In 2016, Ontotext's founder and chief executive officer Atanas Kiryakov was selected as one of New Europe 100 outstanding challengers who have a global impact and change the world with ideas that scale up in the digital world.

The nomination of Atanas Kiryakov at NE 100 reads: *"for giving us smarter information"*.

In an interview with Irina Nedeva for the Bulgarian National Radio, Mr. Kiryakov shared more about what it takes to be one of the hundred outstanding challengers from Central and Eastern Europe and how Ontotext with the help of Semantic Technology is transforming the landscape of information and knowledge management.

Read the full interview below.

BNR: What does your software company do and what do you think is the reason you've been included in the NE 100 list of challengers, among the other challengers: 29 in the "Politics and Society" section, 54 – in "Business", 7 – in "Media and Culture" and 10 – in "Science"? What do you think earned you a place in this prestigious ranking?

Atanas Kiryakov: The socialist legacy, I suppose. On a more serious note, what we do ultimately comes down to Artificial Intelligence. Thankfully, there are a lot of talented IT specialists in Bulgaria, and at Ontotext we have managed to gather in one place software engineers with their research work, formerly carried out in the Bulgarian Academy of Science (BAS). I was myself doing a Ph.D. in BAS before I dropped out to found Ontotext.

So, looking at the big picture, Ontotext succeeded in doing all the interesting things that we are talking about today because we were able to draw from a large body of research and from the work of a number of researchers and experts. This is the reason why, in 2000 we were able very quickly to enter the most fascinating and fast developing areas of the Artificial Intelligence field, namely Semantic Technology. Now, I am aware that this does not say much to our audience, so to explain a little bit what we do, it boils down to analyzing texts with the help of big volumes of data.

Linking Data Is the Magic

BNR: We are talking about open data, data coming from large enterprises and institutions,

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

gathering these data on a global scale and opening them in a specific way, right?

Atanas Kiryakov: Yes, absolutely. At Ontotext, we use Open Data a lot. Actually, we are among the pioneers of the so-called Web of Data, or the Semantic Web. The Semantic Web is a next generation of the Web, a Web for machines, in which data are being published across various places, just like the HTML pages are. The difference is that this new-generation Web is not about documents, it is about facts to which a computer program can refer or which it can query the way it queries databases in order to get good results.

The exciting thing about this Web of Data is that you can simultaneously use large amounts of data published across various places.

BNR: And find connections between these data, that is, find information and match it with other information, from a different database? For example, statistical data with financial data.

Atanas Kiryakov: Exactly. Linking data is the most interesting part of this approach, of this new Web. In fact, the term for this kind data is Linked Open Data. These data are linked between servers and data pieces refer to one another, which makes combining information really easy. For instance, some information about Sofia from Wikipedia can be easily linked to information about Sofia from a database with geographic data, or to economic growth data from the World Bank.

Linking data is the magic that makes things exciting.

In a typical project, we usually collect 10, 15 sometimes 20 data sources to build a big knowledge graph, that is a knowledge structure, which very much resembles the structures we build in our brains. The concepts are connected in a similar way.

Whose Paris?

Atanas Kiryakov: Having information collected from different sources about a place or a person, helps us better analyze a text. My favorite example is with Paris. Let's say we need to understand whether Paris is Paris Hilton; Paris, France; Paris, Texas or Paris, the son of King Priam. Computationally, this is an extremely complex task. And this is what we achieve with the help of semantic web technologies. And that is the reason BBC, Financial Times and many

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

other publishers and enterprises work with Ontotext. They work with us because we provide the technology that can accurately, automatically recognize who or what is mentioned in a text.

And this is an extremely difficult task. The process of a machine automatically understanding who Paris is, involves machine learning, many types of heuristics, Artificial Intelligence. This same process is automatic for your editor, they would have no problem recognizing Paris from the context, as people have this graph, this knowledge, in their heads. However, for a machine to do that, it needs data gathered data from various places – from Wikipedia, from different databases.

It is these data that we manage and link together at Ontotext. We connect texts to additional knowledge. Google do that same thing, too. The difference is that they do it for end users on the Web, and we do it for big publishers, financial institutions, etc.

The Web as an Endless Dataset

BNR: You also help media companies, consultancy companies, administrations. We know that the number of databases that are being open according to the Open Data principles is constantly growing. The question is to what extent this linking of databases and this automatic analysis of information is important. And can we speak of a giant automatic brain, larger than any human intelligence taken separately?

Atanas Kiryakov: It would be rhetoric to just say "Yes." Actually, this is not a giant brain. Rather, it is a giant database. Just like the World Wide Web is a giant library. Come to think of it, 30 years ago nothing of this kind existed. Today, the Web connects information and gives you access to it. In a similar way, the Web of Data gives access to data, serving as a giant repository. The difference is that it stores data. Various people publish data across various places and these data get linked together. Think of this Web of Data as a big distributed memory. A giant distributed archive, an Excel table in which you can endlessly link to other Excel tables.

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

Panama Papers and Linked Data

BNR: Let's have a look at one specific example: the Panama Papers. Creating relationships between data enables inferences that otherwise would have been difficult to make. In the Panama Papers, journalists published information about connections between politicians, businessmen, offshore companies. Have you worked with the data from the Panama Papers?

Atanas Kiryakov: Yes, we have. We were the first, after the International Consortium of Investigative Journalists published part of these data as databases, to publish these data as Linked Open Data, within 3 days. Today, many research centers globally use the Panama Papers through our server, where we have linked the data.

BNR: Well, this is why you are on the NE 100 list.

Atanas Kiryakov: Maybe that's part of the story. Let me give you an example that illustrates the usefulness of linking data. When the Panama Papers were first published, these were only records. Records of people and companies that somehow had come out of these documents. And that was all. This alone doesn't help you much in analyzing the data. What we did was to link the data about these people and organizations from the Panama papers to Wikipedia. Now, for the tech savvy and curious, we also have the service Linked Leaks. This linking makes it possible to query data in a meaningful way. For instance, one can find all the politicians from Eastern Europe who are mentioned in the Panama Papers. And this is possible because we linked the Panama Papers leaks with Wikipedia, that is, with additional knowledge.

BNR: A Who's Who of sorts. And the last question: what is the connection between you and the investigative journalists? Do you investigate or do the journalists who work on the case?

Atanas Kiryakov: The journalists are the ones who make the real investigation, we only help with technologies that do the heavy-lifting of linking data.

If you want to learn how to navigate through the gigantic sea of freely released data from the Panama Papers as well as the Linked Open Data cloud, we recommend listening to Atanas Kiryakov's webinar: Diving in Panama Papers and Open Data to Discover Emerging News. This could empower your understanding of today's news or any other information source.

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com