Mining Governmental Collaboration Through Semantic Profiling of Open Data Catalogues and Publishers

Mohamed Adel Rezk, Adegboyega Ojo, Islam A. Hassan

Mining Governmental Collaboration through Semantic Profiling of Open Data Catalogues and Publishers

Mohamed Adel Rezk, Adegboyega Ojo and Islam A. Hassan

Insight Centre for Data Analytics, National University of Ireland Galway, Galway, Ireland.
{mohamed.adel, adegboyega.ojo, islam.hassan}@insight-centre.org

Abstract. Due to the increasing adoption of open data among governments worldwide especially in the European Union area, a deeper analysis of the newly published data is becoming a mandate. Apart from analyzing the published dataset itself we aimed on analyzing published dataset catalogues. A dataset catalogue or a dataset metadata contains features that describe what the data is about in a textual representation. So, we first acquire data from open data portals, choose descriptive dataset catalogue features, and then construct an aggregated textual representation of the datasets. Afterwards we enrich those textual representations using Natural Language Processing (NLP) methods to create a new comparable data feature "Named Entities". By mining the new data feature we are able to produce datasets and publishers relatedness network. Those networks are used to point similarities between the published data across multiple open data portals. Pointing all possible collaborations for integrating and standardizing data features and types would increase the value of da1ta and ease its analysis process.

Keywords: unstructured data analysis, data mining, collaborative network; open data, e-government.

1 Introduction

Despite the availability of data loaded into open data portals worldwide[1] [1, 2], methods to maximize stakeholders' engagement and ease data integration still not complete [3–5]. We believe that a proper mining of collaboration channels within a single data portal internally as well as between multiple open data portals are not introduced yet. Our work is aiming to develop an open data portals collaboration channels mining framework as shown in Fig. 1. To achieve this, we start with data acquisition by harvesting metadata of datasets published on the portal then restructure and store them in MongoDB[2]. Afterwards we construct textual representation from the dataset metadata's unstructured features, apply DBpedia [6] Named Entity Recognition pipeline called DBpedia Spotlight [7] to extract information that

[1] http://opendatabarometer.org/
[2] https://www.mongodb.com

represent those dataset and their publishers as well. After that we end up with a semantically enriched dataset upon which we can apply our profiling [5] and collaboration opportunities analysis. To illustrate our work, we organized the paper as follows: Section 2 presents a background on Open Government Data, NLP and Collaboration Mining. Section 3 discusses our approach to tackle the research question. Section 4. Discussing our research findings, conclusions and future plan.

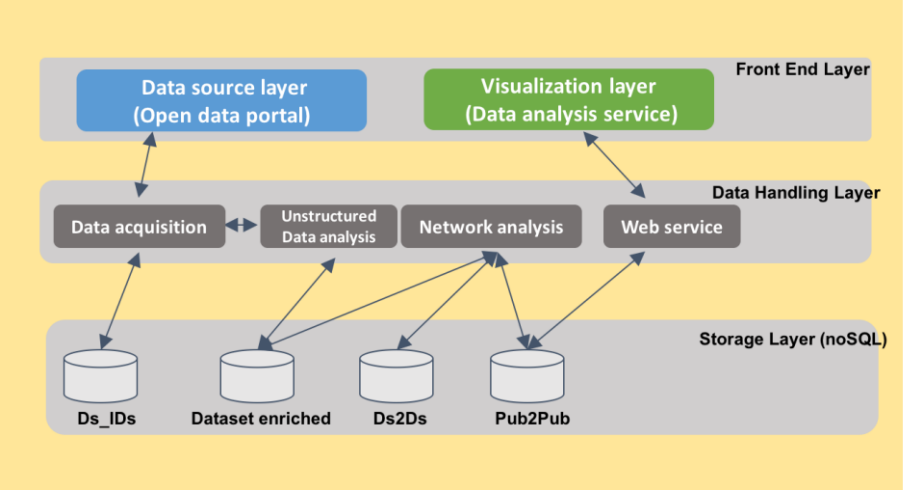

Fig. 1. Open Data Catalogues and Publishers' Semantic Profiling Conceptual Framework.

2 Background and Related Work

Following concepts definitions and a literature review of correlated research areas Open Government Data, NLP and Collaboration Mining:

2.1 Open Government Data

Open Government Data referred to the datasets generated and published by governmental departments "without any restrictions on its usage or distribution" and it doesn't contain any personal or undisclosed data [8]. OGD vary by multiple aspects for example: a) OGD publishing department or agency domain e.g. Agriculture Data, Transport Data, Environmental Data, Financial Data and Telecommunication Data. b) Data format e.g. Excel, Text, PDF, CSV, Theoretically, Government Open Data is operational or administrative governmental data available to use, redistribute, and analyze "in any form without any copyright restrictions" [9]. Regarding the open government working group draft in 2007[3] they generated initial open data principles:

[3] http://public.resource.org/8_principles.html

data must be complete, primary, timely, accessible, machine-processable, nondiscriminatory, nonproprietary, and license-free. Then they generated further open data principles, data must be online and free, permanent, trusted, assumed to be open, documented, safe to open, and designed with public input. Fig. 2 shows the Irish government's open data portal which we used for our experiments[4].

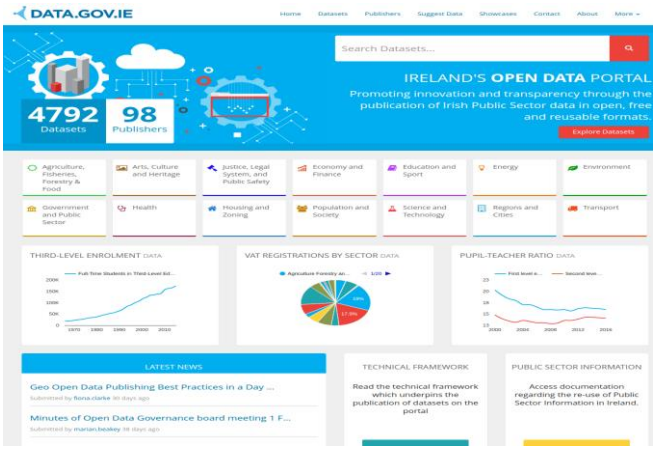

Fig. 2. Irish Government's Open Data Portal.

2.2 Natural Language Processing

Following we discuss the correlated features of Natural Language Processing to our research. Specifically, Named Entity Recognition applications:

2.2.1 Named Entity Recognition

Named Entity Recognition is the process of discovering Named Entities (NE) laying within a given text, a common definition of NE is as follows [10], "an information unit described by the name of a person or an organization, a location, a brand, a product, a numeric expression including time, date, money and percent found in a sentence."[11]. NER applications are implemented using multiple methodologies:

> The Supervised Learning techniques use a big manually categorized dataset. Then this dataset is used for training the recognition algorithm. Supervised Learning techniques apply Conditional Random Fields [12], Hidden Markov Models [13], Decision Trees [14], Support Vector Machines [15] and Maximum Entropy Models [16] The objective of these methods is to identify and categorize related key-words. The unavailability of manually

[4] http://data.gov.ie

categorized datasets and the high cost of generating them, represent a challenging obstacle against Supervised Learning Techniques.

The Semi-Supervised Learning and Unsupervised Learning techniques use either a small categorized dataset for training the algorithm [17], or a clustering based algorithm. Further Unsupervised Learning techniques depend on lingual resources e.g. WordNet, and statistics to solve the NER task as a prediction problem [18].

2.3 Natural Language Processing in E-Government

There are few implementations of NLP technologies in the e-government area. Examples from the works found: A proposed application for gathering crime data from police departments and eyewitness stories and apply NLP technologies with GATE [19]. A system that imitate email answering process automatically or semi-automatically using NLP technologies [20]. Another application presents an original model for incorporating multimedia data to assist e-government tasks [21].

2.4 Mining for Collaboration

In general, due to the great benefits and possibilities of collaboration opportunities mining and discovery research e.g. Process speed enhancing, Standardization and Integration. The detection of possible collaboration opportunities within an organization or across multiple organizations and platforms is targeted in multiple domains. Following the few existing work digging into mining for collaboration area: Mining for collaboration in library domain, the research is harnessing the detection of possible collaboration opportunities with academic professional based on their publications to increase the benefits of students [22]. Collaboration mining between governmental levels and departments based on their objectives, resources and services to increase the government efficiency regarding public policy development and implementation, crisis management, etc. [23]. Collaboration mining tool using agent technology to analyze the collaboration between information on the web to help the tool users to get their desired materials more accurately and faster [24]. Collaboration mining of team members using summaries of successful past projects to increase moderator efficiency to promote project partner's awareness of best way to formulate a proposal for a European research project [25].

3 Semantic Profiling for Collaboration Mining

As shown in Fig. 1 and zoomed in Fig. 3 we have designed a solution pipeline that incorporates Data Acquisition, Data Modeling, Data Analysis, and Data visualization technologies to enable the existence of a collaboration mining tool. We start with inputting the targeted open data portal(s) in which we seek mining for collaborations then we start acquiring metadata (catalogue) of the datasets. Then we restructure the

catalogue to fit into the predesigned storage model (semantic profile), within this model we enhance, filter and exclude less important catalogue features – regarding our use case - and we add new features that are corresponding to our collaboration mining requirements e.g. we add "textual representation" feature by merging original textual features of the data catalogue, we add "Entities" feature to the new catalogue storage model by applying NER over the new "textual representation" feature of the catalogue, we filter features like "author" and "creator" to end up with only "publisher ID" feature, and we exclude "groups" and "tracking summary" features. After constructing and storing the new data model (semantic profile) we start the unstructured data analysis (text mining) pipeline by applying NER algorithm. At the end of that process we generate a comparable feature "Entities" and add it to the new data model to be used for collaboration mining. After that we construct dataset's publisher data model (semantic profiles) which contains aggregated features' values from their published datasets. Finally, we compute relation strengths between dataset publishers based on *comparing* their semantic profiles that we built using the aggregation of unique entities they publish datasets about and store it as shown in Fig. 4 for later visualization and web service usages as shown in Fig. 10 and 11.

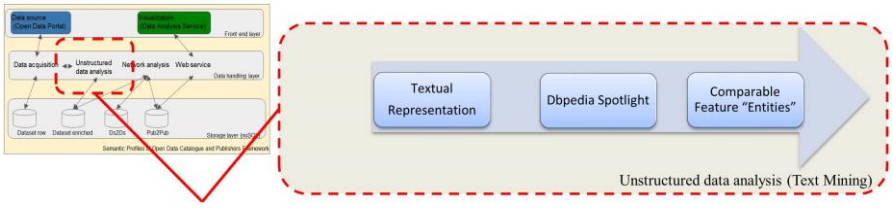

Fig. 3. Unstructured Data Analysis (Text Mining).

Fig. 4. Publisher Collaboration Network.

Following we discuss and represent the results of our Semantic Profiling for Collaboration Mining approach.

3.1 Profiling the Catalogues

By querying the stored enriched metadata of open data portal we are able to generate charts that are profiling the underlying open data catalogue. As an example of those queries we are able to retrieve the named entities detected from mining unstructured textual representations of data catalogues generated by our tool. Those named entities which are originally derivate from dataset metadata are - same as their origin – able to demonstrate a description of the contents of the data portals see Fig. 5 and 6.

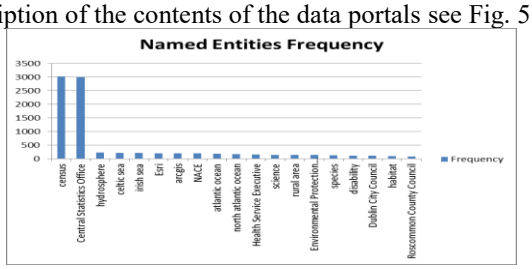

Fig. 5. Top Named Entities Describing the Open Data Portal "data.gov.ie".

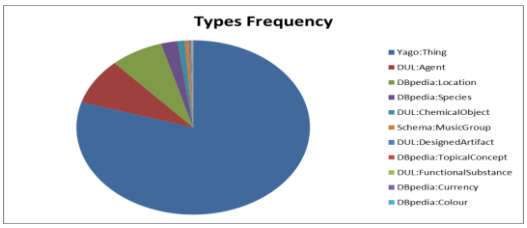

Fig. 6. Top Named Entities Types Describing the Open Data Portal "data.gov.ie".

3.2 Publishers Profiles

Open data publishers are an interesting open data analysis feature; publishers could be governmental departments, councils, etc. which make their profiles a key component of governmental data integration and standardization. An open data publisher's profile is the aggregation of the information extracted from its published dataset metadata. One of the usages of a publisher's profile is to understand more about the domain of the publisher see Fig. 7 for an example.

Open Data Publisher	Named Entities;Frequency
central-statistics-office	Central Statistics Office;2985, census;2863, county;269, NACE;194, Irish;142, rural area;140, disability;97, Irish Travellers;68
department-of-housing-planning-community-and-local-government	department;209, Helena;96, Constitution;38, referendum;35, local authority;35, census;21, county;18, mortgage;15, gener
health-service-executive	Health Service Executive;140, achi;73, icd 10;73, Discharge;62, DRG;62, AM;62, AR;60, Irish;59, ACS;56, birthweight;54, gest
environmental-protection-agency	Environmental Protection Agency;139, REST;102, WMS;74, Water Framework Directive;73, Informatics;59, WFD;43, irish se
marine-institute	hydrosphere;119, atlantic ocean;91, north atlantic ocean;87, celtic sea;85, irish sea;85, CSV;61, biology;56, life science;54,
dublin-city-council	Dublin City Council;105, Dublin City;60, .csv;49, Irish;20, ID;18, Dublin;16, DCC;14, MAP;10, DAT;10, urban planning;9, recyc
roscommon-county-council	KML;85, CSV;85, Roscommon County Council;85, Shapefile;81, Esri;80, Rest API;80, arcgis;80, Roscommon;65, roscommon;
geological-survey-of-ireland	geological survey;58, earth science;55, GSI;53, science;49, lithosphere;43, ireland;38, Irish;32, WMS;29, JSON;28, bedrock;
c2f170ca-63d0-4498-9e81-759827708e97	Ordnance Survey;57, Shapefile;56, CSV;56, KML;56, Rest API;56, Esri;56, arcgis;55, OSi;55, level crossing;3, watercourse;2, S
national-transport-authority	statute;56, Dublin Transport Authority;56, Oireachtas;56, National Transport Authority;56, public transport;55, Peter;54, D
all-island-research-observatory	demography;55, CSV;43, REST;30, JSON;30, SA;22, CSO;19, Service economy;14, population density;13, ED;12, Northern Cro
fingal-county-council	Fingal County Council;53, county;52, .csv;35, Fingal;24, WGS84;20, .zip;5, ID;4, datasheet;3, water quality;2, recycling;2, Du
department-of-education-and-skills	department;50, census;40, county;11, Leaving Certificate;3, .csv;3, ireland;2, national school;2, Junior Certificate;2, xls;2, I
revenue-commissioners	Revenue Commissioners;47, income tax;8, Wicklow;7, Kerry;7, Cavan;7, Monaghan;7, Longford;7, Meath;7, Scheme;7, Kilk
health-research-board	census;42, ICD 10;12, Health Service Executive;9, group sex;3, database;2, disability;2, fibrosis;1, scooters;1, lumbar;1, imn
galway-county-council	Shapefile;42, CSV;42, Galway County Council;42, KML;41, Rest API;36, Esri;36, arcgis;36, county;35, Galway County;19, Galw
national-biodiversity-data-centre	wildlife protection;40, wildlife sanctuary;40, species;40, data centre;40, synecology;40, wildlife conservation;40, wildlife;
sustainable-energy-authority-of-ireland	ireland;38, renewable energies;26, renewable energy;26, Atlas;23, WMS;19, JSON;18, celtic sea;18, irish sea;18, north atla
dun-laoghaire-rathdown-county-council	DLR;18, .csv;7, sbn;6, shp;6, dbf;6, Cycling;5, st;3, ards;3, xls;3, Cherrywood;3, Sandyford;2, WW1;2, A2;2, Dev;2, A1;2, INFO
department-of-communications-climate-action-and-environment	department;36, hydrosphere;35, KML;32, CSV;31, WMS;30, WFS;29, north atlantic ocean;26, atlantic ocean;26, celtic sea;26

Fig. 7. Top Named Entities Describing Data Posted by Top Publishers to the Open Data Portal "data.gov.ie".

3.3 Interlinking Publishers

The resulted publisher profiles are used to mine possible collaboration channels between data publishers at data portal level and among portals level by using the added comparable feature "Entities" see Fig. 8 - 10.

According to our results "marine-institute (129) datasets" and "geological-survey-of-ireland (67) datasets" have the highest relation strength score of (82) which means that they share 82 entities/topics in common. We examined the datasets published by both publishers and we found that for pollution concept/topic there are (7) datasets published by "marine-institute" and (7) dataset published by "geological-survey-of-ireland" and similarly for hydrography concept/topic there are (4) datasets published by "marine-institute" and (18) datasets published by "geological-survey-of-ireland" as shown in Fig. 11 and 12.

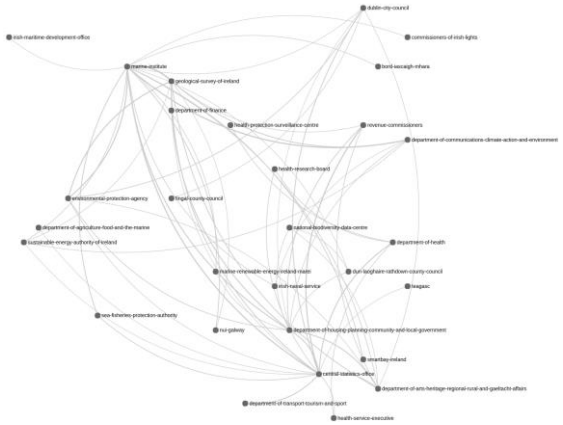

Fig. 8. Publishers Collaboration Network of Open Data Portal "data.gov.ie" with relation strength > 20.

Fig. 9. Publishers Collaboration Network of Open Data Portal "data.gov.ie" – Showing highest relation strength score between "marine-institute" and "geological-survey-of-ireland".

publisherA	publisherB	strength	commons
marine-institute	geological-survey-of-ireland	82	[('DBpedia:Thing', 'pollution'), ('DBpedia:Thing', 'hydrography'), ('DBpedia:Thing', 'habitat'), ('DBpedia:
department-of-communications-climate-action-and-environment	marine-institute	66	[('DBpedia:Location', 'United Nations'), ('DBpedia:Thing', 'hydrography'), ('DBpedia:Thing', 'North Atlar
central-statistics-office	marine-institute	62	[('DBpedia:Thing', 'hake'), ('DBpedia:Species', 'whiting'), ('DBpedia:Thing', 'probability'), ('DBpedia:Loc
central-statistics-office	revenue-commissioners	62	[('DBpedia:Thing', 'Stamp Duty'), ('DBpedia:Thing', 'Exchequer'), ('DBpedia:Thing', 'communication'), ('I
marine-institute	environmental-protection-agency	58	[('DBpedia:Thing', 'aquatic'), ('DBpedia:Thing', 'hydrography'), ('DBpedia:Thing', 'estuary'), ('DBpedia:T
environmental-protection-agency	geological-survey-of-ireland	55	[('DBpedia:Thing', 'topography'), ('DBpedia:Thing', 'hydrography'), ('DBpedia:Thing', 'habitat'), ('DBped
department-of-housing-planning-community-and-local-government	marine-institute	53	[('DBpedia:Thing', 'habitat'), ('DBpedia:Thing', 'Density'), ('DBpedia:Thing', 'Shapefile'), ('DBpedia:Thin
central-statistics-office	department-of-health	51	[('Wikidata:Q12136', 'brucellosis'), ('Wikidata:Q12136', 'influenza'), ('DBpedia:Thing', 'birthweight'), ('W
department-of-housing-planning-community-and-local-government	geological-survey-of-ireland	50	[('DBpedia:Thing', 'habitat'), ('DBpedia:Location', 'Fingal'), ('DBpedia:Thing', 'Shapefile'), ('DBpedia:Thi
department-of-housing-planning-community-and-local-government	central-statistics-office	48	[('Wikidata:Q386724', 'Excel'), ('DBpedia:Location', 'Tipperary'), ('DBpedia:Thing', 'county'), ('DBpedia:T
department-of-housing-planning-community-and-local-government	department-of-arts-heritage-regional-rural-and-gaeltacht-affairs	48	[('DBpedia:Thing', 'wildlife protection'), ('DBpedia:Thing', 'habitat'), ('DBpedia:Thing', 'wildlife conserv
central-statistics-office	department-of-transport-tourism-and-sport	47	[('DUL:Agent', 'iveco'), ('DBpedia:Thing', 'skoda'), ('DBpedia:Thing', 'vauxhall'), ('DUL:Agent', 'piaggio'),
central-statistics-office	dublin-city-council	43	[('DBpedia:Thing', 'tonne'), ('DBpedia:Thing', 'road safety'), ('DBpedia:Thing', 'force'), ('DBpedia:Thing',
central-statistics-office	geological-survey-of-ireland	42	[('DBpedia:Thing', 'copper'), ('DBpedia:Thing', 'aluminium'), ('DBpedia:Thing', 'force'), ('DBpedia:Thing'
department-of-communications-climate-action-and-environment	geological-survey-of-ireland	42	[('DBpedia:Thing', 'geophysical'), ('DBpedia:Thing', 'hydrography'), ('DBpedia:Thing', 'Shapefile'), ('DBp
sea-fisheries-protection-authority	marine-institute	42	[('DBpedia:Thing', 'hake'), ('DBpedia:Species', 'whiting'), ('DBpedia:Thing', 'fish'), ('DBpedia:Location', '
department-of-arts-heritage-regional-rural-and-gaeltacht-affairs	geological-survey-of-ireland	41	[('DBpedia:Thing', 'limestone'), ('DBpedia:Thing', 'Sligo'), ('DBpedia:Thing', 'habitat'), ('DBpedia:Locatig

Fig. 10. Publishers Mined Relations of Open Data Portal "data.gov.ie".

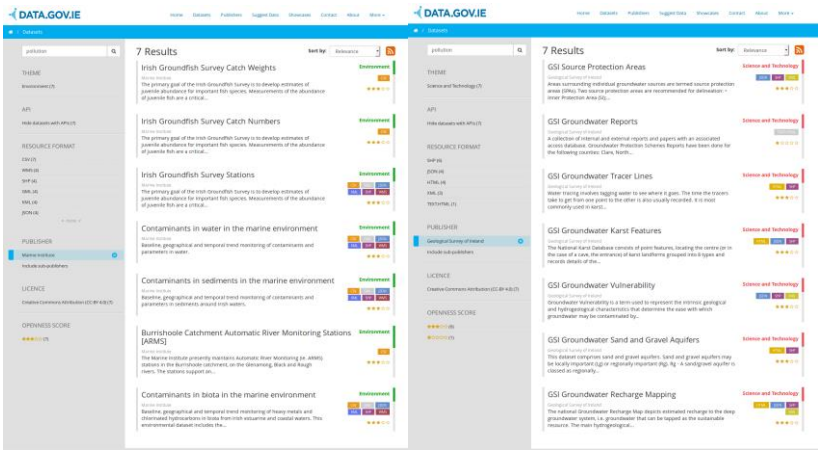

Fig. 11. Datasets shared between Marine Institute and Geological Survey of Ireland around the concept pollution [5] [6]

[5] https://data.gov.ie/data/search?q=pollution&publisher=marine-institute
[6] https://data.gov.ie/data/search?q= pollution &publisher=geological-survey-of-ireland

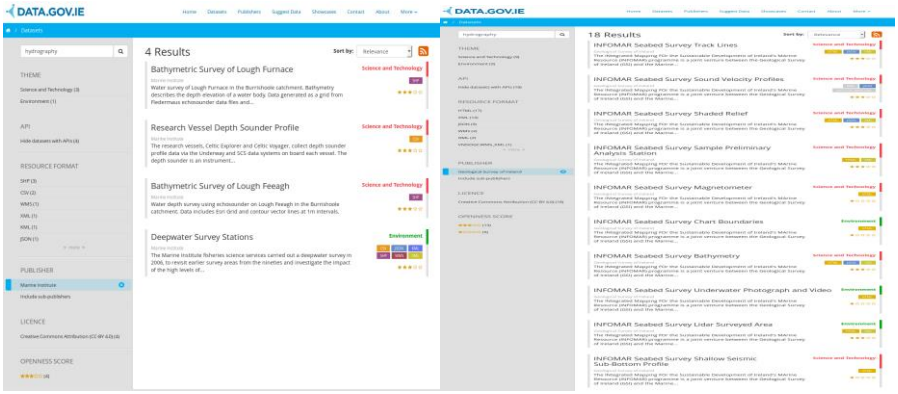

Fig. 12. Datasets shared between Marine Institute and Geological Survey of Ireland around the concept hydrography [7] [8].

3.4 Limitations

Named Entity Recognition area of the work is tightly coupled with the training and the quality of the Named Entity Recognition algorithm. Through this research we have experimented Natural Language Tool Kit (NLTK), Stanford NER and Stanford NER with nGram of (3) enhancement, then we ended up using DBpedia Spotlight as the NE source as through our manual examination of the text analysis phase results DBpedia out performed the other methods in its NE detection quality. DBpedia spotlihght still have its limitations though and we reported one of the issues we faced to their github repository[9].

4 Applications

4.1 Standardization and Collaboration Analysis

Despite most of governments already publishing their data via their open data portals, when a government decides to integrate their data sources over its variant departments and councils, this heterogeneous domain dependent data will consume huge analysis resources and a considerably extended period of time to be fitted into an integrated data repository. Our profiling service will lead the way for data analysts to define integration channels, and necessary concepts standardizations between governmental

[7] https://data.gov.ie/data/search?q=hydrography&publisher=marine-institute
[8] https://data.gov.ie/data/search?q=hydrography&publisher=geological-survey-of-ireland
[9] https://github.com/dbpedia-spotlight/dbpedia-spotlight/issues/407

departments and councils, using the available data published on open data portals. Same example would fit a multinational enterprise as well.

For example "marine-institute" and "geological-survey-of-ireland" share the named entity (pollution), this concept shall be standardized regarding its code and its measurement unit to ease integration and comparability or analysis in general among multiple datasets.

4.2 Intelligent Open Data Portals Exploration

Open data portals are meant to be facing the public in other words the citizens, but citizens can't directly comprehend, and consume this row data [4]. Open data portals profiling service will help citizens to easily and intelligently explore the open data portal using visualized semantic profiles of publishers and datasets.

5 Conclusion and Future Work

Regarding our approach results we believe that we are on the right track to tackle the collaboration mining problem in open governmental data domain, as we are getting interested collaboration recommendations out of our pipeline in a visualized way that is easy to comprehend by general public users of open governmental data.

Our future plan is to overcome the NE limitation by developing a new text analysis pipeline that integrates statistical text analysis, babel.net[10], and DBpedia[11] as our NE source. Also we are planning to replace the string comparison module with semantic relatedness comparison module as the way of calculating relation strength between open governmental data publishers.

Acknowledgments

This paper is partially supported by European Union's Horizon 2020 research and innovation programme under grant agreement No 645860, project ROUTE-TO-PA (Raising Open and User-friendly Transparency-Enabling Technologies for Public Administrations).

References

1. Shadbolt, N., O'Hara, K., Berners-Lee, T., Gibbins, N., Glaser, H., Hall, W., Schraefel, M.C.: Linked open government data: Lessons from data.gov.uk. IEEE Intell. Syst. 27, 16–24 (2012).

[10] http://babelnet.org/
[11] http://wiki.dbpedia.org/

2. Breitman, K., Salas, P., Casanova, M.A., Saraiva, D.: Open Government Data in Brazil. Intell. Syst. IEEE. 27, 45–49 (2012).
3. Mutuku, L.N., Colaco, J.: Increasing Kenyan open data consumption. Proc. 6th Int. Conf. Theory Pract. Electron. Gov. - ICEGOV '12. 18 (2012).
4. Artigas, F., Chun, S.A.: Visual analytics for open government data. 14th Annu. Int. Digit. Gov. Res. Conf. From E-Government to Smart Gov. dg.o 2013. 298–299 (2013).
5. Ribeiro, D.C., Freire, J., Vo, H.T., Silva, C.T.: An Urban Data Profiler. WWW Work. Data Sci. Smart Cities. 1389–1394 (2015).
6. Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., Ives, Z.: Dbpedia: A nucleus for a web of open data. Springer (2007).
7. Mendes, P.N., Jakob, M., García-Silva, A., Bizer, C.: DBpedia spotlight: shedding light on the web of documents. In: Proceedings of the 7th International Conference on Semantic Systems. pp. 1–8 (2011).
8. Janssen, M., Charalabidis, Y., Zuiderwijk, A.: Benefits, Adoption Barriers and Myths of Open Data and Open Government. Inf. Syst. Manag. 29, 258–268 (2012).
9. Kassen, M.: A promising phenomenon of open data : A case study of the Chicago open data project. Gov. Inf. Q. 30, 508–513 (2013).
10. Nadeau, D.: A survey of named entity recognition and classification. Linguist. Investig. 3–26. (2007).
11. Grishman, R.: Message Understanding Conference-6: A Brief History. Proc. COLING. 96, (1996).
12. McCallum, A., Li, W.: Early results for named entity recognition with conditional random fields, feature induction and web-enhanced lexicons. In: Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003 -. pp. 188–191. Association for Computational Linguistics, Morristown, NJ, USA (2003).
13. Bikel, D.M., Miller, S., Schwartz, R., Weischedel, R.: Nymble. In: Proceedings of the fifth conference on Applied natural language processing -. pp. 194–201. Association for Computational Linguistics, Morristown, NJ, USA (1997).
14. Borthwick, A., Sterling, J.: NYU: Description of the MENE named entity system as used in MUC-7. … Conf. (MUC-7. (1998).
15. Asahara, M., Matsumoto, Y.: Japanese Named Entity extraction with redundant morphological analysis. In: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - NAACL '03. pp. 8–15. Association for Computational Linguistics, Morristown, NJ, USA (2003).
16. Hoffart, J., Yosef, M.A., Bordino, I., Fürstenau, H., Pinkal, M., Spaniol, M., Taneva, B., Thater, S., Weikum, G.: Robust disambiguation of named entities in text. 782–792 (2011).
17. Ji, H., Grishman, R.: Data selection in semi-supervised learning for name tagging. 48–55 (2006).
18. Alfonseca, E., Manandhar, S.: An unsupervised method for general named entity recognition and automated concept discovery. … Conf. Gen. …. (2002).
19. Ku, C.H., Iriberri, A., Leroy, G., Ph, D.: Natural Language Processing and e-Government : Crime Information Extraction from Heterogeneous Data Sources. In: The procedings of the 9th Annual International Digital Government Research Conference. pp. 162–170. ACM International Conference Proceedings Series, ACM Press (2006).
20. Dalianis, H., Rosell, M., Sneiders, E.: Clustering E-Mails for the Swedish Social Insurance Agency – What Part of the E-Mail Thread Gives the Best Quality ? 115–120 (2010).

21. Amato, F., Mazzeo, a., Moscato, V., Picariello, a.: Semantic Management of Multimedia Documents for E-Government Activity. 2009 Int. Conf. Complex, Intell. Softw. Intensive Syst. 1193–1198 (2009).
22. Williams, L.M., Cody, S.A., Parnell, J.: Prospecting for new collaborations: Mining syllabi for library service opportunities. J. Acad. Librariansh. 30, 270–275 (2004).
23. Basanya, R., Ojo, A., Janowski, T., Turini, F.: Mining collaboration opportunities to support Joined-Up Government. IFIP Adv. Inf. Commun. Technol. 362 AICT, 359–366 (2011).
24. Wan, L., Chen, J., Gu, D.: An Information Mining Model of Intelligent Collaboration Based on Agent Technology. In: International Conference on Applied Sciences, Engineering and Technology, ICASET 2014. Scientific.net (2014).
25. Palmer, C., Harding, J.A., Swarnkar, R., Das, B.P., Young, R.I.M.: Generating rules from data mining for collaboration moderator services. J. Intell. Manuf. 24, 313–330 (2013).