Preface

Since the beginning of the 21st century, the Internet has become a widely used social platform. Web 2.0, in particular social software, again emphasizes the crucial role of social relations in knowledge production and diffusion. Net surfers collaborate more and more (implicitly and explicitly) to manage the available knowledge on the Internet. Among the various forms of collaborations, two of them receive increasing attention:

- **collaborative search** tools allowing net surfers to benefit from the searches done by others to access relevant information (for example collaborative filtering tools);
- **communities (of interest)** formed by a collective of net surfers with common interests where knowledge is produced collectively and shared (for example the Open Directory Project).

For many specialists, these two forms of collaborations are crucial to manage the large amount of accessible information on the Internet or in intranets. Despite their success in practice, several questions are raised. There are three types of open issues: (i) the different approaches for implementing these forms of collaborations, (ii) the possible ways to make them work in practice, and (iii) their impact on the environments in which they are deployed. The aim of this book is to propose some answers to these different questions.

The first chapter introduces the context of these two forms of collaborations, explains the main reasons for their success and their (mostly ignored) potential dangers, and details each specific contribution. The second chapter studies the concept of “community” as defined by sociologists and demonstrates that its use in the context of the Internet is not really adequate. It proposes the concept of “project-driven mediated collectives” to better describe this phenomenon in the future.

The first set of chapters is related to innovative technologies providing solutions to detect communities of interests and/or to share knowledge. Chapter 3 overviews the recommender systems (a particular type of collaborative search tools), their approaches, their preference indicators, the algorithms used and their predictions mechanisms. In chapter 4, a method is proposed to detect communities of net surfers sharing common access patterns based on latent semantic analysis, and then to recommend Web pages. Another approach to make recommendations is described in chapter 5, where link analysis models that exploit the graph formed by the users and items they rated, are enumerated. Chapter 6 describes an algorithm that detects communities of interests based on the relevance of the documents read by the users. Chapter 7 reviews a P2P system that collectively organizes knowledge. The specific coordination mechanism used to allow different peers to find other peers sharing similar knowledge perspectives is detailed. Chapter 8 introduces a novel algorithm to filter relevant information for a community and shows how it can be integrated in Web observatories.
A second set of chapters studies how tools, communities of interest and knowledge sharing impact the environment in which they are used. Chapter 9 shows how Twitter may change the practices of journalism since the journalists, their sources and their audiences can now be directly connected through microblogging. Two generic theoretical frameworks are then analyzed. In chapter 10, a framework examines the important (and rarely understood) problem of tacit knowledge sharing. The framework introduced in chapter 11 studies how social software can be deployed in organizations or directly on the Internet. Finally, chapter 12 reports an interesting approach that helps members of a large community to better know each other (an important issue for trust building): an autograph book built by the different members and based on a collaborative platform.

Every researcher involved in studying collaboration on the Internet should be interested in this book. The concepts of communities of practice and collaborative search are defined and their theoretical and practical limits studied. The “technical” chapters describe several technologies and algorithms used to solve these problems. The researcher can not only find useful models for his or her own needs, but also an overview of the kind of information exploited by computer programs. The latter is important to understand the hypothesis behind the different approaches, and the inputs and outputs for the users. The “applications” chapters propose several examples of the impacts of communities of practice and collaborative search in real life. It is of course impossible to cover every possible impact, but the diversity of the examples is large enough to inspect the main elements.

This book outlines that studying the different forms of collaborations on the Internet has several dimensions (sociological, psychological, technical, etc.). While there are a lot of works about these different aspects, they are mostly published in different journals and books, and diffused in separate research communities. By treating these dimensions in a same book, the authors hope to build a bridge between these research communities and to initiate fruitful discussions between them.

I thank the authors for their contribution as well as the reviewers for their work and numerous interesting remarks. I will also thank Emilie Vossen for her help to finalize this book.

Dr. Ir. Pascal Francq
November 2009