

Chapter 2

GEEK: Analyzing Online Communities for Expertise Information

Lian Shi

Parque Científico Tecnológico de Gijón Edificio Centros Tecnológicos, Spain

Diego Berrueta

Parque Científico Tecnológico de Gijón Edificio Centros Tecnológicos, Spain

Sergio Fernández

Parque Científico Tecnológico de Gijón Edificio Centros Tecnológicos, Spain

Luis Polo

Parque Científico Tecnológico de Gijón Edificio Centros Tecnológicos, Spain

Iván Mínguez

Parque Científico Tecnológico de Gijón Edificio Centros Tecnológicos, Spain

Emilio Rubiera

Parque Científico Tecnológico de Gijón Edificio Centros Tecnológicos, Spain

Silvino Fernández

Parque Científico Tecnológico de Gijón Edificio Centros Tecnológicos, Spain

ABSTRACT

Finding experts over the Web using semantics has recently received increased attention, especially its application to enterprise management. This scenario introduces many novel challenges to the Web of Data. Gathering Enterprise Expertise Knowledge (GEEK) is a research project which fosters the adoption of Semantic Web technologies within the enterprise environment. GEEK has produced a prototype that demonstrates how to extract and infer expertise by taking into account people's participation in various online communities (forums and projects). The reuse and interlinking of existing, well-established vocabularies in the areas of person description (FOAF), Internet communities (SIOC), project description (DOAP) and vocabulary sharing (SKOS) are explored in our framework, as well as a proposal for applying customized rules and other enabling technologies to the expert finding task.

DOI: 10.4018/978-1-60960-040-2.ch002

GEEK: ANALYSING ONLINE COMMUNITIES FOR EXPERTISE INFORMATION

In the last decades, we are observing a growing need for different types of “multidimensional” expertise. Efficient expertise management is a critical factor of success for organizations. Traditional approaches to this challenge are suffering many difficulties. Firstly, within multinational organizations, enterprises and companies, which are usually built by continuous merging of smaller companies, information systems are heterogeneous, disperse and often redundant. Secondly, a huge number of employees are scattered all over the world, with their expertise-related information spread out. Human resources departments have to manage and deal with hundreds of thousands of employees and countless profiles and areas of expertise. Even if integrated company solutions (such as ERPs, Human Resources Management Systems, etc.) can efficiently manage administrative and personnel information, data about employees’ profiles are decoupled from their daily activity, and have only poor links with their actual expertise. Thirdly, there is a very large variety of competencies with different topics, knowledge and techniques, which means another challenge is to locate given expertise quickly and precisely. The last issue is that traditional IT integration and maintenance of a new system becomes very complex.

Motivated by these significant problems, the GEEK project aims to find a better solution to expertise management. More precisely, GEEK better enables us to extract and infer up-to-date expertise by taking into account people’s participation in virtual communities. In order to achieve this idea, we get help from the power of semantic technologies, which also provide lots of innovation opportunities to implementing the GEEK prototype. There is a range of benefits that can revert into an advanced personnel management. We list the most intuitive ones here: (i) identifying experts that match a given profile; (ii) detecting skill gaps

in order to plan training activities; (iii) building teams with the purpose of internal mobility and agile response to emergency situations.

Semantic Web technologies have reached a maturity that allows managing large amounts of linked data, including information from virtual communities. Consequently, there seems to be potential possibilities for addressing enterprise expertise management by providing precise machine-readable semantic descriptions of the expertise and profiles of employees. The first step is to identify potential information sources on the Web, and to create the mechanisms to gather significant amounts of data. The raw information is usually available in different formats, therefore suitable extraction components are proposed to adapt the data. An ontology is used to unify all the information that comes from different data sources. The next step is to deal with partial descriptions caused by the use of different identifiers for the same resource. This phenomenon is particularly evident when a single individual participates in multiple online communities under different identities (e.g., different email addresses). On the one hand, these repetitive data become a burden for the quality of the collected data. On the other hand, users would be annoyed by redundant answers for a given query. Therefore we apply smushing techniques aiming to identify the co-occurrence of the same person in different communities. A key part of the project is the expertise inference process. We assume that people’s participation in communities is an evidence of their expertise. Then a set of customized rules are used and executed on top of all collected data to derive the expertise. Some mathematical functions aggregate partial expertise evidences into a coherent result. A friendly interface allows users to enter queries and browse the results and the collected data, including experts’ profiles.

22 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/geek-analyzing-online-communities-expertise/50332

Related Content

Fast Single Image Haze Removal Scheme Using Self-Adjusting: Haziness Factor Evaluation

Sangita Roy and Sheli Sinha Chaudhuri (2019). *International Journal of Virtual and Augmented Reality* (pp. 42-57).

www.irma-international.org/article/fast-single-image-haze-removal-scheme-using-self-adjusting/228945

Advanced Visual SLAM and Image Segmentation Techniques for Augmented Reality

Yirui Jiang, Trung Hieu Tran and Leon Williams (2022). *International Journal of Virtual and Augmented Reality* (pp. 1-28).

www.irma-international.org/article/advanced-visual-slam-and-image-segmentation-techniques-for-augmented-reality/307063

Use of Metaverse in Tourism and Hospitality Services

R. Velmurugan, J. Sudarvel, K. Jothi and Ravi Thirumalaisamy (2025). *Cases on Metaverse and Consumer Experiences* (pp. 117-140).

www.irma-international.org/chapter/use-of-metaverse-in-tourism-and-hospitality-services/371261

Integration Visualization Techniques in Groupware Interfaces

Mohamed Daassi, Chaouki Daassi and Marc Favier (2006). *Encyclopedia of Virtual Communities and Technologies* (pp. 279-284).

www.irma-international.org/chapter/integration-visualization-techniques-groupware-interfaces/18086

The Creation and Management of Online Brand Communities

Paola Falcone (2012). *Virtual Community Building and the Information Society: Current and Future Directions* (pp. 15-28).

www.irma-international.org/chapter/creation-management-online-brand-communities/56282