



Smart and human-centred handling of information, knowledge and language

By Bennie Mols Images Ivar Pel



GROUP PASSPORT

RESEARCH FIELD

Data engineering, data management, information systems, text mining, knowledge engineering, knowledge representation, knowledge systems, knowledge extraction

INSTITUTION

HAN University of Applied Sciences

EMPLOYEES as of October 2024)

1 professor (group leader), 7 researchers, 4 lecturers-researchers, 1 developer-researcher, 30 master students (throughout the year).

WEBSITES

Lectureship Data & Knowledge Engineering:
www.han.nl/onderzoek/lectoraten/lector-aat-data-and-knowledge-engineering/

SPRONG project DEMAND:
www.han.nl/projecten/2023/demand/

Project 'Flexible work instructions':
www.han.nl/projecten/2023/flexibele-werkinstructies/

Project SCEPA (Scaling up the Energy Poverty Approach):
www.han.nl/artikelen/2023/10/project-scepa-samen-strijden-tegen-energiearmoede/

The Data and Knowledge Engineering lectureship of HAN University of Applied Sciences focuses on practical applications of existing knowledge. Projects include automating work instructions for SMEs and tackling energy poverty by providing insight into unstructured data.



Stijn Hoppenbrouwers

Whereas academic research groups develop new knowledge, research groups in higher vocational education (HBO), called lectureships, make existing knowledge better and more publicly applicable. Stijn Hoppenbrouwers leads the *Data and Knowledge Engineering lectureship* at HAN University of Applied Sciences, a group of 12 researchers.

‘What binds the group together’, says Hoppenbrouwers, ‘is a love of developing computational applications to deal with information, knowledge and language in a smart and human-centred way. For us, the question is how to apply existing generic knowledge in our teaching and in our collaboration with parties in practice. Then, the focus is not on the leaders, including companies like ASML, but on the peloton, such as small and medium-sized enterprises and municipalities.’

‘The focus is not on the leaders, but on the peloton’



Maya Sappelli

A good example is the two-year RAAK SME project *Flexible work instructions*. This project investigates how work instructions can be made and used in a context-sensitive and flexible way. Hoppenbrouwers: ‘Suppose a heating technician needs to repair a boiler. What, then, is the best way to present a work instruction? From my lectureship, we are investigating how to generate work instructions automatically from various sources. The *Media Design* lectureship investigates how the information can best be presented so that it is usable for different types of boilers, for example. Ultimately, we want to offer SMEs a toolbox to support flexible work instructions.’

Another lectureship with which Hoppenbrouwers works closely is *Applied Science & AI*. ‘That lectureship emerged from our group early this year because it started to grow too big’, he says. ‘The new lectureship is more on the data analytics side. We are more on the linguistic side with knowledge representation, data quality and things like that.’

When it comes to data-driven work, there is still a big gap between what the big tech giants do and what, for example, SMEs do. To close that gap, the SPRONG project DEMAND was created, a collaboration between various institutes, lectureships and business partners, supported by the governing body SIA.



Marijn Siebel



The project is led by the HAN. 'In this project', Hoppenbrouwers explains, 'we can use our expertise in data quality, availability and management. DEMAND is also starting to act as a magnet, prompting external parties to come to us with their questions.'

Ambiguous concepts

Senior researcher Maya Sappelli started five years ago in Hoppenbrouwers' group as a specialist in language technologies like text mining and information retrieval, in which she had gained her PhD. 'What I like about our research group', Sappelli says, 'is that we have a lot of interaction with other parties. At a university, researchers are more concerned with just their own little research domain.'

For example, Sappelli is working on a project with Alliander, the company that develops and manages energy networks. They want to improve the use of information within their own company. 'Part of this is a conceptual framework: what terms are used within the company? And how do employees use those terms in practice? Take a concept like the Dutch term "aansluiting". In Dutch, this is unambiguous: it is a physical connection. At Alliander, however, they also use "aansluiting" in the interpretation of a 'customer'. Detecting those terms with multiple interpretations are important to make knowledge graphs and large language models more usable. Eventually, Alliander wants to build a kind of chat interface that allows employees to efficiently query the company's data sources.'

Searching spider

The headquarters of the *Data and Knowledge Engineering* lectureship is at in the Arnhem location of the HAN. That is where the researchers and lecturers meet physically, but a lot of work is also done from home or from workplaces in the Nijmegen part of HAN. 'We are very flexible about that', says Marijn Siebel,

another senior researcher in the *Data and Knowledge Engineering* lectureship. 'We have a lot of freedom in with whom and how we want to collaborate, including with the other lectureships within HAN.'

As a specialist in query-driven knowledge systems, Siebel's work includes the project SCEPA: *Scaling up the Energy Poverty Approach*. With energy bills rising in recent years, more and more European citizens have run into financial difficulties. Siebel: 'Within the SCEPA project, we are investigating the smartest way to tackle energy poverty in a country, municipality, city or neighbourhood. The project runs until 2027. By then, we want to have a well-grounded framework on how to look at energy poverty. Which information from which sources do we need? How can we make it easier and more intuitive to access this information? I like to think of it as creating a spider that searches for you in a web of information. Ultimately, we want to have a beta version ready, which can be used to take better measures against energy poverty, tailored to the local situation.'

