Classification of intangible social innovation concepts

Nikola Milosevic[1], Abdullah Gok[1], and Goran Nenadic[2]

[1]Manchester Institute for Innovation Research, University of Manchester, Denmark Building, M13 9NG, Manchester, UK
[2]School of Computer Science, University of Manchester, Denmark Building, M13 9PL, Manchester, UK
{nikola.milosevic,abdullah.gok,g.nenadic}@manchester.ac.uk

Abstract. In social sciences, similarly to other fields, there is exponential growth of literature and textual data that people are no more able to cope with in a systematic manner. In many areas there is a need to catalogue knowledge and phenomena in a certain area. However, social science concepts and phenomena are complex and in many cases there is a dispute in the field between conflicting definitions. In this paper we present a method that catalogues a complex and disputed concept of social innovation by applying text mining and machine learning techniques. Recognition of social innovations is performed by decomposing a definitions into several more specific criteria (social objectives, social actor interactions, outputs and innovativeness). For each of these criteria, a machine learning-based classifier is created that checks whether certain text satisfies given criteria. The criteria can be successfully classified with an F1-score of 0.83-0.86. The presented method is flexible, since it allows combining criteria in a later stage in order to build and analyse the definition of choice.

Keywords: text mining, classification, natural language processing, social innovation

1 Introduction

Many social science concepts have abstract definitions, that may be unclear for operational use and for distinguishing entities that satisfy the given definition from those that do not. The physicalist approach suggests to use terms and entities that are more concrete and that are observable. Small concepts are more tangible than the bigger, more abstract ones [7]. Often, it is not possible to use only the tangible concepts, especially for the complex social issues and phenomena that try to deal with these issues.

In this paper we examine automated analysis of one of the high level, intangible social science concepts – social innovation. Social innovation refers to innovative activities, models or services that are supposed to meet certain social need or solve a certain social issue. They are usually executed by the organisations that are primarily social [13, 6]. Social innovation may be performed by both formal and non-formal organisations. They may include technological innovation, but also innovation in inter-actor communication, project management, models, services, and ways to generate output. The major goal of social innovation is to solve societal challenges and improve the lives of the members of the society.

Some of the examples of social innovation projects are:

- Feelif[1] - a product that combines a standard tablet, with an application and a relief grid that allow visually impaired people feel shapes on the tablet. Feelif allows users playing games and use educational content on tablet, addressing social need that visually impaired people were disregarded by smart phone/tablet vendors.
- Real Junk Food project[2] - the chain of restaurants that produces food using ingredients that are getting out of date. Visitors can pay as they feel for the consumed food, while they also employ workers from disadvantaged backgrounds. Real junk food presents organisational innovation.

In the past decade, social innovation has received political recognition. For example, the European Union introduced Social Investment Packages [6] in 2013, that included recommendations for investing in social innovation projects. Since then, the European Union invested a great amount in social innovation projects and organisations. Social innovation projects can apply among others for funds to Horizon 2020, EU Programme for Employment and Social Innovation and European structural and investment funds. Also, certain governments, the European Union and large funding organisations invested in creating databases of social innovation projects and actors [9, 1, 12]. These databases should help funding bodies and policy makers to map projects and determine successful practices and environments for social innovation projects.

However, most of the currently available databases face the following challenges:

1. Thematically focused – The most of the currently available databases are thematically focused to a certain area (e.g. digital social innovation, ageing, homelessness, etc.)
2. Small in size – Majority of the available databases contain between 50 and 1000 projects
3. Limited information – The entries in the databases are limited in terms of features describing a project
4. Relying on a single source – Most of the available databases collect data by human input, either by the project team or the self registration of actors. Also, the majority of the databases do not update data after the project funding have ended.
5. Using different definitions – There are a number of social innovation definitions in literature. Databases use different definition depending of the belief

[1] https://www.feelif.com

[2] http://therealjunkfoodproject.org

of the author's creator or the definition set by the funding body. This makes comparison between databases, database integration and reuse difficult.

The usual way for coding concepts in social sciences is to use human coders that would review and classify the entities. However, human annotation is expensive and its costs can be reduces by almost 80% for certain tasks by utilising natural language processing, text mining and machine learning techniques [11].

In this paper we describe part of the methodology used in European Social Innovation Database (ESID) for classifying social innovation projects, that should address the stated challenges. The database is created in semi-automatic manner, encompassing web crawling, text mining and machine learning. In this paper we focus on the methodology for determining whether a project satisfies social innovation criteria. Social innovation is intangible concept, since people consider social innovation concepts differently and since there are multiple definitions of social innovation. We show that it is possible to create a modular system that takes into the account the underlying concepts of social innovation definitions and allow users to select criteria (and projects) they consider relevant for defining social innovation.

2 Disentangling the definition of social innovations

Discussion about what is social innovation and what are social innovation inclusion criteria is ongoing. Therefore, there are multiple authors proposing definitions of social innovation [2–4, 8, 10, 5]. While nuances between each definition are vastly varying, the broad criteria are about social objectives, social interaction between actors or actor diversity, social outputs and innovativeness. However, different definitions include different combinations and different number of these criteria (e.g. EU is using definition stressing out social objectives and actors interaction). The criteria we used for this work are based on literature review performed by authors. The criteria and their descriptions are presented in Table 1.

By performing exhaustive literature review, it is possible to identify concepts and criteria that compose social innovation definitions. Because there are disagreements between authors on the criteria used for the definition, it is not possible to create a classifier that would satisfy multiple social innovation definition. However, it is possible to create classifiers that can classify underlying concepts and criteria. This allows creation of a modular system in which user can set the definition of social innovation that he would like to use.

3 Methodology overview

Methodology for semi-automated generating of social innovation database consists of two phases. In the first phase we use currently available databases, lists, case study repositories, and mappings of social innovation projects in order to obtain initial data about social innovations. This phase include the following steps:

Element of definition	Criteria description
Objectives	Project primarily or exclusively satisfies (often unmet) societal needs, including the needs of particular social groups; or aims at social value creation. Often no price involved for the main social beneficiary or the innovation is provided to the main beneficiary at cost only. However, there might be examples that price is involved.
Actors and actor interactions	Satisfy one or both of the following: **i. Diversity of Actors:** Project involves actors who would not normally involve in innovation as an economic activity, including formal (e.g. NGOs, public sector organisations etc.) and informal organisations (e.g. grassroots movements, citizen groups, etc.). This involvement might range from full partnership (i.e. project is conducted jointly) to consultation (i.e. there is representation from different actors). **ii. Social Actor Interactions:** Project creates collaborations between "social actors", small and large businesses and public sector in different combinations. These collaborations usually involve (predominantly new types of) social interactions towards achieving common goals such as user/community participation. Often, projects aim at significantly different action and diffusion processes that will result in social progress. Often social innovation projects rely on trust relationships rather than solely mutual-benefit.
Outputs/Outcomes	Project primarily or exclusively creates socially oriented outputs/outcomes. Often these outputs go beyond those created by conventional innovative activity (e.g. products, services, new technologies, patents, and publications), while conventional outputs/outcomes might also be present. These outputs/outcomes are often intangible and they might include the following but not limited to: - change in the attitudes, behaviours and perceptions of the actors involved and/or beneficiaries - social technologies (i.e. new configurations of social practices, including new routines, ways of doing things, laws, rules or norms) - long-term institutional/cultural change
Innovativeness	There should be a form of "implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organisational method". The project needs to include some form of innovative activities (i.e. scientific, technological, organisational, financial, and commercial steps intending to lead to the implementation of the innovation in question). Innovation can be technological (involving the use of or creating technologies) as well as non-technological. The innovation should be at least "new" to the beneficiaries it targets (it does not have to be new to the world).

Table 1. Description of the social innovation criteria used in this study

1. Compose a list of social innovation sources.

2. Crawl the project description pages from the listed sources
3. Crawl the project websites, if they were available in the social innovation source
4. Annotate a set of projects. The projects are annotated whether they satisfy social innovation criteria by human coders
5. Create a machine learning model for classifying projects whether they satisfy social innovation criteria
6. Obtain additional features about the project, such as information about organisations involved, location, etc.

The last step of the phase is out of the scope of this project. The second phase of the project involves crawling of the sources that potentially could contain social innovation projects, such as crowd-sourcing platforms. Also, the second phase is also out of the scope of this paper.

4 Obtaining data about social innovation from existing sources

Firstly, we have identified data sources containing information about social innovation projects and actors. The list contained 93 information sources, however, some of the data sources contained cleaner data than the others. For the clean and relevant data sources, we developed a set of web crawlers that obtained data and stored it into our database. A web crawler (also known as scraper or spider) is a program or automated script which browses the World Wide Web in a methodical, automated manner and collects the content of the visited web pages (in full or targeted parts of them). The data sources could not be directly downloaded, however, they had database accessible on the web, and therefore web crawlers could methodically visit all entity pages and obtain data about them. The development of crawlers can be time-consuming. We started with developing crawlers for the biggest databases. At the moment, we have developed crawlers for 9 databases, however, they contain over 6,000 entities. This presents more than 85% of all entities in the identified data sources.

However, while crawling, we faced a number of challenges. The main challenge for the crawling is that the web sources did not have consistent structures, and wealth of information. Our approach to crawl these data sources was to obtain targeted information that was included in data source about certain entity (project or actor). In order to achieve this, we needed to develop separate crawlers for each data source that is able to locate information of interest on the page.

Certain data sources contained information about both projects and actors (e.g. Digital Social Innovation), however, some data sources contained information only about one entity type (projects – e.g. EUSIC, MOPACT, actors – Social Enterprise UK, Social Innovation Generation).

The crawled data sources and the number of their entities are presented in the Table 2.

Data Source	Number of Projects	Number of Actors
Digital Social Innovation	2,200	2,007
European Social Innovation Competition	90	0
MoPAct	140	0
Innovage	153	0
SIMRA	9	28
European Investment bank social innovation tournament	72	0
Social Innovation Generation (Social Innovation in Canada database)	0	256
Bill and Melinda Gates Foundation	0	444
Social Enterprise UK	0	687
Total	2,664	3,422

Table 2. Description of the social innovation criteria used in this study

5 Data annotation

In order to make a data set for supervised machine learning-based approach that is able to classify social innovation criteria, we organised two data annotation workshops. During the first workshop, 6 annotators were annotating about 40 projects each. Annotators were PhD students and research staff whose research is associated with the are of innovation and social innovation. The text for each project was composed from the texts available on the project websites. For this annotation task, we included only projects whose websites contained between 500-10,000 words. About 20% of the documents were annotated by at least two annotators. For annotation we used Brat rapid annotation tool [14]. The annotators were asked to annotate sentences that present how certain project met defined social innovation criteria (objectives, actor interaction, outputs, innovativeness) and to give a score at the document level for each of the four criteria (as presented in Table 2). The document level marks were in range of 0-2:

- 0 – criteria not satisfied
- 1 – criteria partially satisfied
- 2 – criteria fully satisfied

In the second annotation workshop, we focused more on calculating inter-annotator agreement and finding potential outlier annotator (annotator with high disagreement compared to other annotators). We created a dataset using projects listed as semi-finalists and finalists in EU Social Innovation Competition[3] and European Investment Bank Social Innovation Tournament[4]. The dataset consisted of 40 projects, whose websites were crawled. A subset of the four annotators annotated the whole dataset in the same manner as during the first workshop.

[3] http://eusic.challenges.org/

[4] https://institute.eib.org/whatwedo/social-2/social-innovation-tournament-2/

6 Classification of criteria

The dataset created during both annotation task were used for training and validation of the machine learning-based approach. The classifier is created for each social innovation criteria (objective, actors, outputs, innovativeness). Before applying machine learning algorithm the text was stemmed and stop-words were removed (using Rainbow stop-word list[5]). We have evaluated classification using Naive Bayes machine learning algorithms using the bag-of-words language model. Since dataset was not balanced, having more negative instances than positive, we also performed an experiment with balancing data by oversampling positive instances.

7 Results

Firstly, we present the results of annotation workshop, including inter-annotation agreement and the number of non-relevant projects in the examined data sources. Then we present the results of machine learning classification using the two described approaches.

7.1 Data annotation results

During the first workshop, six annotators annotated 40 documents each on the sentence and document level. During the second workshop, three annotators annotated same 43 documents, while one annotator annotated 30 of these documents. During the workshops about 10% of the data available at the moment of annotation workshops was annotated. Inter-annotator agreement per each criteria is presented in the Table 3. Inter-annotator agreement is calculated on the paragraph level and document level. Inter-annotator agreement on the paragraph level is calculated by examining each paragraph of the text whether it contains annotation of a certain class in both annotated documents. Inter-annotator agreement on the document level is calculated by examining whether both annotators scored the document with the certain annotation type (annotation types translates to four social innovation criteria). Since agreement on score was also fairly low, we have binarised document level annotation, so they score whether the criteria is satisfied (scores 1 and 2) or not satisfied (score 0).

As it could be seen, inter-annotator agreement on the paragraph level is low and therefore these annotations are not useful for machine learning. Sentences are shorter structures than paragraphs and therefore agreement would be even lower. In the paragraph level inter-annotator agreement we looked whether a certain paragraph is annotated by both annotators with the same annotation type. The inter-annotator agreement on the document level is in range of 65%-76%, which is relatively low as well, indicating that the annotation concepts are intangible and that people generally do not completely agree on what is innovative, what

[5] http://www.cs.cmu.edu/ mccallum/bow/rainbow/

Inclusion criteria	Paragraph level agreement	Document level agreement
Objectives	37.50%	76.60%
Actors and Actor interactions	17.20%	65.70%
Outputs	18.90%	66.73%
Innovativeness	19.50%	70.80%
Macro-average	23.27%	69.96%

Table 3. Inter-annotator agreement on paragraph and document level per criteria in the annotated dataset

social objectives are or what social actor interactions are. However, this score is high enough to be used for machine learning.

We also calculated how many projects in each examined data source were false positives (projects not satisfying any social innovation criteria – spam projects). Agreement for detecting social innovation or false positive project is about 85%. The number and percentage of false positive (spam) projects per data source can be seen in Table 4.

Data source	Number of projects	Percentage of false positive projects (false positive/total annotated)
European Social Innovation Competition	90	12.9% (8/62)
MoPAct	140	7.3% (3/41)
Innovage	153	30% (6/20)
Digital Social Innovation	2,200	58% (105/188)
European Investment bank social innovation tournament	72	6.8% (6/87)
SIMRA	9	0% (0/2)

Table 4. Percentage of false positive (spam) projects per data source in the annotated data

The range of false positive projects in data sources ranges between 0% and 58%. As we hypothesised, the data sources are not clean and user-imputed data sources, such as Digital Social Innovation are noisy. Even some expert imputed data sources, such as Innovage contains about 30% of false positive projects.

7.2 Classification results

The final dataset, that is used for machine learning, after both human annotation workshops, contains 277 documents, with the following composition per criteria:

– Objectives – 166 negative, 111 positive instances
– Actors – 189 negative, 88 positive instances
– Outputs – 190 negatives, 87 positive
– Innovativeness – 190 negative, 87 positive

This data was used for training and testing machine learning-based method for determining whether project description satisfies given social innovation criteria. For evaluation was used 10-fold cross-validation. The results of four different classifiers for each criteria are presented in Table 5.

Criteria	TP	FP	FN	Precision	Recall	F1-score
Actors	62	39	26	0.614	0.705	0.656
Objectives	81	36	30	0.692	0.730	0.711
Outputs	61	41	26	0.592	0.701	0.642
Innovativeness	58	42	29	0.580	0.667	0.620

Table 5. Results of classifying criteria using Naive Bayes classifier. The input text was created from the content of projects' websites. The evaluation is performed using 10-fold cross-validation

The results are in range 0.62-0.71 F1-score, having significant amount of false positives and false negatives. As it was previously mentioned, the data set that was used was imbalanced, having higher amount of negative instances than positive. In cases of actors, outputs and objectives, the data set contained more than two times more negative instances compared to the positive ones. The imbalance of data set can affect the performance of classification, especially in case of Bayesian-based method for classification. Therefore, we applied oversampling of positive instances. With oversampled positive instances, the data set contained similar amount of positive and negative instances for each criteria. The classification results over oversampled data is presented in Table 6.

Criteria	TP	FP	FN	Precision	Recall	F1-score
Actors	118	25	15	0.825	0.887	0.855
Objectives	121	25	14	0.829	0.896	0.861
Outputs	122	30	11	0.803	0.917	0.856
Innovativeness	117	29	16	0.801	0.880	0.839

Table 6. Results of classifying criteria using Naive Bayes classifier after balancing data set by oversampling positive class. The input text was created from the content of projects' websites. The evaluation is performed using 10-fold cross-validation

The performance of classifiers across criteria increased after oversampling positive instances. Classifiers with these performance can be applied in production system in order to determine whether descriptions from projects' websites satisfy social innovation criteria and to determine which projects are social innovation and which are not.

Experiments using other algorithms, such as SVM, decision trees and neural networks using Glove embeddings were also conducted. However, Naive Bayes over-performed these approaches. This is likely due to the small data set and the fact that Naive Bayes is able to generalise well even with relatively small data sets.

8 Conclusion

In this paper we presented a methodology to model intangible social science concepts over which definition may be active debate in the field. The case of intangible concept presented in this paper is social innovation. There are numerous definitions describing the concept. Also, using multiple definitions would make data inseparable, therefore it would not be possible to make automated systems for collecting, cataloguing and classifying such items.

The approach we proposed relies on extensive literature review of the concept and disentangling the definitions into the components that they were made of. These components are usually more tangible concepts. Such concepts would have better agreement between coders and will be easier separable for machine learning-based methods. Each of the many definitions of the concept will be made of a certain combination of the identified components. Therefore, it is possible to create a system that is modular and that based on the combinations of components is able to return instances that satisfy particular definition of the concept. In the case of social innovation, the identified components were objectives, actors, outputs and innovativeness. Once certain text is classified whether it satisfies given components, one can retrieve instances satisfying EU definition of social innovation that consists of objectives, actors and innovativeness. Also, instances satisfying other combination of criteria or social innovation definitions can be retrieved.

The performance of classifiers for the social innovation criteria was encouraging, ranging between 0.83-0.86 F1-score. It proves the hypothesis that supervised machine learning can be used to classify whether a project description satisfy certain social science criteria. Also, the percentage of correctly classified instances in the given data set was higher than calculated inter-annotator agreement. However, the percentage of correctly classified instances is agreement with the final data set only, which may be possible due to larger amount of training data and proves that classification results are comparable with human annotators. After project descriptions are classified whether they satisfy given criteria, it is possible to extract other meta-data from text, using named entity recognition tools and information extraction strategies.

The approach can be generalised to cataloguing other social science concepts by decomposing multiple definitions into the component criteria. Different kinds of projects, regulations, laws, or organisations can be classified in this manner. The catalogues of such entities may be useful for further research, collaboration, policy making and governance.

In the future, we are planning to include additional data sources, such as projects from crowd-funding platforms and community building portals (such as Meetup.com) and use the classification model to discover new projects satisfying out social innovation criteria. In addition, we are planning to expand training data set by performing additional annotation workshops. We hope that more annotation will enable other methods to be used and reduce certain domain related biases that may exist with the current model.

Acknowledgements

The work presented in this paper is a part of KNOWMAK project that has received funding from the European Union's Horizon 2020 research and innovation programme under grand agreement No. 726992.

References

1. Bria, F., SESTINI, F., GASCO, M., BAECK, P., HALPIN, H., ALMIRALL, E., KRESIN, F.: Growing a digital social innovation ecosystem for europe: Dsi final report. Brussels: European Commission (2015)
2. Caulier-Grice, J., Davies, A., Patrick, R., Norman, W.: Defining social innovation. A deliverable of the project:The theoretical, empirical and policy foundations for building social innovation in Europe(TEPSIE), European Commission–7th Framework Programme, Brussels: European Commission, DG Research (2012)
3. Choi, N., Majumdar, S.: Social innovation: towards a conceptualisation. In: Technology and innovation for social change, pp. 7–34. Springer (2015)
4. Dawson, P., Daniel, L.: Understanding social innovation: a provisional framework. International Journal of Technology Management 51(1), 9–21 (2010)
5. Edwards-Schachter, M., Wallace, M.L.: shaken, but not stirred: Sixty years of defining social innovation. Technological Forecasting and Social Change 119, 64–79 (2017)
6. European Union: Social innovation (2017), `http://ec.europa.eu/social/main.jsp?catId=1022`
7. Gerring, J.: Social science methodology: A criterial framework. Cambridge University Press (2001)
8. Grimm, R., Fox, C., Baines, S., Albertson, K.: Social innovation, an answer to contemporary societal challenges? locating the concept in theory and practice. Innovation: The European Journal of Social Science Research 26(4), 436–455 (2013)
9. Haak, M., Slaug, B., Oswald, F., Schmidt, S.M., Rimland, J.M., Tomsone, S., Ladö, T., Svensson, T., Iwarsson, S.: Cross-national user priorities for housing provision and accessibilityfindings from the european innovage project. International journal of environmental research and public health 12(3), 2670–2686 (2015)
10. van der Have, R.P., Rubalcaba, L.: Social innovation research: An emerging area of innovation studies? Research Policy 45(9), 1923–1935 (2016)
11. Hillard, D., Purpura, S., Wilkerson, J.: Computer-assisted topic classification for mixed-methods social science research. Journal of Information Technology & Politics 4(4), 31–46 (2008)
12. Luijben, A., Galenkamp, H., Deeg, D.: Mobilising the potential of active ageing in europe (2012)
13. Mulgan, G.: The process of social innovation. innovations 1(2), 145–162 (2006)
14. Stenetorp, P., Pyysalo, S., Topić, G., Ohta, T., Ananiadou, S., Tsujii, J.: Brat: a web-based tool for nlp-assisted text annotation. In: Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for

Computational Linguistics. pp. 102–107. Association for Computational Linguistics (2012)