2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)
Banff Center, Banff, Canada, October 5-8, 2017

Inventorum: A Platform for Open Innovation

Jarosław Protasiewicz
National Information Processing Institute, Warsaw, Poland

***Abstract*—The purpose of this industrial brief is to introduce an information system named Inventorum, which implements the idea of open innovation. Its aim is to increase cooperation between business and academia and in this way boost economic competitiveness. The system recommends innovations, projects, experts, partners, and conferences to the users based on their profiles. Information is served in three ways, namely as (i) fast recommendations, (ii) full recommendations, and (iii) search responses. This information platform has been developed by National Information Processing Institute in Poland. Any institution, company, or independent expert is welcome to join the platform, which is available on the Internet for free use (see http://inventorum.opi.org.pl/en). The paper briefly presents functions, advances, and some technical details of the proposed information system.**

***Index Terms*—information system, open innovation, recommendation, big data**

I. Introduction

Nowadays, almost each business, even local one, has to compete in the global market with other small businesses or huge corporations. The information era enabled us easy access to vast information regarding goods, companies, experts, and the most recent research achievements, which seems to be the most important regarding innovativeness issues. Since innovations help companies to be competitive on the market, cooperation between businesses and academia is perceived as the essential aspect of a business strategy. It could be said that everybody can become an innovator [4], [5].

Two issues, i.e. information influx and innovativeness need, motivated us to propose an information system, which is aimed to implement an open innovation paradigm [1]. The system realises it by recommending innovations, projects, experts, partners, and conferences to research institutions or companies based on user profiles. For example, a company may get cooperation propositions with other companies or research units, find innovations that could boost its competitiveness, or receive offers to take part in reach and development projects. Such profiled recommendations could reduce time need to reach desirable information and concurrently increase its accuracy [2], [3].

This work briefly presents functions, advances, and some technical details of the proposed information system named Inventorum. The aim of the system is to increase cooperation between business and academia and in this way boost economic competitiveness. The web platform has been developed in 2016 by National Information Processing Institute, Poland. Any institution, company, or independent expert is welcome to join the platform, which is available on the Internet for free use (http://inventorum.opi.org.pl/en).

It has to be underlined that this work is an industrial brief rather than a research study. The details regarding scientific assumptions of the system are included in the simultaneous work [2]. Especially, it explains how the system recommends information.

The remainder of this industrial brief is as follows. Section II includes basic information about the system. Next, Section III covers the description of tools provided by the system. Then, Section IV presents some technical details regarding technology, architecture, and usage statistics. Finally, conclusions and literature are provided. In addition, there is an Appendix comprising screenshots from Inventorum.

II. System overview

This section briefly explains how the system provides information to users. There are also illustrative examples.

A. Messaging channels

Inventorum is a place that connects scientists with the industry and the business using intelligent algorithms, as well as ensuring access to comprehensive, consistent and reliable information on innovations, projects, innovative enterprises and scientific institutions, as well as experts and conferences. Information is promoted in three ways, namely as (i) fast recommendations, (ii) full recommendations, and (iii) search responses (Figure 1). Each messaging channel plays a different role in transferring information between users. Their purposes are explained by examples in the next subsection.

B. Information flow in examples

Let us see how the messaging channels may work in reality. For this purpose, a system user is compared with a customer in a shop.

Fast recommendations: Just after the user has authenticated, the system prompts the news on user's desktop, i.e. the most recent and personalised information, which is recommended to the user. It has to be underlined that the recommendations are personalised for the user, and they are available only for this particular person. Such fast recommendations are somehow similar to a situation when the customer visiting the shop sees advertisements on this shop window. For instance, they are displayed on a screen, when its sensors recognise the customer.

978-1-5386-1645-1/17/$31.00 ©2017 IEEE

Full recommendations: If the user was interested in the news, he or she would probably request for more offers. In such case, the full recommendations are presented. It has to be noted that they are personalised regarding the user needs and interests depicted in the personal profile and the profiles of related institutions. It resembles a hypothetical case when the customer receives a unique book containing offers directed only to him or her, after passing through the shop door. The book includes offers comprising the whole selection of available goods at the shop.

Search responses: The user may need more information or can be unsatisfied with the recommendations. In such case, he or she can manually search the system database. Since the platform is equipped with the full-text search engine, it understands the queries posed almost in natural language. As a result, the individual receives a list of information that matches the search criteria. This is similar to a regular shopping at the shop when the customer looks for products without any guidance.

Figure 1. The overview how information is served to the users by the system.

III. System features

This section covers the brief overview of all functions provided by the system, i.e. recommendation mechanisms, search tools, communication among the platform and its surrounding, users authentication, and information authorization.

A. Authentication and the main desktop

Since it is hard to remember each unique login and password required by every information system, the platform allows authentication by using credentials of the most popular social media, i.e. Facebook, Google, and LinkedIn. It is the convenient and safe way of authentication. Of course, a user can create an account dedicated only to this system, which later may be associated with the social media.

After authenticating, the user sees a desktop at first, which a screenshot is shown in Figure 3 in the Appendix. As user's main operations centre, it is composed of the information boxes presenting the most vital information, namely areas of (i) news, (ii) observed objects, (iii) companies, and two toolbars including the most crucial functions, i.e. (iv) searching and (v) communication tools. The news box presents the most adequate and recent information, which are recommended to the user by the fast recommendation mechanism. The box of observed objects covers information that the user marked as valuable to follow. Similarly to the news widget, it shows only recently marked information objects. The third information box presents companies or research institution associated with the user. In the case of the lack of such connections, the user is encouraged to add some. This rule applies to any information box in the system. The upper desktop area includes tools allowing search information by querying in user's natural language. For clarity, requests and displayed information objects are divided into groups, i.e. innovations including legal protections, companies, scientific institutions, experts, and conferences. In addition, the upper widget covers communication tools, which display messages coming from the system or other users.

B. Recommendations

The recommendations provided on user's desktop may be too narrow due to limited space on the news box. The user may display all recommended information by using a function *Show full recommendation* available on the desktop (see a screenshot in Figure 3 in the Appendix). As a result, the system will display all information proposed as the most relevant to the user needs (see a screenshot in Figure 4 in the Appendix). On the right side, there are tools for defining the user profile and his or her preferences. More specifically, it is possible to limit displayed information according to its categories, i.e. innovations, projects, companies, research institutions, experts, and conferences. It may be convenient for someone interested only in some categories. In addition, the user can limit data displayed according to time their validity, e.g. it could be the last month, three months, six months, or without any limits. The recommendations are generated based on data contained in the user profile and the profiles of his or her companies and research institutions. Thus, on the right side, there are also tools for selecting user organisations and editing user's profile by providing personal data like *about me, keywords, documents,* etc.

C. Search and communication tools

If the recommendations are insufficient to the user needs, it is possible to query information by the searching tools localised on the top of each screen (see a screenshot in Figure 6 in the Appendix). What is convenient for users, they can ask the system by using their natural language, e.g. English. The search algorithm uses full-text indexes to look up for information in respect of each its category. If the returned results are too broad due to language ambiguity, they can be precisely narrowed by the specific search fields on the right.

The system informs each user by putting information about events. The events must be somehow related to the user. On the other hand, the platform participants may send invitations to each other, e.g. a call for joining a project or innovation. In this way, information flows among the users as well as between the system and each user.

D. Profiles

It has been mentioned that information is grouped into several categories, i.e. innovations, projects, companies, research units, experts, conferences. Each information type has its unique profile, which is understood as the set of fields and a presentation layout.

Figure 7 in the Appendix contains an example of the expert profile. Apart from common information like personal data, keywords, publications, and documents describing achievements, it includes relations the expert with innovations, projects, companies and institutions, conferences, etc. These data are utilized by the recommendation engine to propose experts for a particular task, e.g. a problem posed by a company.

On the other hand, Figure 8 in the Appendix depicts an example of a research unit that employ the expert discussed above. It can be assumed that the company profile looks almost the same. Similarly to the expert profile, these forms include a name, activity description, keywords, contact data, etc., as well as relations with all possible information objects. These data help to match companies or research units with innovations, projects, experts, partners, and so on. In addition, the profiles may include information specifying what these entities look for in the market, which allows the recommendation engine to form more precise suggestions.

Innovation, projects, and conferences are information objects that connect companies, research institutions, and experts. Figure 5 in the Appendix presents an example of the innovation profile. The project and conference profiles are almost similar to the innovation profile; however, they may have some specific data. Overall, they include descriptive data in many languages such as keywords, description, owners, some dates, and alike. In addition, the innovation profile may cover its legal protections. It is also possible to scrutinise similar innovations to the currently displayed.

IV. Technical details

This section includes some technical details regarding technology and architecture of the system and selected usage statistics.

A. Architecture

The system has exemplary multi-layered architecture. Nevertheless, it contains some advanced non-classical solutions. It covers four layers, namely (i) a data storage, (ii) a data access layer, (iii) a computation layer, and (iv) a user access interface. Figure 2 briefly depicts the overall architecture of the system.

The data storage and data access layers include two mechanisms of information persistence and its serving to the higher tiers. All metadata describing objects, e.g. innovations, companies, and alike, are persisted in a relational database. On the other hand, a distributed file system keeps the objects' contents. The data access layer has to synchronise requests for data in both storages if it is necessary. Seeing that the users usually want to operate on objects metadata, the system can quickly answer their requests because the database is not overloaded by entire textual contents. On the contrary, the distributed file system efficiently stores textual information and optimises it for further processing. The object contents may be swiftly served to users as well as to computational processes, e.g. the recommendation engine.

The computation layer contains the core processing components. The algorithms are mainly responsible data crawling, extraction, transformation to information, e.g. domains classification, entities recognition, and identification. The most crucial element is the recommendation engine, which processes all information, generates full-text indexes, and produces suggestions of information that correspond best to users' needs. A business processes component implements a typical model-view-controller model for handling user requests. Moreover, it synchronises cooperation of the business processes with the recommendation tools and other algorithms regarding the integration of responses to users' requests.

The user access interface covers web interfaces, which were designed having in mind that they must be responsive, i.e. adapt a current layout to users screen size.

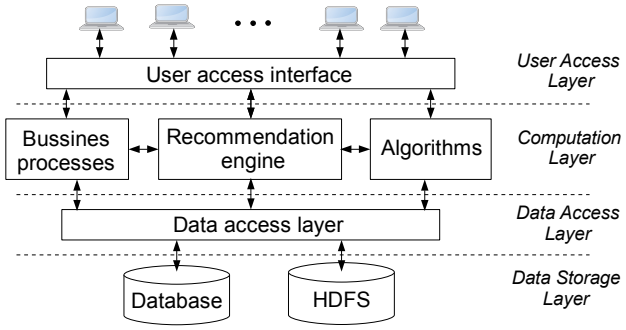

Figure 2. An overall architecture of the system (HDFS stands for Hadoop Distributed File System).

B. Technology

Java Enterprise Edition is the leading technology used to develop almost all layers of the web platform. It is utilised for the construction of the web interfaces, business components, security rules, and data access layer. In addition, it partially occurs in the components implementing processing algorithms and recommendations. Nevertheless, the recommendation engine utilises the Apache Lucene full-text engine; whereas, the algorithms are developed by using of Apache Spark and Apache Hadoop environments, i.e. the open source distribution provided by Cloudera. Textual data are stored and processed in the Hadoop Distributed File System. On the contrary, relational data are preserved in the relational database.

C. Usage statistics

The platform was deployed in December 2015. The tables included below cover usage statistics collected since that time.

Acquiring information regarding innovations, projects, conferences, and research institutions was rather easy because it

originates from our local or publicly open databases, e.g. European Patent Office. As we can notice by analysing quantities in respect of information types in Table I, a more challenging task was to convince companies and users that the offered platform might be a beneficial tool for them.

Although the platform has only about 6 thousand users (Table I), they rather actively use it taking into account that the system is not entertainment media but something presumably useful for a narrow scope of their work. On average, each user has nine sessions each month. During each session, he or she sends about 60 requests to the system (Table II). In a certain approximation, a request may be considered as a user click.

In addition, Table III depicts an average activity of the users in respect of their country origin. It is evident that the system is rather popular only in Poland. The users localised in other countries generate merely nine percent of all traffic to the system. The reason for such results might be the lack of a worldwide promotional campaign and relative immaturity of the system.

Table I
THE VOLUMES OF INFORMATION IN RESPECT OF ITS CATEGORIES AS OF JULY 24, 2017. THE USER QUANTITY INCLUDES 606 EXPERTS AND 1,141 REVIEWERS.

Information category	Quantity
Companies	369
Research Institutions	2,457
Innovations	2,991,722
Projects	21,969
Conferences	18,802
Users	6,029

Table II
AN AVERAGE DAILY AND MONTHLY ACTIVITY OF THE USERS IN 2016.

Average number of	Overall		Per registered user	
	Daily	Monthly	Daily	Monthly
users' requests	9,498	288,890	1.91	58.13
users' sessions	1,486	45,185	0.30	9.09

Table III
AN AVERAGE ACTIVITY OF THE USERS IN RESPECT OF A COUNTRY IN 2016.

No	Country	Requests	
		Number	Percent
1	Poland	3,164,866	91.29
2	United States	128,663	3.71
3	Ukraine	39,474	1.14
4	Germany	25,005	0.72
5	Great Britain	12,340	0.36
6	France	5,955	0.17
7	Russian Federation	5,550	0.16
8	Latvia	4,244	0.12
9	Italy	4,129	0.12
10	Netherlands	2,719	0.08
Others		84,830	2.13
Total		3,466,669	

V. CONCLUSIONS

In this industrial brief, the new information platform called Inventorum has been presented. It has to be noted that this work is aimed to report the product features. Thus, scientific issues are excluded from this paper.

The system tries to implement the idea of open innovation, which is manifested in three aspects, namely (i) open access, (ii) information variety, and (iii) pressure on cooperation supported indirectly by recommendations.

The platform is open for everyone who wants to share or acquire knowledge. It is available for free use on the Internet (see https://inventorum.opi.org.pl/en). Its interface is in English and Polish; however, the descriptive data may be in any language. It has to be noted that legal protections of innovations are usually described in English, German, and French.

Information in the system is grouped into several categories, i.e. innovations including legal protections, projects, experts, research institutions or companies, and conferences. Such variety of information may give a guarantee that every user could find something worth to interest. On the other hand, standardisation of information objects in profiles helps organise information flow among the users.

The most crucial for innovativeness is cooperation among business and academia. Therefore, the system focuses on providing information about possible partners, innovative products, and valuable research achievements. Since people are overwhelmed by data influx, the recommendation engine tries to propose the most relevant data to users' needs. It is realised by (i) personalised news, (ii) full recommendations, and (iii) search reports.

Although the system is complete and ready to use, it requires additional development in various areas. Firstly, it is essential to encourage as many as possible experts, companies, and scientific units over the world to join the platform. In addition, the recommendation engine may need improvement especially in dealing with multilingual textual data.

REFERENCES

[1] M. Curley and B. Salmelin, "Open innovation 2.0: a new paradigm," *OISPG White Paper*, pp. 1–12, 2013.

[2] J. Protasiewicz, "Inventorum - a recommendation system connecting business and academia," in *Systems, Man, and Cybernetics, 2017 IEEE International Conference on,*. IEEE, 2017, p. manuscript accepted for publication.

[3] J. Protasiewicz, W. Pedrycz, M. Kozłowski, S. Dadas, T. Stanisławek, A. Kopacz, and M. Gałężewska, "A recommender system of reviewers and experts in reviewing problems," *Knowledge-Based Systems*, vol. 106, pp. 164–178, 2016.

[4] A. Savoia and P. Copeland, "Entrepreneurial innovation at google," *Computer*, vol. 44, no. 4, pp. 56–61, 2011.

[5] M. Wisnioski, "The birth of innovation," *IEEE Spectrum*, vol. 52, no. 2, pp. 40–61, 2015.

APPENDIX

This Appendix comprises screenshots from Inventorum, namely an example of the system desktop, full recommendations screen, searching tools, innovation profile, expert profile, and institution profile.

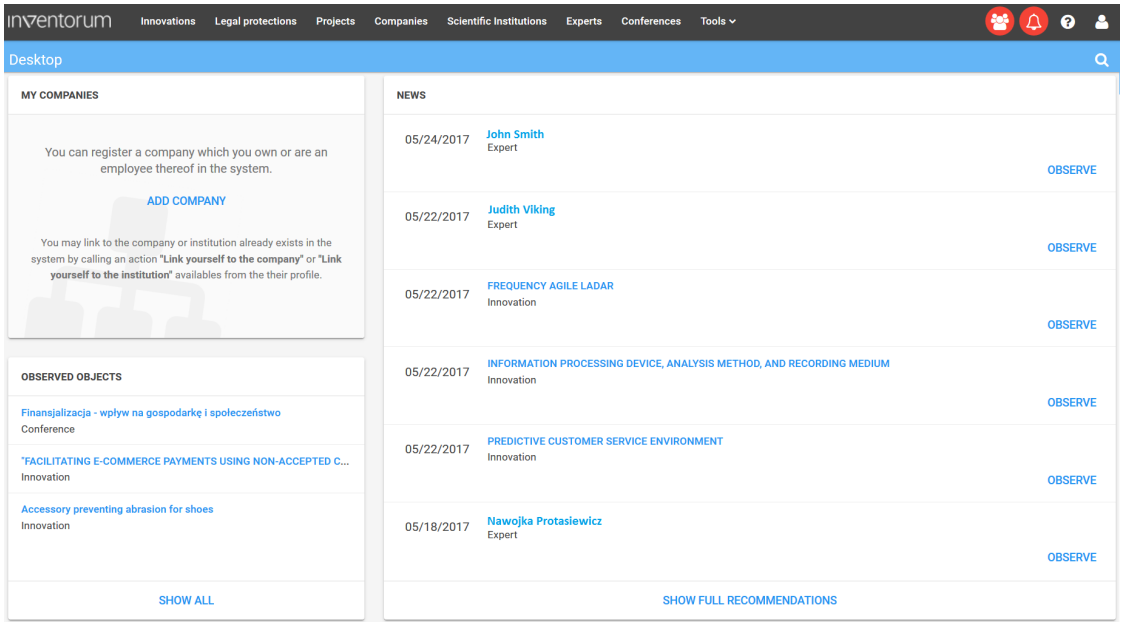

Figure 3. An example of the system desktop, which is composed of four information areas, i.e. (i) news, (ii) observed objects, (iii) companies, (iv) searching and communication tools.

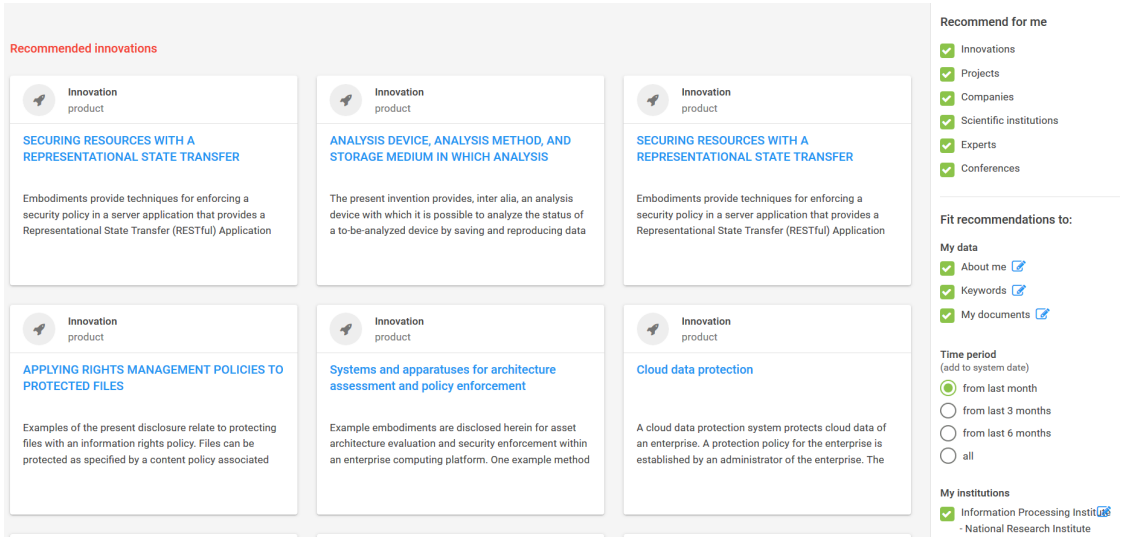

Figure 4. An example of the full recommendation screen: the tools for defining a user profile are on the right, the recommended innovations are in the middle

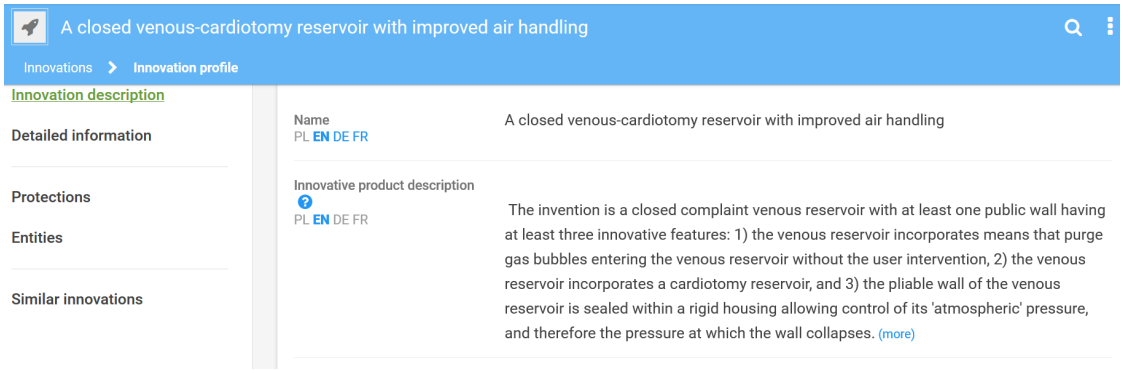

Figure 5. An example of the innovation profile.

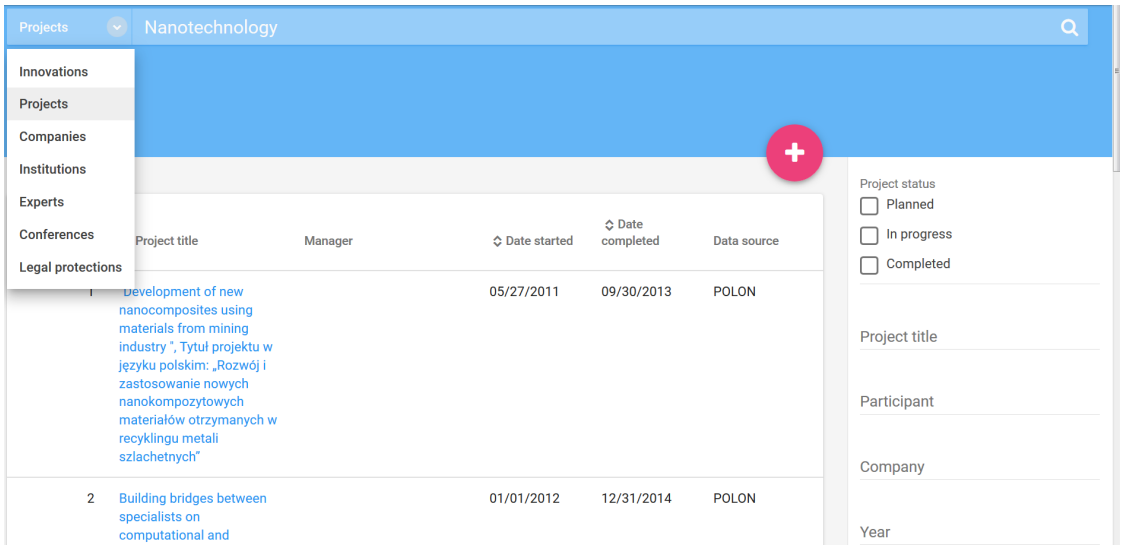

Figure 6. An example of the searching tools: a full-text searching field is on the top, the specific search fields are on the right, the results are in the middle.

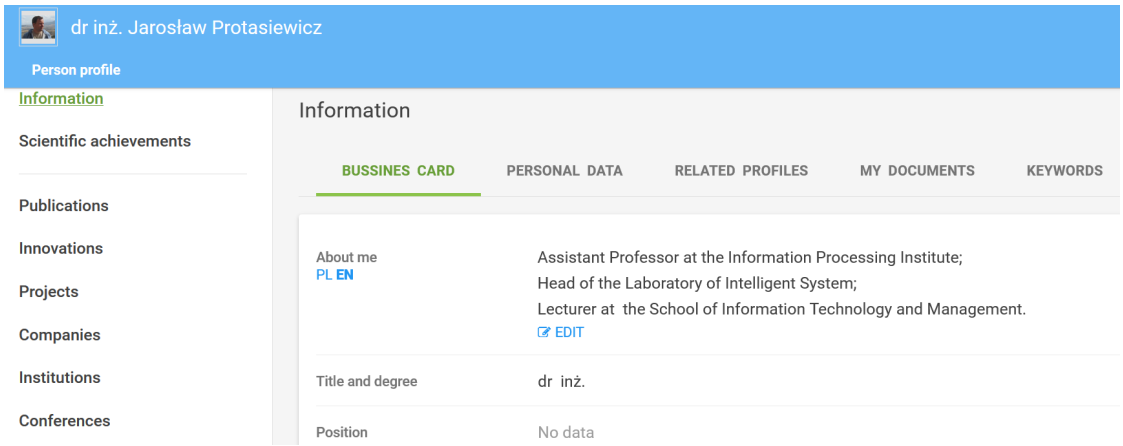

Figure 7. An example of the expert profile.

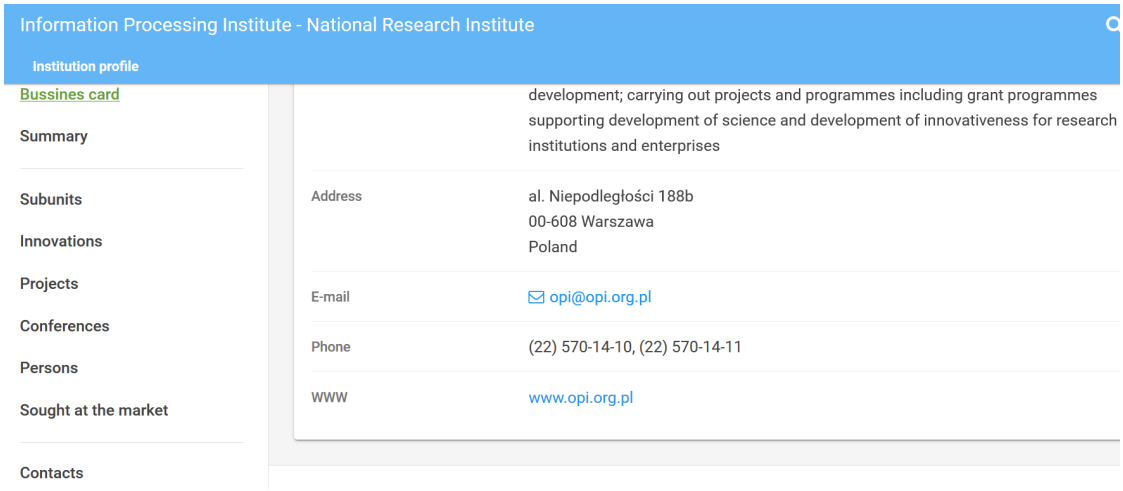

Figure 8. An example of the institution profile.