anthology on Citizen Engagement and Activism for Social Change (pp. 1330-1345).*

www.irma-international.org/chapter/civic-technology-and-data-for-good/295058

An Isolated and Behind-the-Scenes Work Called Conservation: Developing Collaboration and Community Engagement of Conservation Practice in Indonesia

Widiatmoko Adi Putranto, Indah Novita Sariand Regina Dwi Shalsa Mayzana (2022). *Handbook of Research on the Role of Libraries, Archives, and Museums in Achieving Civic Engagement and Social Justice in Smart Cities (pp. 1-15).*

www.irma-international.org/chapter/an-isolated-and-behind-the-scenes-work-called-conservation/291387

ICTs: Ancillary Tools for Indirect Democracy?

Kerill Dunne (2017). *Politics, Protest, and Empowerment in Digital Spaces (pp. 91-106).*

www.irma-international.org/chapter/icts/173915