A PLATFORM TO ENRICH, EXPAND AND PUBLISH LINKED DATA OF POLICE REPORTS

Bruno C. N. Oliveira, Ivan Salvadori, Alexis Huf and Frank Siqueira
Department of Informatics and Statistics
Federal University of Santa Catarina, Florianópolis, Brazil

ABSTRACT

Motivated by transparency policies, government agencies have been publishing more valuable data. In order to take advantage of this initiative, data must be published in accordance with Linked Open Data principles. This work proposes a platform capable of enriching police reports with data available on the Web. This work also presents an evaluation tool, which consumes the linked data provided by a Web API, expands it with data extracted from other sources and provides an interface to assess the relevance of the expansions. Results show that the proposed expansion method might be useful for discovery and evaluation of related information, supporting richer analysis of the published data.

KEYWORDS

Linked Data, Web API, Semantic Web

1. INTRODUCTION

Government data have been published due to the growth of public transparency initiatives. The availability of this data is crucial to provide information to society as well as to encourage the framing of new action plans focused on improving the living conditions of citizens. However, it is essential to publish government data according to certain guidelines in order to allow its proper use. Among these guidelines, we highlight the principles that encompass the Linked Data publication, which seek to semantically connect and enrich information. Furthermore, the Linked Open Data (LOD) publication allows underlying datasets to be easily queried, combined, analyzed and reused.

Currently, researchers have been interested in exploring the challenges and possibilities of providing linked data in view of the increasingly relevant movement to publish government data. Authors as Kalampokis et al. (2013) and Chan (2013) show how to reuse published government data and perform mashups to provide deeper and detailed analysis. Lebo and Williams (2010) present a conversion technique of structured open data into linked data. Entity expansion approaches from heterogeneous sources and semantic similarity between resources are covered by Pereira Nunes et al. (2012).

The main contribution of this work is to propose a platform for enrichment, expansion and publishing of linked data, as well as an evaluation tool to assess the relevance degree of expansions. The platform comprises a data retrieval tool, which obtains non-structured data and converts it into structured data, a semantic Web API to enrich these dataset and publish linked data, and a tool that combines different data sources in order to expand the entities with additional related information. The evaluation tool enables the user to assess the expansions in order to enrich the entities. In this paper, we use data published by the Public Security Secretariat of the state of São Paulo (SSP/SP)[1], in Brazil, which discloses police reports information in a non-structured format. The data sources adopted for enriching information contained in the police reports are the Google search engine, which provides all sorts of information available on the Web, and the Court of Justice of São Paulo (TJSP)[2], which provides data on criminal and civil lawsuits in the same region where the police reports were filed. The results of this study point out that the enrichment, expansion and provisioning

[1] http://www.ssp.sp.gov.br/transparenciassp/
[2] http://tjsp.jus.br/

of linked data can assist the discovery and analysis of new relevant information. Moreover, it allows the effective reusing of data.

The remainder of this paper is organized as follows. Section 2 presents the required background for the understanding of this work. Section 3 describes some related work found in the literature. The semantic Web API for enrichment and publishing of linked data relating to police reports is described in section 4. Section 5 presents a tool to assess the police reports expansions. The assessment method is presented in section 6 while section 7 shows the results of the evaluation study. Finally, section 8 draws some conclusions regarding the results obtained with this work and discusses some foreseen improvements.

2. BACKGROUND

According to Bizer et al. (2009), the web has a huge potential to be a global space of linked data. In this view, not only the documents are available and linked to other documents, but also their content. Thus, data is structured in the form of subject-predicate-object triples in which the subject represents the described resource, the predicate represents a feature of the subject, and the object is the value assigned to this feature. Triples are described based on the RDF (Resource Description Framework) standard, which defines a model to describe the semantics of resources and associate them in order to enable that other information be linked to them (Gandon and Schreiber, 2014).

The concept of linked data is defined by Berners-Lee (2006), who presents a set of best practices for publishing structured data on the Web. These principles seek to establish links amongst different sets of data to form a single global data repository, leading to a powerful integration and reusing of data. Among the principles that conduct linked data publishing, HTTP URIs are used to identify and locate resources, which should be structured in the form of RDF triples. In addition, published datasets should include links to extra information. By sharing the same semantic definition base, through maximum reuse of predicates coupled with URIs, it can be created a network of connections that forms the basis of a novel Web.

Another important concept in the context of Linked Data is the Mashup approach, which combines content from different data sources aiming at engendering new resources. However, the lack of a high-level language to describe queries in heterogeneous sources is one of the issues faced in the development of mashup applications (Magalhães et al., 2013). According to Di Lorenzo et al. (2009), building mashups is often accomplished through several invocations to Web APIs, which do not follow a pattern. In addition, the information retrieved from each API is not linked with the remainder of the data involved in the mashup. This characteristic results in isolated data that hinders the integration process.

Another key concept for the understanding of our proposal is the REST architectural style (Fielding, 2000), which is a collection of principles and architectural constraints concerning the development of distributed web applications. One of the core concepts of REST relies on the resources, which can be addressed by a unique identifier (URI) and available to remote clients through representations. A resource can be a list of films showing in a cinema, user comments in a blog or a profile on a social network, for example. In addition, there are different ways to represent a resource. A representation is a sample of the state of a resource at a given moment of time, which can be described using a variety of formats such as HTML, XML, JSON, and so on.

3. RELATED WORK

Kalampokis et al. (2013) and Chan (2013) show the possibilities of information analysis and data integration due to the growth of open government data available on the Web. A mechanism to convert government data published in CSV format into linked data is suggested by Lebo and Williams (2010). The authors developed a tool with RDF vocabulary to describe how CSV datasets should be transformed into linked data. Apart from CSV, this tool neither addresses other dataset formats nor non-structured data.

In the field of mashups, DiFranzo et al. (2011) proposed an approach to create mashups from linked open government data. The work of Minhas et al. (2012) presents a framework driven to the study and evaluation of mashup tools. The authors sought to improve the current mashup development projects with approaches more focused on the users. They have performed an exploratory and qualitative analysis to study the

strategies of mashup development tools, the technical resources adopted in its implementation, the graphical resources and the support offered to users. Our goal, on the other hand, is to perform a mashup presented by an evaluation tool, so that users can assess the relevance of the enriched content.

Regarding expansion, Pereira Nunes et al. (2012) present an approach to identify relationships between entities in different datasets. The degree of similarity among entities is measured according to the number of paths and the distance between the entities in a graph. Their approach utilizes social networking algorithms to analyze the graphs of datasets. Thus, the similarity between the different resources based on existing semantic connections is explored. In addition, expansion approach using the Bing search engine was employed to identify the correlation between entities based on co-occurrence of terms on the web. Similarly, our platform uses the concept of entities expansion in order to retrieve additional relevant information from different resources and utilizes the Google web search engine to perform one of the expansions.

The approaches proposed by Lebo and Williams (2010) and DiFranzo et al. (2011) consider only structured data as input parameter for mashups generation, whilst our proposal also takes into account non-structured data. The framework proposed by Minhas et al. (2012) allows assess the usability of and supports of mashup tools to guide new development strategies, although disregards the mashup's content relevance assessment. Pereira Nunes et al. (2012) evaluate, by means of algorithms, the connectivity degree between entities. The documents relatedness is measured based on the co-occurence of terms, yet a real checking if the content is indeed about the same object under study is not applied. On the other hand, the purpose of our study is to publish linked data retrieved from non-structured data sources without semantic annotation. In addition, our goal focuses on the assessment, carried out by humans, of entity expansions to properly enrich the published linked data with additional relevant information gathered from other resources.

4. WEB API FOR ENRICHMENT AND PUBLISHING LINKED DATA

The main component of the proposed platform is a REST Web API that offers data from police reports extracted from the website of the SSP/SP system. The Web API provides the information from the reports in a structured form and semantically annotated with terms of vocabularies and ontologies used in linked open data. In addition, police reports are enriched with geographic coordinates of the incident site.

4.1 Data Retrieval and Structuring

The SSP/SP system classifies the police reports according to their category (suspicious death, intentional homicide, robbery followed by murder, among others). Information about the report such as location, category, police station and date of the incident are available. Also, information of those involved (victims and perpetrators) are also provided, such as name, document number, gender, occupation, skin color, marital status, among others. However, detailed description of the fact is not disclosed, narrowing the accurate understanding of the incident.

In spite of being a substantial initiative of information transparency to society, the SSP/SP system does not publish information in a suitable format to be reused by other systems or even analyzed by researchers. The reports are published in a non-structured format in HTML pages created dynamically and with no semantic enrichment, disregarding basic principles of linked data, necessary to achieve an appropriate level of public transparency. Furthermore, a search mechanism to retrieve available data is not supplied, so it becomes unfeasible to fetch police reports based on certain desired features.

We have developed a tool, available at *Figshare* (https://dx.doi.org/10.6084/m9.figshare.3856074), that is able to retrieve police reports from the SSP/SP system and store data into a structured format. As a result, for each report that is processed, the tool generates a document in JSON format, which is more suitable to be consumed by other systems. In addition, the tool also generates documents in CSV format with all the information obtained from police reports, becoming suitable to be processed by data analysis mechanisms.

4.2 Data Publishing

Access to information obtained from police reports retrieved from the SSP/SP website is provided through a semantic RESTful Web API. As depicted in Figure 1, by means of JSON documents generated previously as

well as Google Maps API, the Web API of police reports provides semantic representations enriched with geographical data. Two representations are provided: police reports in JSON-LD (JSON for Linked Data) format (Lanthaler and Gütl, 2012) and maps in HTML format. These outputs are consumed by the evaluation tool described in section 5, and afterwards will be enriched with additional information. Some instructions of how to access and manipulate the APIs are also available in the *Figshare* repository.

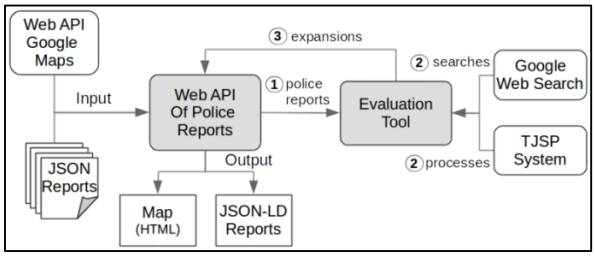

Figure 1. Architecture of the Web API of police reports and Evaluation Tool

The semantic enrichment of provided representations adds semantic annotations to JSON documents, generating JSON-LD representations, which associate objects and properties in the document to classes and properties defined in ontologies. JSON-LD representations are compatible with the RDF data model and can be converted to triples by several existing libraries, for instance, JSON-LD-Java (https://github.com/jsonld-java/jsonld-java). On the other hand, the location pointed out in a police report is enriched by invoking the Google Maps API to convert addresses into geographic coordinates, which can be used to place markers on a map.

Figure 2 (a) shows an example of a police report, categorized as suspicious suicide, represented in JSON-LD. This representation is associated with four ontologies and the entity attributes are related to entities and terms of these ontologies. This example also illustrates the enrichment of a report with its address set as geographic coordinates and with other details about the location, such as the city, state and postal code. Finally, "rdfs:seeAlso" properties are used as expansions to link the report with other related entities. In contrast, Figure 2 (b) depicts the original non structured police report, provided by the SSP/SP system, which has no semantic embedded and none mechanism of LOD to be reused by other resources.

(a)


```
{ "@context": {
   "foaf": "http://xmlns.com/foaf/0.1/",
   "dbo": "http://dbpedia.org/ontology/",
   "schema": "https://schema.org/",
   "ssp": "http://exemplo.com.br/ssp-ontology/",
   "rdfs": "http://www.w3.org/2000/01/rdf-schema#" },
 "ssp:reportNumber": "6923",
 "ssp:dateOfIncident": "16/12/2015",
  ...
 "dbo:victim": [ {
    "@type": ["foaf:Person", "schema:Person"],
    "foaf:name": "JACQUELINE DO CARMO LIMA LEONE",
    "ssp:personId": "OMMITED-INFORMATION",
    "dbo:nationality": "dbpedia.org/resource/Brazil",
    "foaf:gender": "Female",
    "dbo:birthDay": "OMMITED-INFORMATION",
    "dbo:occupation": "Cashier",
    "dbo:skinColor": "Brown",
    "dbo:deathCause": "dbpedia.org/resource/Category:Unsolved_deaths",
     ... }   ],
 "schema:location": {
   "@type": "Place",
   "schema:address": {
     "@type": "schema:PostalAddress",
     "schema:addressLocality": "Santo Andre",
     "schema:addressRegion": "SP",
     "schema:postalCode": "09190-615",
     "schema:streetAddress": "Rua Doutor Henrique..." },
   "schema:geo": {
     "@type": "schema:GeoCoordinates",
     "schema:latitude": "-23.670621",
     "schema:longitude": "-46.533615" }
 },
 "rdfs:seeAlso": "http://.../mulher-baleada-por-ladroes",
 "rdfs:seeAlso": "http://tjsp.jus.br/..."
}
```


(b)

SECRETARIA DE ESTADO DA SEGURANÇA PÚBLICA
POLICIA CIVIL DO ESTADO DE SÃO PAULO

Dependência: 01° D.P. SANTO ANDRÉ **Folha: 1**

Boletim No.: 6923/2015 **Iniciado:** 16/12/2015 10:38hs **e Emitido:** 16/12/2015 10:38hs

Boletim de Ocorrencia de Autoria Conhecida

Natureza(s):

Espécie: Óbito
Natureza: Morte suspeita
Consumado

Local: RUA DOUTOR HENRIQUE CALDERAZO, 321 - PARQUE CENTRAL
CEP: 09190615 - S.ANDRE - SP
Tipo de Local: Saúde
Circunscrição: 01º D.P. SANTO ANDRÉ
Ocorrência: 16/12/2015 PELA MANHÃ
Comunicação: 16/12/2015 às 22:34 horas
Elaboração: 16/12/2015 às 22:38 horas
Flagrante: Não

Vitima(s) e Autor(es)/Vitima(s):

JACQUELINE DO CARMO LIMA LEONE (Vitima) - Vitima Fatal - RG: [illegible] - Natural de: MAUA -SP - Nacionalidade: BRASILEIRA - Sexo: Feminino - Nascimento: [illegible] 23 anos - Estado Civil: Casado - Profissão: CAIXA - Instrução: 1 Grau completo - Cutis: Parda - Naturezas Envolvidas: Morte suspeita

Histórico:

"Os históricos dos boletins de ocorrência são públicos. Em regra, serão fornecidos mediante solicitação ao SIC (Serviço de Informação ao Cidadão), excepcionando-se somente os dados necessários a preservação da intimidade, vida privada, honra e imagem das pessoas, nos termos da Lei nº 12.527 de 18 de novembro de 2011."

Exames requisitados: IML

Solução: ENCAMINHAMENTO DP ÁREA DO FATO

01° D.P. SANTO ANDRÉ

Endereço da delegacia: R. XAVIER DE TOLEDO, 48 - CENTRO - S.ANDRE-SP 09000-110
Telefone: (11)4438-1133

15/06/2016 10:38:07

Figure 2. (a) Example of police report provided by the: (a) Web API in JSON-LD and (b) SSP/SP in a Web page

The proposed Web API allows querying data by using search filters, such as the category of police reports, its year or identification number, as well as filters applied to those who were involved in the incident, such as gender, education level and age. Additionally, the Web API provides map representations with markers locating all police reports selected from the filters mentioned above. For instance, it is possible to position markers in a precise place on a map where crime against women were committed in the year of 2015, as depicted in Figure 3. In this way, the Web API provides custom geographic views of police reports. The representations available in JSON-LD are suitable to be consumed by other systems, allowing such information to be used in different ways and for different purposes. Moreover, HTML representations accurately show the location of incidents on a map, thereby providing a geographic view over a set of reports.

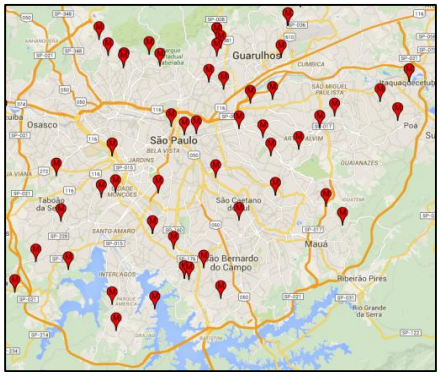

Figure 3. Locations of homicide against women

5. EVALUATION TOOL

The platform proposed in this paper includes a tool that performs a mashup to evaluate the expansion of entities that represent a police report. This tool uses data provided by the Web API of police reports and comprises a web application as an interface for end users and a REST API to manipulate a) the expansions of reports and b) the assessment regarding the relevance of an expansion. Two kinds of expansions are supported: the results from Google web search and the results of lawsuits registered in TJSP. Based on the identification number (available in the report) of those who were involved in the reported incident, the tool performs the expansion of judicial data from TJSP which offers data on lawsuits related to the victims or authors of the crime. Similarly, it retrieves the first result from Google web search based on the category of the police report and the name of those who have been involved in it. Issues of homonyms in the later expansion can be dealt with, since the user can analyze the whole content of the result, namely features of the incident (location, date, time), characteristics of the person (gender, age, skin color, occupation) and the type of crime. This may rule out homonyms who are unrelated to the incident, leading to a more accurate assessment.

Primarily, an assessment is started with the creation of a sample of police reports. The user must inform to the tool a category (homicide, robbery, suspicious suicide, etc.), a year and a desired amount of police reports to be assessed. From the reports provided by the Web API, as depicted in Figure 1 (1), the evaluation tool selects a simple random sample without replacement of police reports that fit the filters passed by the user. Afterwards, both expansions - Google and TJSP (2) - are performed and the obtained data is stored. At the end of evaluation, the result is submitted to the Web API (3), in order to enrich the corresponding entity.

Figure 4 illustrates the user interface of the evaluation tool showing a police report categorized as a suspicious suicide. At the top of the tool, the main information about the report and the parties involved are displayed. Below that, expansions from Google search engine and results of lawsuits obtained from TJSP are presented to the user. Finally, the relevance options for each expansion are provided, so that users can choose the best option according to their analysis. The button "Assess and Continue" submits the assessment result and displays the next enriched police report for evaluation.

Figure 4. User Interface of the Evaluation Tool

6. ASSESSMENT METHODOLOGY

Effectiveness of entity expansions was conducted through qualitative experiments over the dataset under study. We performed the assessment in a manual procedure assisted by the evaluation tool presented in the previous section. As the expansion process seldom fails, the failure cases were disregarded. In the expansion by the Google search engine, failure is defined as a lack of results. On the other hand, if the search engine has at least one result, relevant or not, the result is assessed. In the TJSP expansion, failures are not recorded, since the lack of results is indeed considered as a result, when a person has never been involved in a criminal or civil lawsuit in TJSP.

The population parameters under study are the proportions of police reports, whose expansions were classified by the evaluation tool based on three different ratings, which are mutually exclusive, for each expansion. The expansions from TJSP system are classified as *CRIMINAL* when at least one of those who were involved in the incident was a party to any criminal process. Similarly, when at least one civil lawsuit was found for any of those involved in the incident, yet no criminal process was returned, the expansion is then classified as *CIVIL*. Finally, if no lawsuit was found for any of the parties involved in the incident, the expansion is classified as *NO_PROCESS*. Concerning Google expansions, classifications are ordered in accordance with new relevant information retrieved. The most relevant classification, *WITH_HISTORY*, comprises the police reports for which expansion has gathered the incident detailed description - unpublished by the SSP/SP official system - corresponding to the report. The second classification, *RELEVANT*, covers cases where the expansion has gathered additional information relating to victims and/or perpetrator, though it does not describe the incident in detail. All other cases - when it is not possible to determine with reasonable accuracy whether the incident or parties involved are the same as in the report - are classified as *UNDETERMINED*.

Two categories of police reports were selected to assess the expansions: “intentional homicide” and a sub-category of “Suspicious Death” called “Reasonable questions about suicide or unnatural death”. The motivation for the choice of homicides is the large attention given to this type of crime by the media and public opinion, whereas in cases of suicides it is expected a more limited coverage by the media. The experiment covers police reports filed in the year of 2015.

Faced with the extensive process of assessing the high volume of reports, we have estimated, for each selected category, confidence intervals for the classification of expansions, using samples generated by the tool. In order to figure out the sample size required for 95% confidence level and a sampling error of 3%, a pilot study with one hundred police reports for each category was executed. In reference to the estimation of confidence intervals for each category of police reports, we have used a sample sized according to the classification parameter which has required a larger sample. For both categories, the *NO_PROCESS* classification of TJSP expansion determined the sample size. Hence, 748 reports were randomly picked amongst 4.485 murder reports, whereas 334 reports from 1.274 suspicious suicides were picked likewise.

7. RESULTS

Confidence intervals that generalize the proportions measured in the experiments for the entire population of police reports are shown in Table 1 and Table 2. The results of expansions, especially those based on Google, are interesting because, although the detailed description had not been provided by the official SSP/SP system, the history of some crimes was found in 17,78% of the police reports, with a margin of error of 2,74% at 95% confidence. In many cases, the expansions with history brought, besides the description of the incident, the full names of suspects and information about the personal life of victims and authors of the crime. It is worth mentioning that, in these experiments, Google expansion considered only the first result available. Moreover, we measured at 95% confidence and 3,248% margin of error, that parties involved in 28,877% of police reports categorized as homicide had already been involved in lawsuit(criminal and/or civil) registered in TJSP. The qualitative assessment of the expansions reveals a significant proportion of police reports that have been enriched with relevant information obtained from Google search engine and from the TJSP system. It is also expected that higher rates of relevant expansions can be obtained by applying heuristics, including other search engines as well as expanding the scope of expansion beyond the first result.

Table 1. Confidence interval at 95% for assessments of Google expansions

Classification	Homicide	Suicide
WITH_HISTORY	17,78 ± 2,740	6,287 ± 2,603
RELEVANT	8,824 ± 2,033	0,8982 ± 0,869
UNDETERMINED	73,396 ± 3,167	92,81 ± 2,770

Table 2. Confidence interval at 95% for assessments of TJSP expansions

Classification	Homicide	Suicide
CRIMINAL	22,460 ± 2,991	10,479 ± 3,285
CIVIL	6,417 ± 1,756	5,090 ± 2,357
NO_PROCESS	71,123 ± 3,248	84,431 ± 3,888

These statistics paint a quite interesting view of a predilection of homicide cases by the media: among 21 expansions with the history of reports categorized as suicide, 7 of these expansions brought descriptions of homicides, 3 about accidental deaths and 2 described murders in robbery cases. These results have been found even with the search query including the string “suicide”. Applying the confidence interval for these 12 cases, we can estimate for suicide reports with 95% confidence that Google expansion brings a detailed description apparently divergent of the original category of the report for between 1,597% and 5,589% of cases.

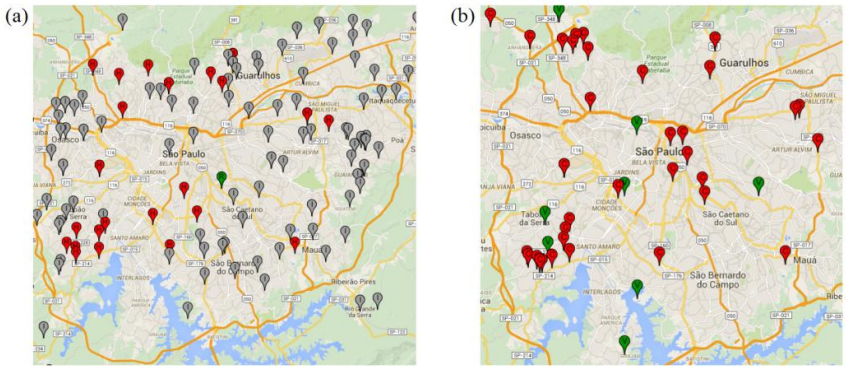

Figure 5. Geographical relevance of (a) Google expansion, and (b) TJSP expansion

Figure 5 shows the geographic coordinates of homicides from the sample concerning both Google expansion (a) and TJSP expansion (b). The expansions which were assessed as *UNDETERMINED* are identified with gray markers and the letter "I". Green ("R") and red ("H") markers indicate the geographic coordinates of homicides assessed as *RELEVANT* and *WITH_HISTORY*, respectively. Red markers identified

by the letter "C" represent police reports which have at least one party involved in the incident with criminal lawsuit in TJSP. Green ("V") markers represent instances in which some of those who have been involved in the incident have only civil lawsuits. Thus, the representations on a map for a set of expanded police reports can aid new exploratory analysis.

8. CONCLUSIONS AND FUTURE WORK

In this paper, we have introduced a platform for the enrichment, expansion and publishing of linked data, starting from a non-structured public dataset of police reports. Up to the writing of this paper, 28.992 reports have been extracted, structured and published in a public repository available at *Figshare* (https://dx.doi.org/10.6084/m9.figshare.3856074). The structured information is semantically enriched and provided by a Web API. Then, the content of each police report is expanded to retrieve other similar web documents with relevant information. In addition, an evaluation tool is presented in order to show how we assess the expansions and combine them with the resources.

Some outcomes of the assessment for entity expansions are also presented in this paper. During the evaluation process, we could identify some police reports registered in a category incompatible with the fact described in the expanded results. We have also found criminal and/or civil lawsuits relating to those who have been involved in the report, especially victims of homicide. These findings are only viable due to the content gathered from Google and TJSP expansions and can be important to work on deeper analysis and further investigation about the incidents.

Although the platform was developed in the domain of police reports, we intend to build a generic platform to expand and publish enriched data. Additionally, future work can also include additional data sources into the platform, for instance, new web search engines, social networks as well as other governmental systems. Instead of manually assess the quality of expansions, we might also apply techniques for named entity recognition, enabling the evaluation, in an unsupervised manner, of the relationship between expansions and the original entity.

REFERENCES

Berners-Lee, T., 2006. *Linked Data*. [Online] Available at: http://www.w3.org/DesignIssues/LinkedData.htm.

Bizer, C., Heath, T. & Berners-Lee, T., 2009. Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems*, pp.1-22.

Chan, C.M.L., 2013. From Open Data to Open Innovation Strategies: Creating E-Services Using Open Government Data. *46th Hawaii International Conference on System Sciences (HICSS)*, pp.1890-99.

Di Lorenzo, G., Hacid, H., Paik, H.-y. & Benatallah, B., 2009. Data integration in mashups. *ACM SIGMOD Record*, p.59.

DiFranzo, D. et al., 2011. The Web is My Back-end: Creating Mashups with Linked Open Government Data. In *Linking Government Data*. New York. pp.205-19.

Gandon, F. & Schreiber, G., 2014. *Resource Description Framework (RDF) Model and Syntax Specification*. [Online] Available at: http://www.w3.org/TR/rdf-syntax-grammar/.

Kalampokis, E., Tambouris, E. & Tarabanis, K., 2013. On Publishing Linked Open Government Data. *Proceedings of the 17th Panhellenic Conference on Informatics*, pp.25-32.

Lanthaler, M. & Gütl, C., 2012. On using JSON-LD to create evolvable RESTful services. In *Proceedings of the Third International Workshop on RESTful Design - WS-REST '12*. New York. p.25.

Lebo, T. & Williams, G.T., 2010. Converting Governmental Datasets into Linked Data. *Proceedings of the 6th International Conference on Semantic Systems*, pp.1-3.

Magalhães, R.P. et al., 2013. QEF-LD - A Query Engine for Distributed Query Processing on Linked Data. In *Proceedings of the 15th International Conference on Enterprise Information Systems*. pp.185-92.

Minhas, S.S., Sampaio, P. & Mehandjiev, N., 2012. A Framework for the Evaluation of Mashup Tools. In *Proceedings of the 2012 IEEE Ninth International Conference on Services Computing*. Washington. p.8.

Pereira Nunes, B. et al., 2012. Can Entities be Friends? *Proceedings of the Web of Linked Entities Workshop in conjuction with the 11th International Semantic Web Conference*, pp.45-57.