Chapter I
Open Formats, Open Information and Future Trends in Software Engineering

Teemu Saarelainen
Kymenlaakso University of Applied Sciences, Finland

ABSTRACT

The amount of information surrounding us is ever increasing. Usable information is our most valuable asset both in our professional and personal lives. People are constantly struggling to find the information needed in various every day situations. Search engines have given us the tools to cope with the problem so far, but what lies ahead? This chapter describes some of the current trends in information management in software systems and discusses the aspects of using natural language, open formats and open source software in the field of software engineering while keeping in mind the needs of our future society.

INTRODUCTION

This chapter concentrates on the concept of open information management with a software engineering point of view. The significance of finding valid information in our every day lives is becoming more and more important, which has influenced the field of software engineering probably more than any other area of research and development. The purpose of the chapter is to give the reader some background on the development of open formats as well as discuss some of the advances made in web technology and search engines that affect our way of using the internet and accessing information. Interoperability and data portability aspects are also covered and their importance in the software definition and design phases are examined. Some conclusions are drawn based on the current state of open information management in software systems and future trends are briefly discussed.

BACKGROUND

We live in a world full of information, some of which is totally useless, some has a meaning, some is important and some is absolutely vital for us. In order to use this massive amount of information and find the bits we need, we must have the means to filter out certain parts effectively. In the world of internet the answer today is Google. We simply 'google'[a] everything using search criteria described in our own natural language. Interaction with practically any other human being across the globe is a reality and networking with other people using social web applications is a natural part of our lives. How do these systems work? What gives them the ability to grasp users and be so effective? In our everyday lives, we don't give much thought to the inner mechanics of search engines or social web sites – we expect those systems to 'just work', but for computer scientists it's a field of endless research and development.

Before search engines became so effective in the very late 1990's and after the millennium, we used to find the information we needed using internet-directories, which sorted different web-pages and sites into different subdirectories and categories. This system relied totally on the creator of the web site or page having the particular content and the maintainer of the directory to have the correct information in the correct category. In practice, this proved to be an inflexible, limited and difficult system to use, and resulted in outdated directories with outdated information. The solution to the problem was search engines, which at first used simple keywords and a database of indexed documents. Now, the creator of a web page no longer had to report his/her new content into a directory, but instead the search engine crawled through millions and millions of web pages and automatically indexed them according to their content. No keywords are required and we can take advantage of natural language. What data can be exploited in these intelligent systems?

The growing trends in computing today are open information and open software systems. There is a certain amount of ideology behind this, but it doesn't mean that everything related to software is free in the future, which is a common misconception. In practice, open information access is still rarely available for us and the cause often is in the data itself or in the software. To aid in these efforts, open data formats have been suggested and adopted in software systems and organizations more and more. Data exchange between applications is required in many fields, which gives rise to standardized xml-based formats.

Open source software is another hot topic and the concept has even changed the business models of traditional software companies. European countries have traditionally led the development and adoption of open source software. The Finnish government, for instance, has given a statement which says that in all new software acquisitions for the government, open source systems should be preferred and interoperability taken into account.

In software engineering, the processes and best practices are open for everyone to use for their own needs. Thus, using the best current knowledge and the development of even better methods is possible for anyone. Even though making specifications is quite formal and writing program code is governed by strict rules, the most important tool in the documentation is natural language. Model-driven-architecture is a reality and program code can be generated from UML-models, and vice-versa. So, why is it that we still cannot use natural language to construct programs?

Computers only handle data in bits. The rules of the Turing machine govern computers even today. Computer scientists and programmers have managed to build systems that somehow understand natural language as inputs. Nevertheless, a totally new level of abstraction is needed in order to develop a system which could be programmed using something similar to our own natural language.

6 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/open-formats-open-information-future/27788

Related Content

Systems Development by Virtual Project Teams: A Comparative Study of Four Cases
David Croasdell, Andrea Foxand Suprateek Sarker (2003). *Annals of Cases on Information Technology: Volume 5 (pp. 447-463).*
www.irma-international.org/chapter/systems-development-virtual-project-teams/44558

Co-Diffusion Effects in Software Sourcing Arrangements
Niharika Dayyala, Faruk Arslan, Kent A. Walstromand Kallol K. Bagchi (2020). *Information Resources Management Journal (pp. 33-52).*
www.irma-international.org/article/co-diffusion-effects-in-software-sourcing-arrangements/262969

Visualization as a Knowledge Transfer
Anna Ursyn (2019). *Advanced Methodologies and Technologies in Library Science, Information Management, and Scholarly Inquiry (pp. 423-437).*
www.irma-international.org/chapter/visualization-as-a-knowledge-transfer/215944

Adaptive Four-dot Median Filter for Removing 1-99% Densities of Salt-and-Pepper Noise in Images
Xin-Ming Zhang, Qiang Kang, Jin-Feng Chengand Xia Wang (2018). *Journal of Information Technology Research (pp. 47-61).*
www.irma-international.org/article/adaptive-four-dot-median-filter-for-removing-1-99-densities-of-salt-and-pepper-noise-in-images/206214

Enhancing New Product Development Effectiveness With Internet of Things Origin Real Time Data
Samir Yerpudeand Tarun Kumar Singhal (2018). *Journal of Cases on Information Technology (pp. 21-35).*
www.irma-international.org/article/enhancing-new-product-development-effectiveness-with-internet-of-things-origin-real-time-data/207364