HELSINKI UNIVERSITY OF TECHNOLOGY
Faculty of Electronics, Communications and Automation

Johan Tötterman

INTERNET STANDARDIZATION: A PARTICIPANT ANALYSIS

Thesis submitted for examination for the degree of Master of Science in Technology

Espoo 17.8.2009

Thesis supervisor:

Prof. Heikki Hämmäinen

Thesis instructor:

Timo Ali-Vehmas

TEKNILLINEN KORKEAKOULU — DIPLOMITYÖN TIIVISTELMÄ

Tekijä: Johan Tötterman

Työn nimi: Internet-standardointi: Osallistuja-analyysi

Päivämäärä: 17.8.2009 | Kieli: Englanti | Sivumäärä: 8+87

Tiedekunta: Elektroniikan, tietoliikenteen ja automaation tiedekunta

Professuuri: Tietoverkkotekniikka | Koodi: S-38

Valvoja: Prof. Heikki Hämmäinen

Ohjaaja: Timo Ali-Vehmas

Tässä diplomityössä tarkastellaan tieto- ja vientintätekniikka-alan (ICT) standardointityötä. Tutkimme erityisesti Internet-standardointia. Aikaisempi tutkimus osoittaa, että standardit luovat perustan teknologioiden yhteentoimivuudelle ja, että ne mahdollistavat sekä tietojen että informaation välittämisen järjestelmien välillä. Standardit toimivat siten edesauttajina ja katalyytteinä, joista on sekä taloudellista että teknistä hyötyä. Aikaisemmat tutkimukset osittavat myös, että usean johtavan toimijan ydinstrategiaan kuuluu nykyään standardien kehitystyöhön osallistuminen. Osallistuminen standardointityöhön saattaa myös olla avain tulevaan kaupalliseen menestykseen. Tässä työssä tutkimme ensisijaisesti mikä Internet-standardoinnin menestyksen mittari ja perustaso on. Vertailemme myös tiettyjen yhtiöiden standardointipanostuksia ja niitten yhteyttä yhtiöiden kaupalliseen menestykseen. Lisäksi tutkimme myös erikseen suomalaisten toimijoiden saavutuksia IETF:ssä.
Kirjallisuuskatsauksessa otamme tarkemmin esille, missä, miksi ja miten ICT-alan standardeja kehitetään. Selvitämme myös tutkimus- ja kehitystyön yhteyttä standardointiin sekä määritämme motivaatioita standardointityöhön osallistumiselle. Osana tätä diplomityötä suunnittelemme ja kehitämme ohjelmiston sekä tietokannan joka antaa meille mahdollisuuden tallentaa, käsitellä ja tutkia useita eri Internet Engineering Task Force:n (IETF) työprosesseja ja -dokumentteja. Hyödyntämällä tietokantaamme ja kehitettyä ohjelmistoa voimme mitata ja analysoida useita IETF:n standardointiprosessin näkökulmia ja myös tutkia lähemmin siihen osallistuvia yrityksiä.
Tuloksemme osoittavat, että Suomen ICT-klusterin aikaansaannokset IETF:ssä ovat verrattain hyvät. Lisäksi voimme todeta, että Cisco:n saavutukset voidaan pitää menestyksen mittarina lähes kaikkia IETF:n standardointityön osa-alueita tarkasteltaessa. Tulostemme perusteella ehdotamme myös, että osallisujien kaupallisella ja standardointityön menestyksellä on yhteys.

Avainsanat: Internet, tieto- ja viestintätekniikka (ICT), standardi, standardointi, Internet Engineering Task Force (IETF), ICT-klusteri Suomi

HELSINKI UNIVERSITY OF TECHNOLOGY

ABSTRACT OF THE MASTER'S THESIS

Author: Johan Tötterman

Title: Internet Standardization: A Participant Analysis

Date: 17.8.2009 Language: English Number of pages: 8+87

Faculty: Faculty of Electronics, Communications and Automation

Professorship: Tietoverkkotekniikka Code: S-38

Supervisor: Prof. Heikki Hämmäinen

Instructor: Timo Ali-Vehmas

This thesis examines standards-setting in the Information and Communications Technology (ICT) industry. Special attention is given to Internet standardization. Previous research suggest that standards lay the ground for compatibility, interoperability, and interchange of data in the ICT field. Standards thus function as enablers and accelerators with both economical and technological benefits. Previous research also suggests that participating in standards development and influencing the outcome by contributing to the standardization process have become core strategic choices of many leading players. Participating in the development of a winning standard can be critical to later business success. In this thesis we will therefore aim to clarify what the benchmark for success in Internet standardization is. We also compare selected organizations' standardization activities to figures measuring their success on the marketplace. The standardization achievements of the Finnish ICT cluster are also given extra attention.

Our literature study elaborates on how, why, and where ICT standards are developed. The relationship between Research and Development (R&D) and ICT standardization is clarified and we also establish motivations for participating in the standards development process. As a part of this thesis we design and create a database that enables us to retrieve and process all working documents related to the Internet Engineering Task Force (IETF) standardization process. Using the database and custom tools created for this task allows us to measure and analyze several aspects of the IETF standardization process and the participants active therein.

The results suggest the Finnish ICT cluster has performed comparatively well within the IETF, that Cisco's achievements can be considered the benchmark for success regarding virtually all aspects of IETF standardization, and that there is a linkage between participants' success in standardization and their merits on the marketplace.

Keywords: Internet, Information and Communications technology (ICT), Standard, Standardization, Internet Engineering Task Force (IETF), Finnish ICT Cluster

Preface

Writing this masters thesis represents a formidable personal achievement to me and it is a fine ending to my *slightly* elongated studies at the Helsinki University of Technology. This thesis was written as a part of my work at Nokia's Corporate Development Office in 2008.

During the course of writing this thesis I have received support from many directions and thus find myself indebted to everybody who have helped me. First and foremost I would like to thank Mr. Timo Ali-Vehmas at Nokia for making this work possible and for working as my instructor. My sincere gratitude also goes to Prof. Heikki Hämmäinen for supervising this thesis.

Special thanks also goes to Mr. Roope Kylmäkoski and all my colleagues and friends who have helped me complete this thesis.

Helsinki, 17.8.2009

Johan Tötterman

Contents

List of Acronyms and Concepts

Acronyms

3GPP	-	3rd Generation Partnership Project
ACTS	-	European Advanced Communications, Technologies and Services Programme
AD	-	IETF Area Director
ARPA	-	U.S. Advanced Research Projects Agency
ARPANET	-	U.S. Advanced Research Projects Agency Computer Network
BOF	-	IETF Birds-of-a-Feather meeting
BSI	-	British Standards Institution
CEN	-	European Committee for Standardization
CEN/STAR	-	CEN Standardization and Research
COPRAS	-	European Co-operation Platform for Research and Standards
CSNET	-	U.S. National Science Foundation Computer Science Network
DARPA	-	U.S. Defense Advanced Research Projects Agency
DIN	-	German Institute for Standardization
DOD	-	U.S. Department of Defense
DSAB	-	Distributed System Architecture Board
DTI	-	UK Department of Trade and Industry
ETSI	-	European Telecommunication Standards Institute
FNC	-	U.S. Federal Networking Committee
FRICC	-	U.S. Federal Research Internet Coordinating Committee
GMPLS	-	Generalized Multi-Protocol Label Switching
HIIT	-	Helsinki Institute for Information Technology
HIP	-	IRTF Host Identity Protocol
HTTP	-	HyperText Transfer Protocol
IAB	-	Internet Architecture Board
IAB	-	Internet Advisory Board
IAD	-	IETF Administrative Director
IANA	-	Internet Assigned Number Authority
IAOC	-	IETF Administrative Oversight Committee
IASA	-	IETF Administrative Support Activity
ICANN	-	Internet Corporation for Assigned Names and Numbers
ICCB	-	Internet Configuration Control Board
ICT	-	Information and Communications technology
ICT SHOK	-	Finnish Strategic center for science, technology and innovation in the field of ICT
IEEE	-	Institute of Electrical and Electronics Engineers
IESG	-	Internet Engineering Steering Group
IETF	-	Internet Engineering Task Force
IMT-2000	-	International Mobile Telecommunications-2000 protocol
INTEREST	-	European programme for Integrating Research and Standardization

Acronyms continued

IP	-	Internet Protocol
IP/MPLS	-	Internet Protocol/MultiProtocol Label Switching
IPSEC	-	IP Security Protocol
IPV6	-	Internet Protocol Version 6
IRSG	-	Internet Research Steering Group
IRTF	-	Internet Research Task Force
ISO	-	International Organization for Standardization
ISOC	-	Internet Society
ISTAG	-	Information Society Technologies Advisory Group
ITU	-	International Telegraph Union
ITU	-	International Telecommunication Union
JTC1	-	International Organization for Standardization Joint Technical Committee 1
LAN	-	Local Area Network
LTBP	-	International Organization for Standardization Joint Technical Committee 1 Long Term Business Plan
MPLS	-	Multi-protocol Label Switching
NCP	-	ARPANET Network Control Protocol
NSF	-	U.S. National Science Foundation
NSFNET	-	U.S. National Science Foundation's Computer Science Network
OAUTH	-	Open Web Authentication protocol
OECD	-	Organization for Economic Co-operation and Development
OGC	-	Open GIS Consortium
OMA	-	Open Mobile Alliance
OSI	-	Open Systems Interconnection Reference Model
POC	-	Push To Talk Over Cellular
PPPEXT	-	Point-to-Point Protocol Extensions
R&D	-	Research and Development
RACE	-	Research in Advanced Communications in Europe
RFC	-	Request for Comments Document Series
RG	-	Internet Research Task Force research Group
SDO	-	Standards development organization
SSH	-	Secure Shell network protocol
TCP	-	Transmission Control Protocol
TCP/IP	-	Transmission Control Protocol / Internet Protocol
UCLA	-	University of California at Los Angeles
USPTO	-	U.S. Patent and Trademark Office
VITA	-	VMEbus International Trade Association
VPN	-	Virtual Private Network
W3C	-	World Wide Web Consortium
WG	-	IETF Working Group
WLAN	-	Wireless Local Area Network

1 Introduction

Information and Communications Technology, including Internet technology, is becoming ubiquitous. We are experiencing ongoing horizontal and vertical convergence in the field. The horizontal convergence can be described as the convergence of technologies, the ICT field is a blend of basic IT technologies, telecommunications technologies, and consumer electronics. It includes systems, tools, and applications dealing with the capture, representation, accessibility, processing, security, transfer, interchange, presentation, management, organization, storage and retrieval of information [82]. Vertical convergence on the other hand can be described as the progress of ICT into all aspects of everyday life, e.g. home, administration, economy, and education. Many conventional tasks are now utilizing the capabilities offered by ICT. The broad scope of information and communications technologies constitutes interoperability as the real impact received from standards.

The ICT industry has changed immensely in the last decade. The way people use and share information has changed and still evolves constantly. The challenge is to develop ever more complex information systems under shrinking schedules. Many ICT applications are interdisciplinary by nature and involve several different technologies. This reflects on standardization and makes it a complex discipline, as ICT systems in general cannot be covered by a single standard. Close co-operation between experts in various domains is demanded, both in product and in standards development.

So why do we need standards and what is the value of standards and specifications? The wide scope, complexity and ICT's all-pervasive nature directly influence the value composition of ICT standards. It is impossible to define the value of ICT standards in one simple statement, as there are several different types of qualifications needed, to assess their value. The International Organization for Standardization's (ISO) Joint Technical Committee 1 (JTC1) documents the following taxonomy for the value domains of standards and standardization: Economical and commercial value, technical value, and political and public interest value [81]. Hurd and Isaak also conclude that when standards are done right, we have a social and economic benefit for all stakeholders involved [63].

From an economic point of view, standards provide efficient dissemination of innovation and can be seen as a precondition for economic growth. Standards also create economics of scale and they can in that way reduce costs of products and services. Standards bring technical value through enhanced quality, they may efficiently reduce variety, and standards also promote understanding of technology by providing and disseminating information and thereby empowering the spread of new technology. Political and public interest value is created through increased interoperability, ensuring competitiveness on the market, and creating products that meet safety and security requirements. [10] [136] [143]

ICT standards do not grow on trees. The development of standards, i.e. standardization, requires tremendous efforts from a multitude of stakeholders. ICT standards are developed in several alternative environments. Cargill divides the ICT standards development field into five basic variants: trade associations, national- and international formal standards development organizations (SDO), consortia, alliances, and the Open Source movement [19]. Some of the above-mentioned are open, some proprietary. In this thesis we focus on voluntary ICT standards that are developed through an open and transparent process, where contributions are accepted from any interested party and decisions are made through some kind of consensus mechanism. The standards development processes can be seen as a very specific type of open collaboration between companies, users, consumers and other stakeholders [7]. Gaining consensus between participants in the standardization process will in theory produce the most sound and technically competent standards. Whether or not to participate in the standards development process thereby boils down to the question if a stakeholder is to pursue its ideas on their own or bring them into a common knowledge pool [9]. Among common incentives for participation in standardization is that you are awarded a voice in the process and can accordingly affect the content of the standard to be aligned with your own assets. You also get an inside view on the development of new technologies and can concurrently build standards conformance early on. A final major motivation for participation is the supply of insider knowledge, early access to information is considered particularly important [38].

This thesis will dive deeper into ICT and Internet standardization through the Internet Engineering Task Force. We will try to quantify the value of ICT standards and the value of participation in standards development using data collected from IETF standards and working documents. The IETF data will be stored and processed in a database designed as a part of this thesis work. The database will be referred to as the *Project Database* in this text. Special focus will also be set on analyzing the achievements of the Finnish ICT cluster in the IETF.

Previous research involving data mined from a SDO's working documents is fairly sparse. Simcoe [127] analyzed the time required to reach consensus in IETF working groups and Leiponen [99] [100] focused on a firms ability to influence standardization in the Third Generation Partnership Project (3GPP) through different co-operative schemes. The value of standards, network effects, strategies and tactics in standardization, the relationship between Research and Development, innovation and standardization, and the motivations for participating in standardization are however discussed in several papers, e.g. Techapalokul, Alleman and Chen 2001 [131], Katz and Shapiro 1985 [90], Besen and Farrell 1994 [6], Brusse 2005 [16], and Blind 2006 [7].

1.1 Research Problem

Using knowledge from existing literature and the data we have acquired from the IETF working documents, we aim at giving an answer to the following question: *What is the benchmark for success in IETF Internet standards development work?*

Related to the principal research problem we furthermore try to establish in more detail: *What is the success of the Finnish ICT cluster in IETF standardization work?*

Finally, to give a business viewpoint on our topic we set out to establish: *What is the correlation between participating in IETF standardization and the participants' success on the marketplace?*

1.2 Research Scope

The scope of this thesis is outlined to cover ICT and Internet[1] standards jointly elaborated by diverse actors in voluntary standards development organizations. Specific weight is set on participation in standards development. We study the internal operations and the achievements of the participants in IETF Internet standardization. Based on prior literature and our own analysis we try to establish the value of, and the keys to success in standards development. Patent issues, royalties, and intellectual property rights discussions are disregarded in this thesis.

1.3 Research Method

As a part of this thesis we design and create a database for the purpose of storing all IETF standards, working documents, and other relevant data regarding the IETF standards development process. Moreover, we also design and create a set of tools for extracting and parsing data from the working documents. The collected data is primarily concerning the working document authors and their organizational affiliations. Other aspects of the IETF standardization process is also covered, and some additional data about the process itself is also gathered. The assembled data will be used as the kernel of our analysis.

1.4 Thesis Structure

This thesis is divided into six chapters. The chapters cover the following topics:

Chapter 1 introduces the reader to the field of research, outlines the scope, and clarifies the research problem for this thesis.

Chapter 2 elaborates on several topics based on relevant literature. The background and specific characteristics of ICT standardization are discussed. The different venues for standards development, SDOs, consortia, and fora, are also covered. Further discussion about ICT Standardization's link to innovation and R&D, the value of participating in ICT standardization, and finally some problems regarding the current ICT standardization paradigm are also covered in this chapter.

[1] The term *ICT standardization* is commonly used in academic literature cited in this thesis. Internet standardization can be considered a part of the ICT standardization domain and we will therefore use the aforementioned term to cover both ICT, IT and Internet standardization in this text.

Chapter 3 introduces the Internet Engineering Task Force to the reader. We document the background and some major achievements of the IETF. A closer look is also taken at the working methods, structure, and field of work. IETF members, participants, and stakeholders are also identified.

Chapter 4 leads us into the data gathering part of this thesis. A data gathering strategy and tools for the task are documented in this chapter. We also reflect on the types of documents and data we are to collect. This chapter also covers issues about the limitations and reliability aspects of the collected data. Some more technical issues concerning the data-mining task are also reflected upon.

Chapter 5 is the analysis part of this thesis. Here the reader is presented with the strategy how the correlation between success in standardization and success on the marketplace and the accomplishments of the Finnish ICT cluster are going to be quantified.

Chapter 6 consists of the results of the analysis and the conclusions of this thesis.

2 ICT Standardization Paradigm

This chapter discusses ICT standardization. We take a look at the background and the underlying principles of standards and standardization. The value of standards is discussed and the factors that differentiate ICT standards development from other standardization domains are also elaborated upon.

2.1 Background on Standardization

Traditionally standards have mostly defined attributes of tangible objects. Standards specified dimensions, materials and other physical attributes of products. Some intangible elements were also treated, but those were primarily result-oriented, like performance and safety. Early standards were created by domain experts within the industry that produced the products involved. The scope of standards was for the most part national and a single standard was enough to solve the issue at hand. Innovation cycles were rather slow and standards setting was a wise investment, as decisions on standards often were expected to be long-lived. [138]

Following the Second World War the standardization scene started to progress. In the U.S. the American National Standards Institute (ANSI) came into prominence. Several other countries also strengthened their national standards bodies to be a part of their industrial policies. The growth of the internationalism after the Second World War also led to the establishing of the International Organization for Standardization (ISO). The improved standardization field enabled advanced homogeneity of markets, thereby also enabling economies of scale and the ability to innovate more quickly. Users' needs were served better, and consequently the way for even more innovation and further increased sales was paved. [19]

In the information and communications industry, standards have been particularly important since the beginning. The early telegraph communication systems were burdened with the lack of interoperability. Telegraph lines tended to cease whenever they reached a national border. The need for interoperability was however noticed and the International Telegraph Union (ITU) was established as early as 1865 [79]. The International Telegraph Union later developed into the International Telecommunication Union and continued publishing recommendations for communications technology.

The ICT industry is also distinguished by the fact that each new generation of technology brings increased complexity. Hence, standards are considered pivotal, because they allow interoperability of products, services, and hardware and software from different parties. It can be argued that the Internet would not have achieved its current ubiquitous presence, where it is accessible from almost any type of computer platform and device, if it did not use widely accepted technical standards in its networking infrastructure and supported services [59].

Initially standards were created primarily in trade associations. These associations were gatherings of professionals who were experts in a particular field. These groups were set up to create specifications embodying their wisdom for the sake of their professional community [19]. From trade associations standardization work transferred in parallel with the evolving industry value chain to an initially national and later international formal standardization process in SDOs like the ISO, the ITU, the ANSI, and the Institute of Electrical and Electronics Engineers (IEEE) [17]. Today the standards development field is more diversified. Universality accompanied by technological and commercial development promote a variety of standardization patterns, we have non-proprietary and proprietary standards, formal *de jure* standards developed by one of the publicly recognized SDOs, and on the other hand market induced *de facto* standards [113]. Especially ICT standards development can be characterized by a hybrid system between formally set and market selected (de facto) standards, e.g. Farrell and Saloner [46], Vercoulen and van Wegberg [145] and Jakobs [87] discuss this diversity and different features of the ICT standardization processes.

Standardization is also closely linked to research and development (R&D) and innovation. It has been recognized that the interface between standards and R&D is crucial to both activities. Especially in ICT research and development, an essential part is to assure that standardization and research are proceeding in parallel, thereby enabling cross fertilization between the two [16]. Classic discussion also involves whether standardization actually should occur *ex ante* or *ex post* of R&D, e.g. Kristiansen elaborates on the timing between R&D and standardization [96]. Moreover, the aspect of reduced risk brought by technical standards provide suppliers with a more secure set of interfaces around which to design a product, and thus standards may also encourage R&D of even more new components [54].

Most standards in use today have been created through voluntary consensus processes in which any stakeholder is entitled to participate [137]. Motivation for participation is discussed in several papers. Hurd and Isaak conclude that participants are provided with a view into the market, participants also have a voice in the process to ensure that the standard accommodates their plans. Being proactive can furthermore help to set the tone in the entire standardization process [63]. Participating in standardization also has strategic and competitive implications. By contributing to the standards development process participants can largely affect he competitive environment they operate in [6].

Standards are not developed without an objective, underlying drivers are repeatedly present. These reasons might be very diversified, but some of the primary reasons that give momentum to standardization can anyhow be isolated and divided into two main categories. Market tied motivators, and non-market tied motivators. Some market-tied motivators for standards are: network effects [90], path dependence [103], compatibility and interoperability factors [45], and product differentiation. Non market-tied factors are often regulative in nature and government tied. It is also possible that the need for a standard rises from within a specific industry or community. In the ICT domain all of these factors have certainly affected standardization. Standards can also emerge without being specifically developed, a *de facto* standard can arise as a result of a dominant design or a monopoly situation.

Finally, to be clear on what a standard is we will use the following definition of a standard in this document. First we will however establish the meaning of the term *Technical specification*. A technical specification is in this thesis referred to, as a description contained in a document which lays down the characteristics required of a product, including levels of quality, performance, safety, dimensions, and conformity assessment procedures. A *Standard* is a technical specification fulfilling the aforementioned criteria and that is approved by a recognized standardization body for repeated or continuous utilization. [143]

In this thesis we refer to *Standardization* as a consensus-driven activity, carried out by, and for the interested parties themselves. The process should be based on openness and transparency within independent organizations, and aim to establish the voluntary adoption of, and compliance with standards [29].

2.2 The Value of Standards

ICT standards are growing increasingly complex as the field is distinguished by hitherto distinct technologies converging into ubiquitous entities. Accordingly, the value composition of standards also increase in complexity and value is created at multiple distinct levels. However, the basic taxonomy for the value of standards seems quite durable. T. R. B. Sanders quantified some of these general aims of standardization in an ISO publication from 1972 [123]. Sanders established that standards bring value as follows:

- Standards reduce the growing variety of products and procedures
- Standards enable communication
- Standards contribute to the functioning of the overall economy
- Standards eliminate trade barriers
- Standards contribute to safety, health, and protection of life
- Standards protect consumer and community rights

Sanders' points are still valid today. As a continuum, we will here further project them on the current ICT standardization field and also bring in some of the most recent nuances of standardization value.

Sanders' first point stated that standards bring value as a tool to reduce variety of products and procedures. According to Swann variety reduction accomplish two different functions [130]. First, variety reduction empowers economies of scale by minimizing the wasteful growth of marginally differentiated products. Businesses can operate on a larger, more efficient scale of production and consequently goods can be produced at lower cost. Secondly, Swann mentions an even more important role for variety reduction. Variety-reducing standards can help to reduce the risks faced by suppliers as standards bring more certainty about the future direction of the industry. Reduced risks often imply more competition, but standards are seen as instrumental in the development and growth of new markets and thus benefit is seen in its entirety, both for producers and customers. Reducing variety can on the other hand also become a trade-off between choice and price.

Standards and standards development furthermore help to reduce the variety of procedures and working methods. Blind raises the issue of research joint ventures. He states that companies active in R&D are likely to also participate in standardization to be able to further build on their previous activities and to reach marketable products and technologies compatible with other offerings on the market. Research and development expenditure is therefore decreased as standardized procedures reduce a duplication of efforts. [7] Standardization can thus improve the success and reduce the costs of introducing new technologies to the market.

The second point was that standards enable communication. In ICT communication is enabled by the means of compatibility and interoperability standards. In the Internet and ICT domains standardized interfaces and protocols are the underlying instrument for successful communication. These domains are marked with the distinction that the utility a user derives from consumption of the good increases with the number of other agents consuming the good [90]. This is called a network effect [102]. In a market characterized by network effects, compatibility standards emerge to help grow the size of the network and, hence, bring value throughout the domain. Compatibility can boost the amount of complementary goods, leading to products becoming cheaper and more readily available [45]. Standards promoting interoperability bring value as it can be assumed that devices work together and communicate in an expected way, and thus users do not need to choose a specific technology or replace equipment as often [40].

Standards also contribute to the functioning of the overall economy. The German Institute for Standardization (DIN) conducted a substantial research in 2000, studying the economic benefits of standardization and specifically the benefits of standards for the economy as a whole [38]. The study concludes that the efficient dissemination of innovation through standards is a decisive factor for economic growth. Further macroeconomic value can also be reached, as standards not only act as positive stimulus for innovation, but they also provide information and thus reduce uncertainty. Standards thereby also mitigate the risk of R&D investments. The openness of standards thus play a central role as they act as a public infrastructure for innovation [130]. A UK study, conducted by the British Standards Institution (BSI), estimates that around 13% of post Second World War UK productivity growth can be attributed to the gain achieved by the increased availability of information and knowledge through standards [136].

Standards also help to eliminate trade barriers. International standards that focus on compatibility, product information and measurement are expected to increase international trade. The DIN study from 2000 summarizes that the very existence of standards has a positive effect on trade, and that international standards act as catalysts in diffusing new technical knowledge and thereby provide further advantages for the international trade [38]. Trade barriers are also diminished as standards encourage market entry and enhance competition by clearly defining what is required to serve a market [84]. Standards can furthermore affect international trade because they act as a form of non-price competition.

To build long term value the standardization process should comply with certain conditions. Standardization should be done in such a manner, and only when it provides added value. Standards development should also be based on evidence, economics, and experience. According to Neelie Kroes, the European Commissioner for Competition Policy, it is important that standardization agreements are made based on the merits of the technologies involved. If factors other than the technical merits influence the standardization process the risk is that the resulting standard becomes inferior and possibly burdened with anti competitive effects to the market. Striving towards interoperability, through a transparent and well-documented process, will advance competition between technologies from different companies, and in that way lock-in situations will be prevented. [97]

Finally, on a generic level, standards also introduce many benefits to society at large. There are many standards directly contributing to safety, health, and the protection of life. In many markets where the playing field is outlined by standards, increased consumer and community rights can be seen. When the core technologies are standardized, value can be added when competition is shifted to the implementations level.

2.3 Specifics of ICT Standardization

There are several issues that distinguish ICT standardization from standards development in many other industry domains. The complexity and short product life cycles stress the need for timeliness, functionality, and interoperability. A standard should be available before the technology is obsolete. Cargill among others have noticed that in some cases the product life cycle is in fact shorter than the standards development cycle and that this reality leads to some of the idiosyncrasies of ICT standards development [17].

Short product life cycles not only stresses the speed of, but also directs a requirement of flexibility on the standardization process. Several SDOs have reacted and introduced more flexible solutions including, remote on-line sessions, mailing lists, fats-tracks, and workshop agreements to accelerate the standardization process [8]. David and Shurmer further discuss some of these procedural reforms SDOs have had to make in order to speed up their processes [35]. The high speed of technological change has also led to, not only an increased amount of standards, but sometimes also to less transparent and more complex processes.

A further characteristic of ICT standardization is that technological convergence is blemishing the divergence between both products and services. Previously used methods to define the field of work for different SDOs might not be accurate any more. In closely related, or converging, technical areas conflicts between standards setting organizations have consequently been noted. Formal liaison statements and co-operation agreements have been developed to clarify the situation. However, the pace of technological advance again makes this especially challenging in ICT standardization. [35]

Complexity is also added by the fact that the ICT is a truly international industry. Standardization is characterized by participation of both multinational organizations, national governments and other stakeholders. The technology itself has also experienced several generations of evolution in many ICT domains. Setting complex technology standards within and between different generations of technology in an international domain is considered especially demanding. [50]

As a direct continuum to the complexity of both the ICT technology and standards development, we can also see an increased diversity among the standard setting organizations in the industry. Updegrove reasoned that the problems introduced by new ICT technologies can rarely be covered by a single standard and not always even by the scope of a single SDO [138]. Furthermore, Updegrove goes on to conclude that increasingly complex collections of standards developed by several SDOs have to be dealt with. This has in turn led to the situation where we have a large number of SDOs, consortia, and fora with complex inter-relations forming a complicated web with many overlapping activities [87]. Also see Figure 1 for an overview of the current ICT standards setting bodies and their inter-relations as presented by Jakobs [87].

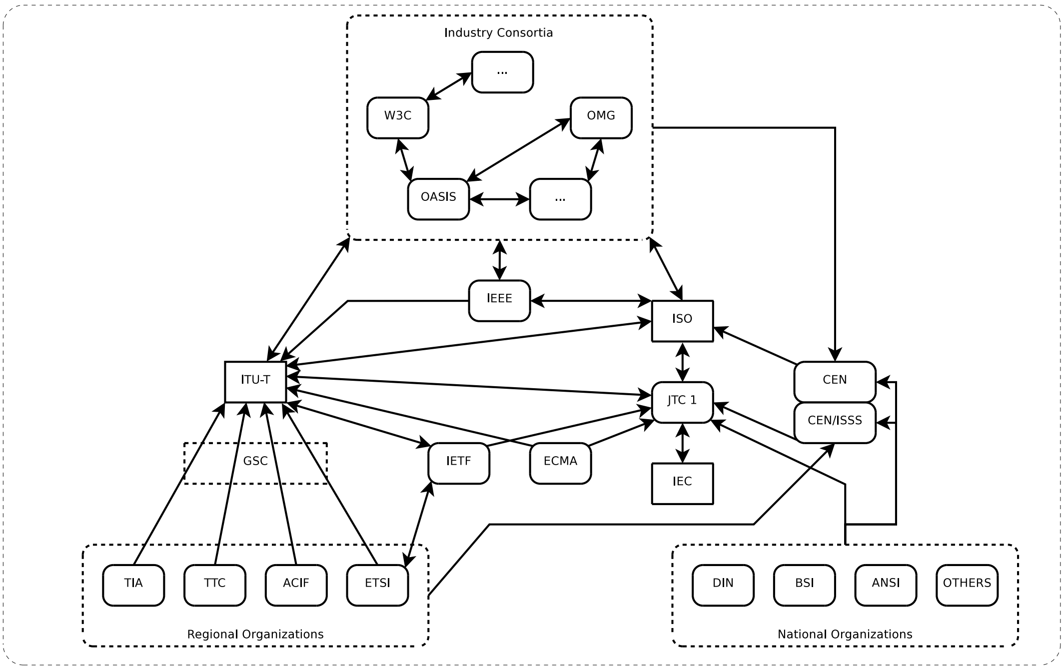


Figure 1: An excerpt of the ICT standardization universe [87].

The considerable variety of SDOs in the industry has also affected the way standards are developed. Vercoulen and van Wegberg analyze the selection process of standards [145]. They elaborate on weather standards are or should be selected and developed by a purely market driven or by a purely negotiated process. They conclude that in an environment like the ICT domain, characterized by dynamics, complexity, and a dispersed web of SDOs, consortia and fora, the standards selection process will most likely be a hybrid between purely market driven and purely negotiated processes. In the aforementioned configuration both negotiation and market competition will play a part and situations where a trade-off between the compatibility qualities of a standard and the time it takes to develop the standard may well arise [144].

The international and networked nature of the ICT industry gives rise to some more specific attributes of the standards development process. Funk argues that the success of firms is strongly related to the evolution of standards. Funk states that, in networked markets an early and large installed base resulting from an early domestic standard will positively influence the likelihood of the standard becoming adopted world-wide [49]. The balance between co-operation and competition, de facto and de jure standardization, and timeliness are key values in creating successful ICT standards.

Finally, accounting for the horizontal convergence, i.e. the far-reaching nature of ICT technology into all parts of everyday life, gives the standardization some further distinctions. The European Information and Communications Technologies Standards Board (ICTSB) discusses the importance of standards and interoperable infrastructure on the way towards a fully connected Information Society [69]. The International Electrotechnical Commission (IEC), ISO, ITU-T and the United Nations Economic Commission for Europe (UN/ECE) go on to propose that standards are not only a technical issue, on the contrary, ICT standards are a fundamental tool for the success of the Information Society in its entirety [71].

Governmental involvement represents another interesting aspect of ICT standardization. Governments tend to have strong policy interests in ICT standards, but at the same time governments are also end users of standards. Especially Internet standardization has a close relationship with the U.S. government, as early IETF standardization activities were carried out by government-funded researchers [55]. In Europe, the EU commission's interest in ICT standardization is divided into three categories: standardization in support of regulation/legislation; Standardization in support of EU policies, not being embedded in legal frameworks; Standardization policy in support of the competitiveness of EU industry [68]. Optionally, governments may also deal with antitrust liabilities in some cases [105].

2.4 ICT Standardization's Link to Innovation and R&D

Standards development is closely connected to R&D and innovation. Interfacing between the two bring mutual benefits. Standardization is an effective way for research projects to reach the industry, users, consumers and other stakeholders, and in that way disseminate their results competently. Standards development organizations in turn can gain benefit from receiving prevalent information from research projects rapidly.

Standards and standardization work provide new information and can thereby have a considerable role in motivating knowledge intensive activities such as R&D and innovation [136]. There are several motivations for participating in standards development, it is commonplace to participate in order guide the process and in that way influence the outcome of the standard. However, regarding R&D, participation can be motivated by being there to learn [130]. In relation to this, Cohen and Levinthal discussed the dual nature of R&D [28]. They concluded that investment in R&D is not only motivated by a direct pursuit of new process and technology innovation, but also by the fact that R&D enhances a firms ability to embrace and exploit externally available information. E.g. identifying and exploiting external information available in a standards development organization can give a R&D project early access to new technologies. See Figure 2 for a clarification on the information flow between R&D and standardization.

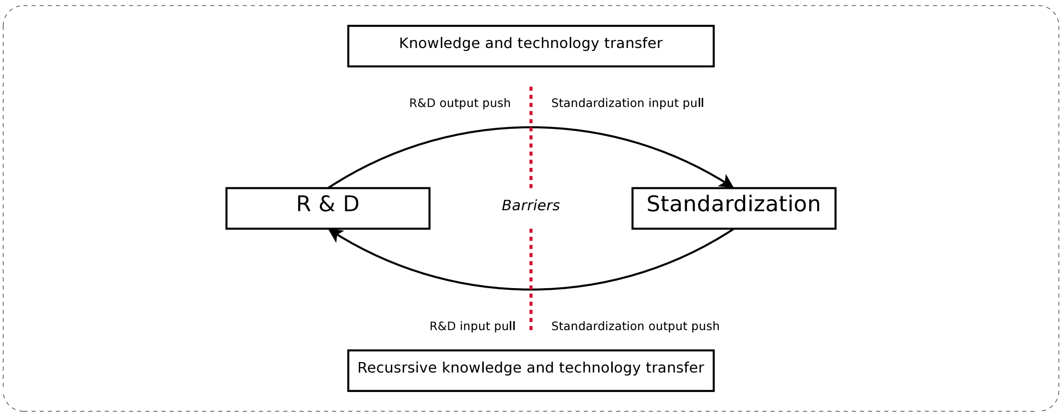


Figure 2: The relationship between R&D and standardization [78].

Information provided from standards and standardization work also bring significant value to R&D and innovation by reducing uncertainty and thus reducing risks [56]. The DIN report argue that by participating in standardization, organizations can in addition to reducing their economic risk, also decrease their R&D costs as results are distributed among them and research is therefore not duplicated [38].

A well-designed interface between R&D and standardization will make standards available earlier and thus bring benefits to both the industry and the public [16]. The Information Society Technologies Advisory Group's (ISTAG) report reasons that moving quicker on standards is essential to systems interoperability and user adoption, and that standards should hence be among the targets of R&D efforts [77].

In Europe the EU has aimed at creating a systematic link between R&D and standardization [143]. There are both ongoing projects and previous success stories. The success of the GSM mobile communications system as a global standard is a good example of a project where the interface and interdependence of R&D and standardization was ably taken into account and put to use. From the start, in 1982, two objectives were identified as paramount to GSM, a pan-European effort was required to obtain *critical mass* in the production of equipment, and secondly, at the same time, competitive supply could develop as the system was to be well specified and open [114]. Programs such as the Research in Advanced Communications in Europe (RACE) and Advanced Communications, Technologies and Services (ACTS) provided important contributions to standardization [95] [117]. Pelkmans summarizes that the openness of the GSM network compatibility standard, its non-proprietary nature, and a measurable form of private and public co-operation to create investor confidence, surely was among the crucial elements in its success [114].

Furthermore, in the EU a special project was run in order to improve the interface between research and standardization. The Co-operation Platform for Research and Standards (COPRAS) project concluded in their guideline report that standardization work provide the opportunity to create exposure among a huge community of external experts. Interfacing with standards development may also lead to technologies developed by a project being successfully embedded in future standards [29]. The aim is thus to enable easy transfer of R&D results to the standardization field and subsequently also to identify the need for R&D consortia to accompany the importance of this issue [41]. COPRAS has also published a thorough guideline document titled *Standardization Guidelines for IST research projects interfacing with ICT standards organizations*, to assist researchers in interfacing with standardization [29].

Other significant European projects focusing on the link between R&D and standardization are the CEN Standardization and Research (CEN/STAR) and the Integrating Research and Standardization (INTEREST) programs. The CEN/STAR project is in place to develop a more efficient link between European Co-operative R&D and European standardization [43]. They focus on both co-normative research that interacts directly with ongoing standardization and pre-normative research relating to activities which are likely to generate issues needing standardization in the future. The INTEREST project on the other hand aims to identify the relevant dimensions to be considered to successfully integrate research and standardization [143].

Gauch also elaborates on the issue, that these programs should distinguish and communicate the incentives, benefits, and the strategic importance of standardization to the researchers [51]. To contribute to the issue, the INTEREST project has also published a comprehensive guide titled *A Guide to Standardization for R&D Organizations and Researchers* [78].

A final extreme version of the interface between standardization and R&D can be noted in the Open Source movement. There co-operative models have been established allowing firms to combine their R&D efforts to develop a common technology and business version that all the participants will support [19].

Standards will furthermore help to increase innovation by enabling companies to create on top of existing specifications rather than spending valuable resources developing a proprietary specification on their own. Standardized technologies also work as great equalizers between big vendors and smaller players, the entry barrier to the market is lowered when the core technologies are standardized and therefore also smaller players can enter the market and succeed. [89]

2.5 Venues for Developing Standards

Developing standards is an intricate undertaking. The process involves organizational rules, legal challenges, and interests of different stakeholders. There are also several different types of venues where standards are developed. In this thesis we divide these venues into five different categories based on their corresponding processes. According to Cargill the categories are: trade associations, national formal organizations, international formal organizations, consortia, and alliances [17]. In this section the categories will be presented in a chronological order. Cargill also argues that each category arose as a response to a specific need in an evolving industry value chain. In line with Cargill and Bolin we further divide the field into a formal arena including trade associations and national/international formal organizations, and an informal arena consisting of consortia, fora, and alliances [19]. The formal standard development organizations enjoy some kind of official accreditation, whereas the informal counterparts are more market driven and in most cases put in place by industry stakeholders.

First we take a closer look at the formal standard developing organizations. According to the Organization for Economic Co-operation and Development (OECD), the basic principles that underlie the formal process are (1) open and broad representation of stakeholders; (2) due process to all participants; (3) provision for public scrutiny and comment; and (4) decision-making by consensus [113]. The knotty rules of the formal process can however, from the ICT industry's point of view, lead to excess inertia in the standardization process. The open and broad representation of stakeholders with possibly differing interests can make consensus reaching a slow process. Farrell and Saloner stated that the negotiation process can occasionally be described as a war-of-attrition between the participants [46]. Reaching consensus can get troublesome, especially when stakeholders have vested interests. E.g. Kolodziej refers to a dispute between AMP, Inc. and AT&T regarding Fiber Distributed Data Interface (FDDI) connectors, delaying the IEEE consensus reaching process on a 802 standard by a year and a half [94]. Provision for public comment and various appeal processes can also prolong the standardization process.

Formal SDOs are typically multi-disciplinary, on a national level a single SDO can be responsible for standardization across all industry lines and international SDOs are also often specialized on a whole industry line [35]. Their processes are designed to cope with very varying issues and therefore the processes may fit the specific needs of the ICT standardization field poorly. Formal SDOs have however become conscious of the value of timeliness and dynamics. E.g. the ISO/IEC procedures now include a fast-track where a standard developed in some other SDO can achieve the ISO *International Standard* status more quickly [70]. Furthermore, van Wegberg argues that the proliferation of competing consortia and alliances, resulting from a possible slow down of standardization in a grand coalition type of formal organization, may have an adverse effect on the degree of compatibility and technical superiority of the standards [144]. There is fundamentally a trade off between the greater compatibility achieved by the scrutiny of the formal process and the agility and speed of the more focused standardization work in the consortia and alliances. See Figure 3 for a clarification of this relationship as presented by van Wegberg.

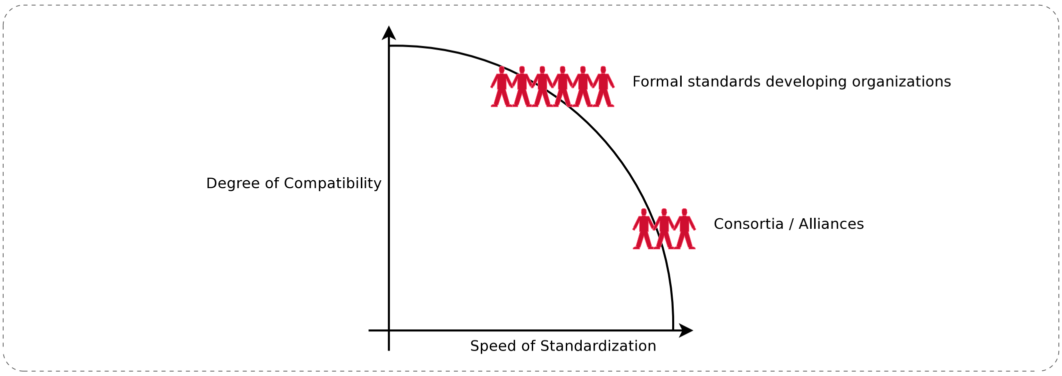


Figure 3: Size of the coalition versus the time needed to reach consensus. [144]

Some formal standards development organizations and trade associations developing ICT standards are the Institute of Electrical and Electronics Engineers (IEEE), the International Organization for Standards (ISO), the International Telecommunication Union (ITU), and the European Telecommunication Standards Institute (ETSI). Their respective most relevant ICT related work is the 802 Local Area Network (LAN) and Wireless Local Area Network (WLAN) standards by IEEE, the Open Systems Interconnection Model (OSI) by ISO, the International Mobile Telecommunications-2000 (IMT-2000), ISDN, DSL, H.323 multimedia and VoIP recommendations by ITU, and European radio bandwidth regulatory work and developing the GSM, TETRA, and xDSL standards by ETSI.

Today, much of the key standardization work in the ICT field is currently carried out by industry consortia and alliances, rather than in formal international or national standards organizations. The complex and dynamic nature of the ICT industry with fast innovation and short product life cycles are explanatory factors for this diversification of the standardization field [87]. The industry consortia and alliances are seen as more dynamic, focused, and business-oriented than the formal SDOs. Updegrove states that in 2007 there were in total more than 500 informal ICT standardization organizations in operation [138]. These consortia and alliances range from small, closed vendor clubs to very large, institutionalized, global organizations. Some operate on an invitation only basis while some are completely open organizations. Common to these consortia and alliances is that they can be described as collections of like-minded stakeholders who come together to act as advocates for some distinctive issue [19]. Because consortia and alliances usually are more focused and consist of groups of congenial participants, reaching consensus can also be much faster.

Ray Alderman, executive director of the VMEbus International Trade Association (VITA), reasons that consortia and alliances could completely displace formal SDOs, as the formal arena appears to value the process more than the results, whereas the informal counterparts are less process and more result oriented [10]. A classic example where the process became the rationale for standardization and thereby partially accounting for the failure of the standard is the ISO OSI model [18]. On the other hand we have the success of the Internet and the Internet Protocols, accompanied by the more simple, result oriented, standards development process conducted by the IETF [52].

The European Committee for Standardization (CEN) maintains a comprehensive list of ICT standards consortia [42]. CEN lists almost 240 consortia and alliances that fulfill the following requirements: (1) are international in outlook and scope, (2) have active and international membership, (3) work is of importance to the areas of ICT standardization or its processes. The proprietary, single-vendor, groups are omitted. Most of the consortia involved in Internet related standardization fulfill these criteria. In the context of this thesis, the IETF is however chosen as the most relevant consortia for further analysis.

The IETF is an independent global consortium and they are generally responsible for standardizing the Internet's technical architecture. The IETF is described as the archetypal consortium in the ICT standardization field [19]. The IETF standardization process is fully open and has been designed to provide swift solutions to immediate problems [86]. Participation is open to any interested individual, specifications are openly available at no cost throughout the process, participation is virtually free as e-mail distribution lists are used for discussions. Demonstrated interoperability is also required for all different technical implementations [85]. Jakobs concludes that the unprecedented importance of the Internet in today's economy also puts the IETF in a special role compared to other standard developing organizations [87]. Jakobs further elaborates that the success of Internet standardization has led to many other large consortia, such as the World Wide Web Consortium (W3C) and the Organization for the Advancement of Structured Information Standards (OASIS), have chosen to base their process on that of the IETF. Rutkowski describes the Internet standards development process as by far the best in the business [122]. He also goes on to illustrate it as more than just a standards process, but rather as a distributed collaboration and innovation engine. The IETF and its working processes are further discussed in Chapter 3 in this thesis.

Fast technological change and the ever increasing diversity in the ICT domain has also resulted in an increased number of standardization activities. Increased standardization activities and a larger number of standards developing organization thereby also imply more complex relations between the venues where standards are developed. The current webbing of liaison agreements and co-operation arrangements between SDOs, consortia, and fora can be considered complex. This situation may also lead to reduced transparency for companies and stakeholders interested in active involvement. [8]

Competition between different standards development entities in the ICT area has also been noticed. Jakobs identify situations where both consortium vs. consortium and consortium vs. SDO competition occur. He further states that actions should be taken to avert competition in standards development and instead improve co-operation and co-ordination. [87]

2.6 Participating in ICT Standardization

Developing standards in the ICT domain is a co-operative and competitive effort of several industry stakeholders. Formal SDOs and several consortia speak for open and broad participation of all stakeholders, whereas, on the other hand, some consortia and alliances have closed fee-based membership policies. Whatever the participation policy might be, it is commonly noted, that for many interest groups there are significant benefits of participating in standardization, these benefits are far reaching, accumulating not only to specific participants, but also to the broader economy as a whole [56]. Participants in standards development benefit from joint learning, but their social and political abilities are also important in the standardization game [100].

To be able to understand participation in standards development better, we try to present a taxonomy for the stakeholders involved in the process. A stakeholder is in this context defined as, any group or individual who can affect or is affected by the achievements of the standard development organization's objectives [48]. There is no single comprehensive typology for the participants and stakeholders in the standardization process. In this text we will therefore elaborate on a taxonomy based on the stakeholder theory presented by Mitchell, Agle, and Wood, that was put into the standardization context by de Vries, Verheul, Willemse and modified by Updegrove [108] [36] [137]. The taxonomy for the stakeholders in standards development is defined as follows:

- Vendors
- Commercial end users
- Government entities
- Individual practitioners of the technology in question
- Universities and academics
- Individual consumers

Each of these categories have different motivations for engaging in standardization work, and therefore also the level of participation varies between these groups. De Vries, Verheul, Willemse identify the importance of open and balanced stakeholder participation in the ICT standardization process [36]. They also go on to conclude that the stakeholder representation is in practice often very unbalanced, regardless of the specifically important issue of broad stakeholder involvement based on the deep impact of ICT standards on many of the stakeholder groups. Bolin further argues that the best possible standards emerge by gaining consensus among an open and broad stakeholder representation [10]. In an empirical study by Chiao, Lerner, and Tirole, the authors conclude that among the 60 SDOs they analyzed, more than half of the standards development organizations (57%) consist of corporations only. 8% consist of both individuals and corporations, and 25% consist of corporations and others. Only one organization consisted of all three types of stakeholders, however almost all of the selected organizations (92%) had corporate members. [25] In this section we will now examine more closely the interests and motivations for participation of each identified stakeholder group.

The first group, the Vendors are the ones with the most at stake regarding participating in standards development. Vendors have the most to gain, and vice versa also the most to lose. Funk and Methe argues that firms who create the winning standards reap substantial rewards compared to firms that back unsuccessful standards [50]. Vendors thus have the largest incentive to dedicate the resources needed to exercise influence on standards development [137]. Vendors and other commercial entities therefore, for the most part, form the largest and most economically influential group of members in most SDOs, consortia and fora [139]. Proactive participation can involve both blocking and undermining competitors contributions as well as supporting and promoting one's own assets. Leiponen has analyzed the competitive and co-operative aspect of firms participating in standardization and she concludes that firms are advised to engage in a broad co-operative approach if they wish to actively contribute to, and align standardization outcomes in a, for them, favorable way [100]. The vendors are also the most active participants in the standardization process. E.g. in a study of ISO, ITU and IETF working groups, almost 60% of all participants represented product vendors or service providers [88]. Vendors are also the stakeholders most likely to have vested interests in technologies being standardized. Strong vested interests can lead to prolonged wars-of-attrition between participants in a SDO [44]. The issue of membership fees is also a concrete factor affecting participation in standards development. Leiponen concludes that it is important to ensure that SDOs remain open for all industry actors and that membership fees do not become prohibitive to small and resource-constraint players [100]. Bolin concludes that from a vendor's point of view, participating in open standards development boils down to being asked to co-operate with competing vendors, sharing information with them, and even donating intellectual property for the good of the market or for the sake of technology, ultimately however increasing the overall market size and uptake [10].

Commercial end-users represent the second stakeholder group involved in standards development. This group does not participate in standards setting as actively as the vendors, but as these corporate purchasers often make significant investments in standards-based products, they therefore also have a clear incentive to take part in the standards development process. These stakeholders can also reap profit from participating by influencing vendors to shape coming products towards their own needs, and by doing so be able to implement innovative solutions ahead of their competitors [10]. From an SDO's point of view, commercial end-users provide value to their process by contributing real-time market input regarding customers' requirements, they also represent the *first adopters* and thus provide merit for possible future wider implementation of a standard [137]. End-users also have a major role in the innovation process linked to standards development as they seldom have vested interests in technologies being standardized and thus are more likely to vote for the best solution based on purely its technological excellence [84]. By participating, commercial end-users can also evaluate the value of a given standard to them and possibly chart the road towards adopting the standard in the future.

Governments and government entities participate in standards development for several different reasons. Swann identifies two idealistic motivations for governments to engage in standards development [130]. The first motivation is to correct the typical imbalance in participation. Vendors are most often over represented in proportion to the other stakeholders. Swann argues that an imbalance in participation might cause standardization to be shortsighted. Governments should therefore participate and help in ensuring the representation of the lesser stakeholders. The second motivation for government participation is to keep the standards infrastructure in good working order. Governments want to prevent situations where e.g. markets fail to produce working standards. In addition to the aforementioned motivations Updegrove states that the interests of government members are also in many cases similar to those of other end-users [137]. This derives from the fact that many government entities also represent large purchasers of standardized products. Jakobs also notes the patriotic aspect of government involvement, he argues that a national standard introduced into the global domain will sustain the prospects of the domestic economy it originated from [85]. There are also risks involved in government participation. The governmental technical expertise might sometimes be vague, but the impact of government intervention can anyhow have a large consequence on the outcome of which technology becomes the standard [56]. Finally, we note the fundamental difference in the U.S. and the EU approach to government participation. The U.S. tend to be timid against government involvement and embrace market determinism, while Europeans on the other hand tend to trust the government more than the market [10].

From governmental stakeholders we move on to academic participants. Universities and academics have been involved in Internet and ICT standards development since the beginning. This stakeholder group is most likely to participate in SDOs with open membership, universities and academia are not likely to pay high-priced memberships fees when the standards development activities are being conducted in proprietary consortia. From an SDO's point of view, participation of university and academic members can be very attractive in the form of their level of expertise, credibility, and other valuable resources they bring to the process. [137] Hence, some alliances and consortia have lower membership fees for universities and other academia.

Consumers and end-users represent a stakeholder group which differs quite substantially from the aforementioned. De Vries, Verheul, Willemse argue that especially in the IT sector, but also in other sectors, end users play a major role in standards development [36]. The central issue involves improving the transfer of end-users' needs to the SDOs. Updegrove furthermore gives the end-users and consumers special attention since they are the stakeholders least able to adapt, work around, and supplement impediments and inadequacies in the standards which they have to work with [137]. End-user and consumer participation may also bring value to other stakeholder groups. E.g. vendor participants may benefit by learning about end-user requirements early on in the standards development process. Bolin writes that some consortia have successfully been able to produce standards more rapidly when they have adopted a requirements-based process that places users in the middle [10]. However, there currently seems to be a lack of user participation in standards development.

Not all stakeholders choose to participate in standards development. The final group we want to mention here are the non-participants, also called the free-riders. According to Blind, there are generally significant incentives for companies and other stakeholders to participate in standardization processes, which are generally higher than the incentive to behave as a free-rider [7]. Open standards can however show many attributes of public goods and may therefore also offer benefits to non-participants and possibly to the society as a whole. Weiss and Toyofuku analyzed free-ridership in the 10BaseT (Twisted Pair Ethernet) standards setting process. They isolated both the costs and the benefits of free-ridership. Among benefits are savings on the costs, or no costs at all of developing standards. There is also the opportunity to learn about the market and compatible products as they develop, before making any own commitments, thus lowering the risks of entering a market [148]. On the other hand, Weiss and Toyofuku concluded, that free-riders may see losses as they will likely enter the market later and they are also possibly collecting additional costs of having to create something more innovative compared to the early entrants' products.

So far in this section we have listed motivations and benefits for several stakeholder groups participating in standards development. In practice however, most SDOs, consortia, and fora accept only organizations, including commercial, academic, and governmental entities as members. In some SDOs, on the other hand, membership is also accepted the individual level. Examples of established SDOs in the ICT field accepting individuals as members are the IEEE and the IETF [72] [60]. The IETF actually lacks formal membership altogether, anyone can register for a meeting or join a working group mailing list. The principle of openness is most often the primary reason a SDO will accept individuals as members [139]. Some consortia, such as the IEEE, have a dualistic role of being, on one hand, a standards developing organization and on the other hand also a professional association, therefore being obliged to accept also individuals as members. Motivations for individuals to participate in standardization work are diverse. Commonly these motivations can however be divided into two categories, participating on behest of ones employer or participating for a personal reason. When individuals participate on behalf of their employers, it is commonly expected that their goals are largely aligned with those who finance their participation [137]. Personal reasons for participation can be out of pure professional interest, include personal education, to make a difference within one's area of expertise, or to just network and get connected with other domain experts [139]. The IETF is a prime example of an SDO where, at least in its first decades, experts participated as objective individual experts driving standardization by purely technological considerations, regardless of their corporate affiliations [101]. The concept of standardization *gurus*, that is people who have been involved since the earliest stages, is present especially in the Internet standardization domain with its academic roots [88]. Many of the gurus then benefit from their vast domain experience by running own consultancy firms. Some SDOs are also successfully incorporating *a one member* - *one vote* rule, where participating individuals can truly influence the outcome.

2.7 Challenges Related to ICT Standardization

ICT standards setting is not a straightforward task. The standardization work involves many difficult aspects from a participants point of view. Both technical and market related problems are also discussed in many articles. A great part of the challenges included in ICT standardization are related to the fast pace the industry evolves.

In their article, *Standardization: A Failing Paradigm*, Cargill and Bolin focus especially on problems around the excessive increase in the number of standards and standards development organizations [19]. They conclude that the rapidly increasing amount of SDOs, fora, and consortia have its roots in the technology boom, a time when standardization increased in popularity. In that period, new entrants saw standards development as a way to compete on the market. If there was not any existing SDO suitable for a specific standardization need, a new SDO was swiftly chartered to better suit the specific preferences at the time. Cargill and Bolin also noted that some companies which might have been in danger of loosing ground on the market if any standard in their domain was successfully developed, started their own standardization efforts just to produce a competing standard, thereby consequently fragmenting the market. The authors furthermore conclude that if this unmitigated output of standards, especially competing standards continues, the market will fragment to the point where interoperability will become impossible and the needs that are being met are not technical, but rather fulfill the providers' market-positioning requirements [19]. As an example Egyedi bring up two failed attempts by Sun Microsystems to standardize Java, ISO JTC1 standardization failed in 1997 and standardization in the European Association for Standardizing Information and Communication Systems (ECMA) failed in 1999 [39]. He concludes that Sun Microsystems' motives were not technically but market oriented and thus a solid precondition for successful standardization was completely lacking.

The fact that ICT technology and the ICT domain itself tend to evolve very swiftly put further stress on the standardization process. Product life-cycles shrink and the standardization cycle can be longer than the product life-cycle. Therefore speed is an important dimension of SDO performance and participants see excessive delays as real problems. Several papers cover the problem of delays and the comparative slowness in standardization processes. We have reviewed papers i.a. by Simcoe, Farrell, Wegberg, and Farrell and Saloner [127] [44] [144] [46]. Simcoe analyzed the slowdown of Internet standards development. He studied the performance of the IETF and focused on the time required for committees or Working Groups to reach consensus. Simcoe furthermore stated that the cost of delay often have to be weighed against the goal of creating a high-quality specification, and promoting its widespread adoption. Farrell analyzed the process of reaching consensus in standards development and concluded that friction is unavoidable when competing firms need to agree on technical details [44]. Farrell also goes on to describe formal standardization as a war-of-attrition, which in turn suggests that delays are expected in reaching consensus. Kolodziej conclude that the political aspect of standardization can become especially entangled when vendors already have a vested interest in the technology being standardized [94]. Some delays for developing standards are listed in Table 1. Further analysis on the IETF delays is presented in Section 5.2.

Table 1: Delays for standards. [127] [44]

SDO	**Average delay**	**Year**
IEEE	7 years	1981
ISO	6-7 years	1987-1991
IEC	5 years	1994
IETF	2 years	2000

The fragmented nature of the ICT standardization scene further add to the challenges. Wegberg presents some reasons behind this phenomenon [144]. His paper argues that having large, industry-wide standardization coalitions, will slow down the standards development in comparison to having smaller, competing, coalitions. Decision making can easily be stalled by competitors with diverging interests politicizing the standardization process, and in that way holding the SDO to ransom in order to get their preferred technology selected as the standard. Under the circumstances where standards are developed in several competing consortia or fora, can in turn lead to lesser interoperability and compatibility. Furthermore the problem of complex relations between SDOs, consortia and fora becomes apparent. Wegberg finally conclude that there basically is a trade off between speed of the process and the level of compatibility achieved in the standardization process [144].

The diversity, and perhaps overlapping scopes, of different standard setting organizations present even more tribulations for ICT standardization. E.g. standardizing a technology such as Push To Talk Over Cellular (PoC) needs to engage several SDOs. Ali-Vehmas lists the 3^{rd} Generation Partnership Project (3GPP), 3GPP2, the Open Mobile Alliance (OMA), and IETF as principal bodies relevant to PoC standardization [1]. Each SDO essentially has a clear mandate, but in practice the work plans are not fully inline and there are still several complicated areas. E.g. co-operational agreements and the definition of scope for the involved SDOs need to be sorted out.

Vinaja and Raisinghani also bring up the challenge related to governmental issues [146]. National interests might place a standardization effort on hold up for prolonged periods of time and the outcome might be that different countries end up with non-interoperable solutions. Major obstacles might also rise from differences in governments' and SDOs' preferences, as standards often have impact on regulative legislation and vice-versa.

Some authors have also analyzed the hurdles related to the co-operative and competitive arrangements put up by the participants in the standardization process. Jakobs, Procter, and Williams looked, among other things, at user participation in the standardization process. They found this to be a somewhat controversial issue. They established the general opinion to be that user participation could be a cure-all for many problems, however, at the same time several participants expressed differentiated and often opposed views. A closed circle of technical experts is according to the authors more likely to co-operate smoothly on a technical problem without bothersome interference from *outsiders* [88]. Leiponen analyzed private and semi-public co-operative arrangements put in place by participants both inside and outside an SDO. More specifically she also analyzed participants' abilities to influence technical standardization through these co-operational arrangements and her results pointed to that they are significant [100]. Some co-operational systems might also be put in place with the goal to exclude certain competitors from the market.

The complex environment outlined above might present a major obstacle for those who are considering active participation in ICT standardization, as well as for those who are looking for a standard that best suits their needs. Jakobs conclude that when considering this complexity of the ICT standardization universe, *Where to participate?* really becomes a relevant question. [87]

3 Internet Engineering Task Force

This chapter introduces the Internet Engineering Task Force to the reader. By taking a look at the birth of the Internet itself we establish where the IETF has it roots. We also cover the inner beings of the IETF, namely the field of work, the structure, the members, working methods, and the documents the IETF produce. Using the more generic knowledge of standardization presented in the previous chapter as a basis, we now dig deeper into the venue where Internet standards are created.

3.1 Background

The Internet Engineering Task Force first saw daylight in 1986, however, it was born with a legacy. Everything actually traces back to 1957, when the Soviet Union launched the Sputnik satellite into orbit. As a reaction to Sputnik, which was an unpleasant surprise for the United States, the U.S. Department of Defense (DoD) introduced the Advanced Research Projects Agency (DARPA[2]) in 1958, with the mission to assure that the U.S. maintained a lead in developing state-of-the-art technology for military capabilities [58]. One of ARPA's fields of interest was research on data networks that would enable resource sharing between major computational resources. As a result, the ARPANET project, was launched in 1967. In 1969 the first computer was connected to the ARPANET at the University of California at Los Angeles (UCLA). In the same year Stephen Crocker publishes the first document in the *Request for Comments* (RFC) series, the RFC 1 titled *Host Software* [32]. RFC 1 discusses the host software and initial experiments on the ARPANET network [120]. In the 1970s the idea of open-architecture networking was first introduced by Robert Kahn and a special *Internetting* program at DARPA was initiated [98]. This approach where networks of different kinds could be interconnected using a meta-level architecture laid down the ground for the ARPANET Network Control Protocol (NCP) and would eventually lead to the development of a Transmission Control Program (TCP) by Vinton Cerf and Robert Kahn in 1974 [24]. This allowed for the the original ARPANET to gradually evolve into the Internet.

Inspired by the success of the Department of Defense's ARPANET, the U.S. National Science Foundation (NSF) arranged a meeting, in 1979, to discuss the benefits of computer network services like electronic mail, file transfer, and remote log in that had greatly enhanced research productivity and had generated a strong community spirit among ARPANET sites [37]. The ARPANET was however only available to a very limited community of university laboratories and government institutions. Thus, the NSF started their own Computer Science Network (CSNET) in 1981 to provide similar services and a gateway to the ARPANET available to all computer researchers [112]. A fundamental success factor was that DARPA offered to

[2]Originally ARPA, but in 1972 the name was changed to the Defense Advanced Research Projects Agency (DARPA). In 1993, DARPA was re designated the Advanced Research Projects Agency (ARPA) and in 1996 the name was again changed to the Defense Advanced Research Projects Agency (DARPA).

make its protocol software (TCP/IP) available to the CSNET project, in return to ARPANET being a component of the CSNET.

NSF aimed at further improving the academic computing infrastructure by launching a supercomputer program in 1984 [112]. The goal was to make supercomputers accessible to researchers countrywide. A principal part was the creation of NSFNET in 1986, a high-speed inter-network that would be linked to the ARPANET. The major novelty brought by the NSFNET was that it would be open to all academic users. This, in addition to the critical decision that the ARPANET standard protocol since 1983, the TCP/IP, would be mandatory on the NSFNET, laid the grounds for success [98]. An equally important success factor was that NSF decided on the participation of the private sector in the NSFNET project.

Fundamental to the early and rapid growth of the Internet (ARPANET, CSNET, NSFNET, et al.) was the the free and open access to the basic documents, especially the specifications of the protocols. A key step was the establishment of the Request for Comments series and the publication of RFC 1 by Stephen Crocker in 1969. As the roots of the ARPANET and the Internet was in the university research community it was natural to promote the academic tradition of open publication of ideas and results. [98]

The RFCs played a central role, as they created a feedback loop between the research teams involved in developing the networks and services. Open and free access was provided to the RFCs which in turn fostered the feedback cycle and led to new revisions of the documents being developed. Early RFCs covered topics regarding both the DARPA and NSF networks and technologies. This open dialog between the research community and the open and free access to specification and discussion documents surely contributed to the fact that both DARPA and NSF chose to support and develop interoperable networks and ended up under the same Internet organizational infrastructure.

The growth of the Internet came accompanied by an increased research community interested in its development. Coordination bodies were needed, and in 1979, an informal committee, called the Internet Configuration Control Board (ICCB), was formed by DARPA to guide the technical evolution of the Internet protocols [23]. In 1984, the management of the Internet research program at DARPA initiated a change, the ICCB was disbanded and replaced by the Internet Advisory Board (IAB) [124]. The IAB originally consisted of 10 research task forces and Jon Postel, as RFC editor and *protocol czar*. During 1984 to 1986 the task forces evolved, some were closed and new ones were established. The *Gateway Algorithms Task Force* evolved into the *Gateway Algorithms and Data Structures Task Force* which in turn, in 1986, was split up into the *Internet Architecture Task Force* and *Internet Engineering Task Force* [11]. It is worth noting that, at this point, the IETF was only one task force among many others.

The IETF held its fist meeting in 1986, then under the supervision of the Internet Activities Board. The meeting was attended by 21 people. The top three IETF areas of concern at the time were protocol development and stabilization, protocol conformance, and to be an implementers support organization [55]. In the beginning, the IAB and the task forces were sustained by the US government Federal Research Internet Coordinating Committee (FRICC) and later Federal Networking Committee (FNC) [23]. Research that led to the birth of the Internet had been U.S. government funded all along through e.g. the DoD's DARPA initiative and the NSF's networking projects. The first IETF meeting with non-government vendors attending was the fourth IETF meeting in October, 1986 [57].

The first restructuring of the IAB and the IETF took place in two phases in 1989 when the Distributed System Architecture Board (DSAB) was merged into the IAB and the Internet Research Task Force (IRTF) was founded to consider long-term research problems in the Internet. The structure was also changed to leave only two of the task forces, namely the IETF and the new IRTF operating under the IAB. Several of the other task forces were converted into Research Groups (RG) under the IRTF. Two steering groups were also established between the IAB and the two task forces, the Internet Engineering Steering Group (IESG) and the Internet Research Steering Group (IRSG), to manage IETF and IRTF activities. The IAB was however the final authority for choosing steering group members and setting Internet standards [11]. See Figure 4 for an organizational structure development time line.

Another restructuring was commenced in 1992 when the Internet Society (ISOC) was formed [57]. The IAB charter was revised and the IESG became the final authority for all standards decisions. IAB's activities was also placed under the auspices of the ISOC. This basic organization structure is still valid and will be further discussed in Chapter 3.3.

3.2 Role and Scope

Defining the scope of a standards development organization, consortium or forum is not necessarily an overly laborious task. The comprehensive list of consortia and fora, published regularly by the CEN, is a good place to find basic information about these entities [42]. On the CEN list we can e.g. establish the scope of the Bluetooth Special Interest Group, their objective and scope is the development of the Bluetooth wireless technology. However, formalizing the scope of the IETF is not as straightforward. From the CEN list we learn that the IETF is a large open international community of network designers, operators, vendors, and researchers concerned with the evolution of the Internet architecture and the smooth operation of the Internet. In this section we will take a closer look at what this exactly means.

Bradner summarizes the role and scope of the IETF as being *above the wire and below the application*, w.r.t. layers in the Internet Protocol Suite [12]. However, he goes on to explain that the definition of the wire-layer is getting fuzzy, and mentions technologies like Multi-protocol Label Switching (MPLS), Generalized Multi-Protocol Label Switching (GMPLS) and Virtual Private Networks (VPN), as technologies smudging any crisp borders between the layers. His conclusion is that it is gener-

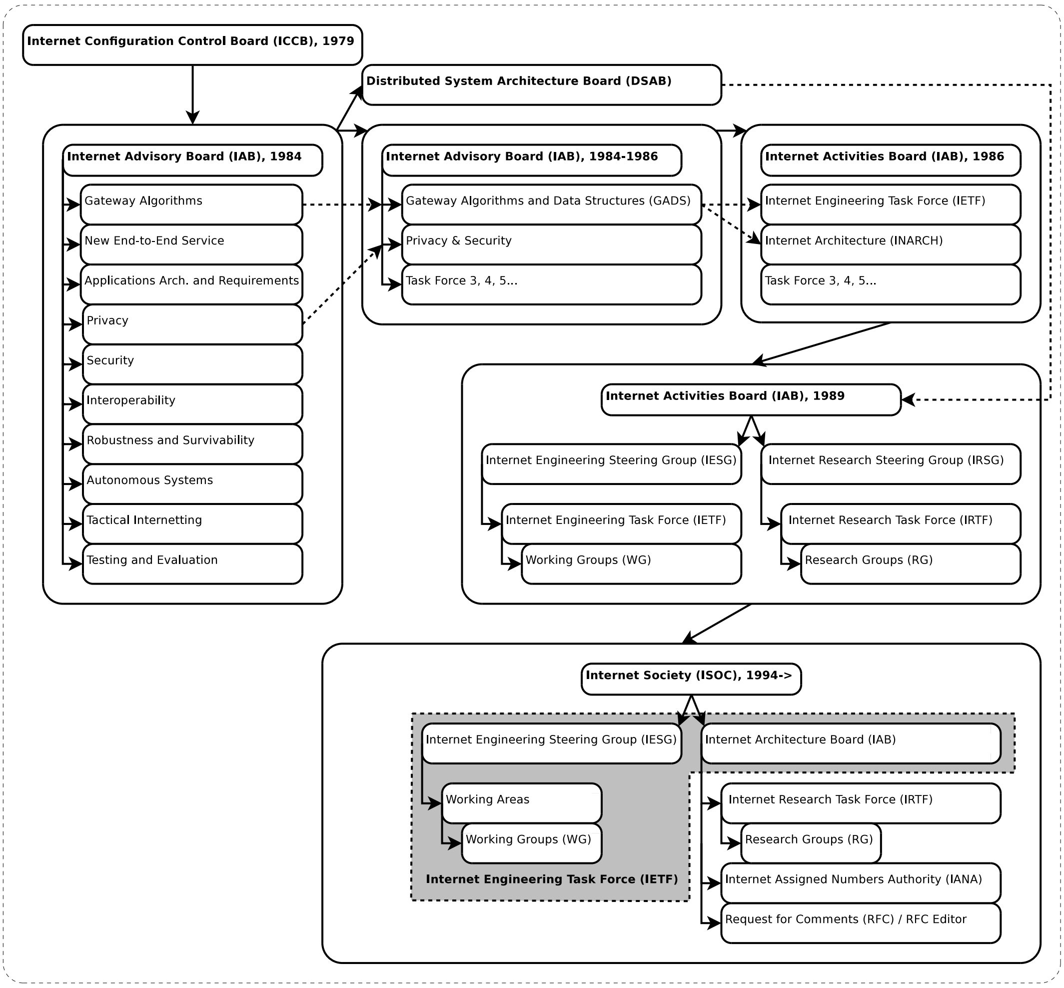


Figure 4: Internet community coordination mechanisms, the birth of the IETF.

ally hard to define the scope of the IETF exactly, and that there is constant edge exploration of the IETF scope.

The goal of the IETF is however concisely stated in RFC 3935 titled *A Mission Statement for the IETF* [2]. It reads, the goal of the IETF is to make the Internet work better. Furthermore, in the Feb 2009 issue of the IETF Journal, IAB chair Olaf Kolkman elaborates on the role of the IETF in a letter to the UN Under-Secretary-General [93] [92]. Kolkman states that the primary function of the IETF is the development, standardization, evolution, and maintenance of the Internet Protocol (IP) suite of technologies. Moreover he also mentions that the IETF conducts its mission with a clear focus on technology and also strives to stay away from policy-making, leaving that to organizations with more expertise in the area.

The scope if the IETF is also impacted by the other SDOs operating in the same domain, often with at least partly converging interests. In parallel with the growing importance of the Internet, the number of other bodies also developing standards for the Internet gas grown. According to RFC 4677, there are also a fair number of standards bodies that ignored the Internet for a long time and now want to get involved in standardizing Internet related technologies [60]. RFC 4677 goes on to document that the IETF in general tries to have cordial relationships with these other significant standards bodies in the domain. This is not always a straightforward task, as many of the other bodies have very different structures compared to the IETF.

Liaison agreements are put into place to manage relations to other SDOs. The IETF maintains several formal liaison relationships with a number of other organizations involved in the development of Internet technologies [93]. IETF liaison management is handled by the Internet Architecture Board (IAB). RFC 4052 describes the best current practice for IETF liaison supervision. The need to engage in direct communication or joint endeavors with various other formal organizations is handled with these agreements [33]. In RFC 4052 it is also stated that IETF liaison relationships are to be kept as informal as possible and that they should carry demonstrable value to the IETF's technical mandate. Successful liaisons prevent duplication of standards development efforts and also provide a framework for information of inter dependencies between IETF's and other SDOs' work. The current map of IETF liaisons can be seen in Figure 5 as published in the IETF Journal.

3.3 Structure

Internet standards are developed in an interesting environment. The Internet Engineering Task Force is not a single organizational entity, it is rather a subset of a larger set of closely connected organizations, societies and groups. A brief introduction to some of these entities and their interrelations was given in Section 3.1 and Figure 4. RFC 2026 The Internet Standards Process is also a good source for basic information about these entities [13]. Further information can be found in RFC 2028 titled *The Organizations Involved in the IETF Standards Process* [61].

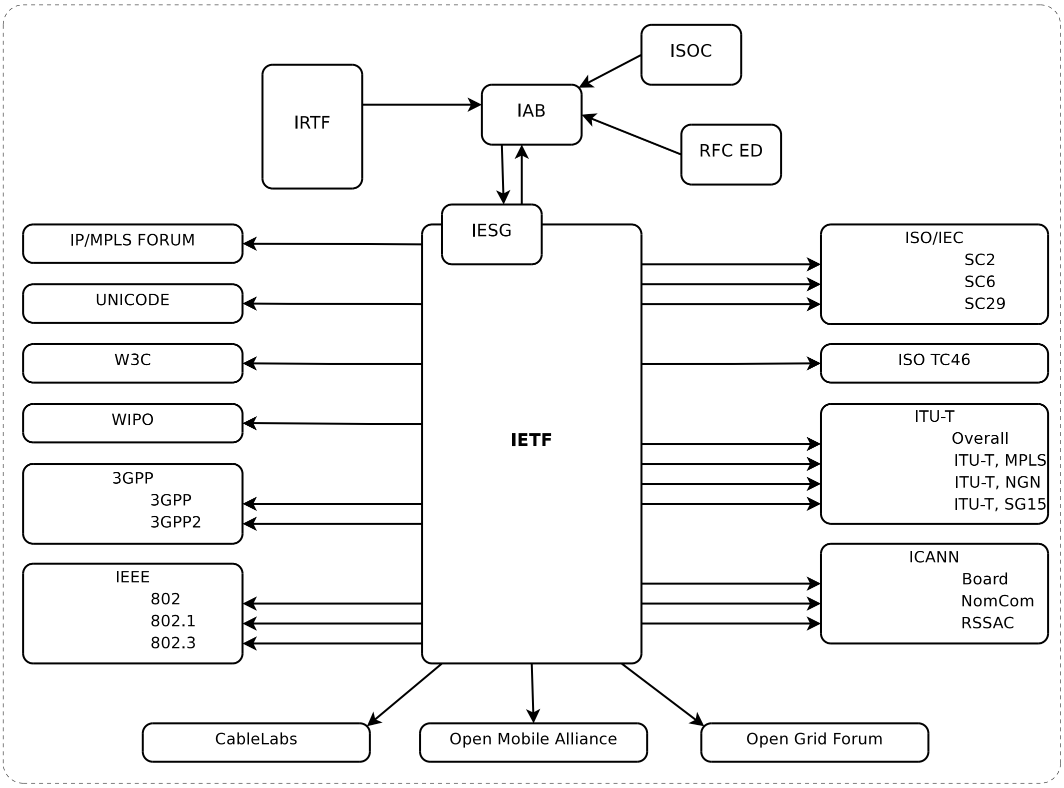


Figure 5: IETF and IAB Liaisons [93].

The structure of the organizations involved in Internet standardization has evolved significantly over time. Initially, in 1979, there was just the Internet Configuration Control Board (ICCB), but today the organizational view is much more checkered. The entities involved are:

- The Internet Society (ISOC)
- Internet Engineering Steering Group (IESG)
- Internet Architecture Board (IAB)
- IETF Administrative functions (IASA) / (IAOC) / (IAD)
- Internet Research Steering Group (IRSG)
- Internet Research Task Force (IRTF)
- Internet Assigned Number Authority (IANA)
- IETF Trust
- IETF Working Areas / Area Directors

The Internet Society (ISOC) is a global nonprofit organization founded in 1992. ISOC's mission is to work for the open development, evolution, and use of the Internet. ISOC works attentively with other Internet organizations, service providers, network operators, root server operators, and other affiliated parties to preserve the elements that have been the foundation for the Internet's success [83]. ISOC's role is also to provide leadership in addressing issues affecting the future of the Internet. ISOC has some 80 organizational and more than 28,000 individual members spread all around the world. One of the drivers behind the founding of the ISOC was to provide support for the IETF. ISOC supports the IETF by functioning as the organizational home of the IETF, the IAB, and the IRTF and also by funding the IETF. IETF's primary source of funding is revenues generated from IETF meetings. Apart from that, ISOC is the only other source of funding for the IETF, the RFC Editor function, the IETF legal support. ISOC itself is funded, among others, by organizational members which include companies whose products and services depend on the standards developed by the IETF. [83]

The actual work in the IETF is carried out in the working groups which in turn are grouped into working areas. The working areas are managed by Area Directors (AD) and the ADs are members of the Internet Engineering Steering Group (IESG). The IESG is responsible for the technical management of IETF activities and the Internet standards process. IESG carries responsibility, and does the final approval, for specifications' progress along the Internet *standards track*. [75] The IESG is also a part of the ISOC and it functions as a source of advice and provides guidance to the ISOC Board of Trustees and Officers, concerning technical, architectural, and procedural subjects relating to the Internet. Another major task for the IAB is to support the IETF and lead the long term planning and defining of the overall architecture of the Internet. Oversight and management of the process by which internet standards are created is also provided by the IAB. [21]

Moreover, there is a support structure called the IETF Administrative Support Activity (IASA), providing administrative support for the IETF technical community. IASA also sits under the organizational umbrella provided by the ISOC. The IETF Administrative Support Activity includes the IETF Administrative Oversight Committee (IAOC) and the IETF Administrative Director (IAD). The IAOC carries out the administrative responsibilities of the IASA and the IAD has the day-to-day operational responsibility of providing fiscal and administrative support [74]. In practice the IASA, IAOC, and the IAD are responsible for preparing an annual budget for the IETF, the negotiation of contracts and other IETF administrative and support agreements, and keeping a detailed public accounting to separately identify all funds available to, and all expenditures relating to the IETF [4]. The IASA however has no authority over the technical work and standards development in the IETF.

The Internet Research Task Force (IRTF) and the Internet Research Steering Group (IRSG) are in place to focus on long-term research related to Internet architecture, technology, protocols, and applications [80]. The IRTF will not commence in standards setting. On the contrary, an IRTF research group is expected to be more long-lived and produce a sequence of results over time. Results of IRTF research groups may however well be used as input for IETF working groups and thereby be brought in for standardization. [147] Organizationally the IRTF sits under the IAB.

The Internet Assigned Numbers Authority (IANA) is also closely related to IETF standards development. The IANA is responsible for co-ordinating the numbering systems used in the technical standards and protocols on the Internet. [64] IANA's day-to-day responsibilities include the assignment of IP Addresses and top-level domain names [115]. We should note that the IANA is only one of several protocol registries operated by the Internet Corporation for Assigned Names and Numbers (ICANN). The IETF - IANA organizational separation can also be considered a unique arrangement, as SDOs generally undertake their protocol parameter registration functions themselves [106]. The IETF - IANA relationship is documented in a Memorandum of Understanding, RFC 2860 [22].

Regardless of the fact that intellectual property considerations related to standards development have been outlined from the scope of this paper, we shortly want to present the IETF Trust. The IETF trust was founded in 2005 and it acts as the administrative custodian of all copyrights and other intellectual property rights relating to the IETF Standards Process. [15]

As mentioned, the technical work in the IETF is divided into Working Areas led by Area Directors. The ADs are responsible for ensuring that the working groups in their area produce coordinated, consistent, and timely output as a contribution to the overall results of the IETF [14]. The Area Directors with the IETF Chair make up the IESG. The IESG and the ADs are in a key position related to the technical work in the IETF as they approve BOFs, new working groups, and make decisions considering the creation or advancement of a specification on the IETF *standards track* [60]. At the time of writing there are eight working areas in the IETF. The working groups are divided into areas as follows: the Applications Area, the General Area, the Internet Area, the Operations and Management Area, the Real-time Applications and Infrastructure Area, the Routing Area, the Security Area, and finally the the Transport Area. There are totally 115 active working groups, divided in to the aforementioned areas.

3.4 Members, Participants, and Stakeholders

The Internet standards process is an open one, and there is no formal membership in the IETF. Any one may participate in the work either on-line or at the meetings. Participation is, as opposed to participation in many other SDOs, at the individual level. Therefore we will in this section look a little closer at who actually participates, and what types of participants there are in the IETF.

Participants are primarily people, not companies. In practice however, almost everyone represent some kind of interest group or organization. RFC 4677 lists some common classifications for IETF participants as follows [60]:

- Network Operators and ISPs

- Networking Hardware and Software Vendors
- Academics

Many SDOs have quite severe membership fees and thus effectively exclude smaller companies, research and education organizations, and personal participation. IETF does not have any membership fees. There is a registration fee when attending a meeting, but remote participation is also possible for people who are financially constrained or if traveling is not an option. Membership fees thus does not exclude participants nor restrict the range of views that can be offered to the standards development process in the IETF [30].

As any other interest group, people representing network operators and ISPs find value in participating, as they are able to shape the outcomes of protocols in accordance to their own preferences. IETF work cover not only network protocols, but also many other parts of ISPs' and operators' businesses. The input of operators is also considered quite valuable to keep IETF work vibrant and relevant [60]. RFC 4677 goes on to state that many of the best operations documents from the IETF come from real-world operators, not vendors and academics.

Networking Hardware and Software Vendors represent the largest interest group behind IETF participants. These entities are also likely to have the largest economic interest in the IETF work. Using the project database we gathered some simple statistics on the proportions of the above-mentioned groups of participants. We looked at the 25 most active organizations counted by RFC authors' affiliations, meeting participants' affiliations, and mailing list activity. The results are presented in Table 2. Looking at the results it is also obvious that most of IETF participants represent an organization from the Networking Hardware and Software Vendors segment.

Table 2: Dividing the top 25 most active organization into groups

Type	**Authors**	**Meet. part.**	**Mailing lists**
Network Hard-/Software Vendors	19	19	18
Network Operators and ISPs	2	4	2
Academics	4	2	5

Academic participants have always been active in IETF work. Initially academic participants were very well represented in the IETF. Their relative share of all IETF participants have however dropped significantly over time. To analyze this aspect, we made a very simple and rough estimate on the relative amount of academic representation in the IETF by counting the relative amount of emails from *.edu* domains on the IETF mailing lists, again using the project database. The results are presented in Table 3. Regardless of the reduced share of academic participants, IETF is still a top place for academics to get information by participating in working groups in their field of interest. Academics and researchers are also well represented in IRTF Research Groups.

Table 3: Percentage of mails from *.edu* domains on the IETF mailing lists

Year	**% of all mails**	**% of .edu+.com mails**
1990	30.14	43.81
1995	12.82	19.87
2000	7.28	10.96
2005	4.85	7.79
2008	2.21	4.32

Finally, to sort out the geographic distribution of the IETF participants, we take a look at the data regarding IETF meeting attendees in our project database. The IETF community initially consisted of researchers and academics from the U.S. Further along the line representatives from commercial organizations also joined in. A majority of them also from the U.S. Today however the IETF community is a truly global one. We use the project database to analyze the geographical distribution of meeting attendees from the three previous IETF meetings. IETF meetings 72, 73, and 74 were attended by people from 61, 65, and 58 different countries respectively. The U.S. had by far the largest share of meeting attendees in the IETF, the percentage share of U.S. attendees in the aforementioned meetings were 40%, 53%, and 58% respectively. Japan and China are the second and third largest IETF countries measured by meeting attendance, their percentage share of IETF meeting attendees are roughly 10% for Japan and 8% for China.

3.5 Work Flow and Working Documents

The IETF lacks an unambiguous document defining its standards process. There are several documents dealing with different parts of the process. Carpenter gives us a guide to these process documents and he draws up a structured way of looking at the them [20]. In this section we will describe the current IETF work flow and the working documents related to it, in a chronological manner, from bringing new work to the IETF, to having an RFC published.

The process of bringing new work to the IETF starts by recognizing a technical problem that is believed to fall within the IETF scope of work. The technical problem might be completely new, or there could already be a solution for the issue. Projects started elsewhere are brought to the IETF the same way as entirely new matters. Deciding if a topic actually belongs to the IETF includes ensuring that there is, or will be, a critical mass of interest in the IETF community to do the work, that the IETF is the most suitable and competent SDO for the topic, and that there is both IETF management support and community consensus to start the work [110].

Rule-of-thumb IETF acceptance criteria for new work are stated by Crocker in his article *Making Standards the IETF Way*. The four criteria to be met are: Competence, Constituency, Coherence, and Consensus. A specification should be technically sound and consistent with the overall Internet architecture. There should be a set of potential providers and of potential users and an indication that they will, in fact, use the services defined by the specification. The specification must also be written clearly and cleanly. Finally, the specification must reflect an adequate consensus of the technical community. [30]

Raising initial community interest includes publishing an Internet-Draft. This is also the first of the two internet standards-related publications. Internet-Drafts are public working documents readily available to the Internet community, thus facilitating the process of review and revision [13]. The first three steps included in the drafting process are documented in RFC 4677 as follows:

1 Publish the document as an Internet Draft.

2 Receive comments on the draft.

3 Edit your draft based on the comments and re-publish.

Finding the correct place for the work within the IETF also needs to be accomplished. Working group charters outline what is in-scope and out-of-scope for the IETF working groups. The charters can be consulted in order to determine, if there already exist a working group suitable for dealing with a specific technical issue. However, if there is no suitable working group, an existing group can be re-chartered or a completely new one can be formed.

A common way to commence new work in the IETF is to have a IETF Birds-of-a-Feather (BOF) session at an IETF meeting. BOF sessions are in many cases organized with the intent to form a WG. The ADs approve BOF sessions and there almost always are one or more Internet Drafts as a base for the BOF. [111] In other words, a BOF session can be described as a meeting which permits *market research* and technical *brainstorming* in the IETF community [14].

Not all work within the IETF is strictly standards development and therefore not all IETF submissions are targeted to become Internet Standards. Thus all RFCs are not Internet Standards, but on the other hand, every Internet Standards is published as an RFC [62]. There are also different requirements for submissions that do aim to become Internet Standards, all six kind of RFCs listed in Figure 6 have different requirements, go through different processes, and have different rankings in the IETF [60].

Informational, Experimental, and Historic RFCs are not considered any kind of standards in the IETF. There are a lot of these non-standards track documents in the RFC series, covering topics from discussion of new research topics to status memos about the Internet [13]. Only RFCs of the Proposed, Draft, and Internet Standards kind are regarded as standards within the IETF.

```
Internet-Draft +---> Expired/Removed
               |
               |
               +---> RFC --+--> Non-Standards Track Maturity Levels
                           |     o "Experimental",
                           |     o "Informational"
                           |     o "Historic"
                           |
                           +--> Standards Track Maturity Levels
                                 o "Proposed Standard"
                                 o "Draft Standard"
                                 o "Internet Standard"
```

Figure 6: Categories for IETF working documents

As mentioned, draft versions of IETF documents are published during the specification development cycle. These Internet-Draft documents are made available for comments and review as they are placed in the Internet-Drafts directory. Internet-Drafts are discussed actively on the IETF mailing lists, in BOF sessions, and Working Group meetings. Internet Drafts expire after a six-month period if it is not replaced by a newer version or recommended for publication as an RFC by the IESG. An Internet-Draft document do not have any formal status at all, and they can be replaced or removed whenever. [13] All documents that become RFCs are initially published as Internet-Drafts.

The first three steps in the drafting process were listed earlier in this section. They are often repeated a few times in order for the IETF community to be satisfied and reach rough consensus on a draft. Subsequently, when a draft document is considered ready for RFC publication, the following steps are to be taken:

4 An Area Director is asked to take the document to the IESG.

5 Changes deemed necessary by the IESG are made.

6 The document is published by the RFC Editor.

The work flow and working documents mentioned in this section will serve as the foundation for our data gathering strategy documented in Section 4.1. There we will look further into the Internet-Drafts, the RFCs, the activity on the IETF-mailing lists, and the participants at the IETF meetings, thus covering all aspects of the day-to-day standardization work within the IETF.

3.6 73^{rd} IETF Meeting

The 73^{rd} IETF meeting was held in Minneapolis, MN, USA in November 2008. As a part of the thesis work, I attended the meeting. The capital goal was to gain an understanding of the current state of the IETF and also learn more about the work and working methods in practice.

To gain an understanding of the current work in the IETF I attended nine working group meetings and BOFs. The meetings were all quite different by nature and in this section we summarize some of the issues dealt with.

The Open Web Authentication (oauth) BOF showed a textbook example of a case where an, already started, external project is brought to the IETF, and development control of the protocol is handed over to the IETF. The authors of the oauth Internet-Draft already had running code and a number of implementations, thus bearing at least some vested interest in the project. This kind of work brought to the IETF is put under public scrutiny from the IETF community and development of the protocol might even be started from a clean slate, disregarding the existing implementations. In this particular case rough consensus was however achieved that the IETF community will start working on oauth and that an incremental improvement cycle was preferred over a clean slate start. There is of course a multitude of reasons why you would want to bring existing work to the IETF, amongst the mentioned was that a published RFC certainly gives the technology credibility especially among governments, institutions and banks.

A returning issue in several of the meetings was a concern of the relevancy of the work. The question, *Who is going to implement this?*, was asked in a majority of the meetings I attended, when discussing different technical solutions. Many meeting chairs also clearly seemed to be aware, and actively brought up this issue. However, already occurred misalignment was also discussed. E.g. in the Session Initiation Protocol (SIP) WG meeting it was noted that the IETF standardization work, and the way the market is implementing the technology, are somewhat conflicting. SIP is a particularly interesting case as, if and when, the market vendors create their own de facto solutions, the whole standards development work can be rendered obsolete. Therefore, the IETF might need to focus even more on aligning their work with the requirements of the market and vendors to avoid situations where proprietary extensions are created on top of a specification. The aforementioned scenario could easily lead to a very complex and possibly un-implementable technical framework.

Another aspect of the standardization work that was discussed during the meeting, especially when working on completely new technologies, was the overall feasibility of the technology. As the IETF states in its mission statement, the goal is to always produce high quality, relevant technical and engineering documents [2]. However, several polls in the meeting rooms on, who of the participants intend to implement the technology at stake, was answered with silence. Besides implementation issues quality matters were also discussed in a some meetings. A situation where a poor quality specification forces the vendors to come up with proprietary bug fixes, thus destroying interoperability was also discussed.

Scope of work related issues were also debated in the meetings. It seemed that it is not always a straightforward task to reach consensus on a topic where the attendees are clearly outside the target users of the technology at hand. Additionally, discussion about how to motivate the IETF community to work on technologies that are relevant but not necessarily close to the preferences of the IETF participants, was also held.

Fast moving technological evolution, a characteristic feature of the ICT and Internet domains was also elaborated upon. Speed was considered somewhat of a necessity also in the standardization work. If the IETF takes too long to create a specification, there is the possibility that vendors on the market has to come up with their own proprietary solutions or some other SDO creates a competing specification. In both cases interoperability will most likely suffer. In addition to being swift in standardization work, it was also discussed that striving for simple solutions is preferable over creating overly complex specifications. One IETF attendee was heard saying that, adding two digits to the page count of a specification takes away two digits from the revenues related to that specification. Added complexity in specifications tend to reduce their adoption.

4 Data Construction

This chapter documents the efforts made to define and gather data and meta data about standardization work in the IETF. We cover what types of data there are available and furthermore also look at how to successfully retrieve and process this data. Finally this chapter also reviews the supposed limitations and reliability of the data we are going to use in our analysis in Chapter 5.

4.1 Data Gathering Strategy

As a part of this thesis we will conduct an analysis on the organizations, companies and people active within the IETF. A primary goal is to measure the participation of different stakeholders through their meeting attendance, mailing list activity and the amount of documents they have authored or co-authored. Readily available data of this kind is scarce or nonexistent, hence we have chosen to collect the data needed for our analysis using our own programming and database tools. In this chapter we will define what types of data and where from it will be gathered.

There are four primary sources of IETF data and in the scope of this thesis we will explore all of them. The sources are the two types of working documents, the RFCs and the Internet-Drafts, the working group mailing list archives and the attendance lists from the IETF meetings. A part of the open process principle stated in the IETF mission statement, is a commitment to make the working documents, the working group mailing lists, the meeting attendance lists, and the meeting minutes publicly available on the Internet [2]. This open availability of the information is a precondition for our data mining task.

Before collecting IETF data from the Internet, we establish some more specific goals on exactly what information we aim to collect from the sources available. Our first priority is to collect data from the IETF working documents, that includes the RFCs and the Internet-Drafts. Information about IETF document authors and their affiliations are considered particularly important, as the RFCs and Internet-Drafts are the primary result and output of all work in the IETF. Both the Internet-Drafts and the RFCs are archived and available. The essential information however lies in the content of the documents, so we need to mine it from there. The types of data and meta-data to be extracted are listed in Table 4. At the time of writing, there are 5334 RFCs and 53911 Internet-Drafts archived in the IETF document repository.

In addition to authoring documents, we can see that much of the work goes on as discussion on the IETF mailing lists. The discussion is open to everyone and all participants can express their opinions. The mailing lists and the discussion there, can be said to represent the IETF at large [60]. Information about IETF mailing list activity is valuable as the majority of new topics and ongoing work is discussed there. The mailing lists are also archived and we set out to collect relevant data from them. Data types to be gathered from the IETF mailing lists are specified in Table 5. At the time of writing, there are 563 mailing lists available in the IETF

Table 4: Working document data types.

IETF document identifier	RFC 4120, draft-ietf-dna-cpl-02...
Document type	RFC, INTERNET-DRAFT
Publication date	1994-12-01, 2008-01-01...
Working group affiliation	sip, avt...
Authors' names	Jane Doe, John Doe...
Authors' organizational affiliations	Cisco, Nokia...
Authors' countries	US, FI...

Table 5: Mailing list data types.

IETF mailing list name	16ng, 6lowpan...
Working group affiliation	16ng, 6lowpan...
Publication date	1994-12-01, 2008-01-01...
Author's name	Jane Doe, John Doe...
Author's organizational affiliation	Cisco, Nokia...
Author's country	US, FI...

archive, containing approximately 1.4 million emails.

The final IETF data to be gathered in this task is information about IETF meeting attendees. There are three meetings per year and the attendee lists are published as a part of the proceedings from each meeting [76]. In addition to the attendees' names, the attendee lists also include the attendees' organizational affiliations or their email addresses. This enables us to collect meeting attendee data that is comparable with the working document and the mailing list data. Meeting attendance data is also relevant as the majority of working groups meet at the IETF meetings. The face-to-face meetings can be considered an important channel to influence IETF work, as all attendees can affect decisions made by taking part in the consensus process [60]. From the IETF proceedings archive we collect the information listed in Table 6. Machine readable meeting proceedings are available from IETF meeting 29 onwards. Attendee lists from IETF meetings 72, 73, and 74 also include country information, thus allowing us to conduct some geographical analyses on the attendees.

Working with the above-mentioned data we are able to quantify the activities and efforts of the people and companies participating in IETF work. Thereby we can draw variable pictures from different angles of what goes on inside the IETF. In addition to the data collected from within the IETF, we will also gather a set of data from outside the IETF. The additional data is however linked to the IETF documents or the companies affiliated with the IETF documents, in such a way that we can classify the working documents by importance and also give the affiliated companies

Table 6: Meeting attendees data types.

IETF meeting number	IETF35, IETF73...
Meeting date	2001-03, 2008-11...
Attendee's name	Jane Doe, John Doe...
Attendee's organizational affiliation	Cisco, Nokia...
Attendee's country	US, FI...

Table 7: Additional data and meta-data.

Market performance	Net sales
Market performance	R&D expenditure
Market performance	Stock performance
RFC significance	Amount of cross-references
RFC significance	Amount of patents issued
RFC significance	Amount of patent applications
RFC significance	Amount of references in academic literature
RFC popularity	Amount of hits using Google search

other external measurable characteristics in addition to their IETF contributions. The types of additional data to be gathered is listed in Table 7.

The IETF data is particularly appealing in this context, as the output of practically all parts of the standardization process are publicly available. The data also covers a significant time span, thereby allowing us to analyze changes and trends over time. The additional data, e.g. patents referring to IETF documents and affiliated companies' stock performances, also allows us to search for connections between activity in the standardization work and corresponding success in other parts of business.

We believe that the IETF data, in accordance with the goals of this thesis, allows us to identify the most influential participants in the standard-setting process, and also establish a benchmark for success in IETF standardization work. Supporting our view is the research by Leiponen on telecommunication standards-setting in the 3GPP, where she concluded that a company authoring change requests and their success in doing so, is one valid measure of their influence in standards development [100].

4.2 Mining the Data

Defining the data to be gathered was a straightforward task. Extracting the actual information from the different sources proved to be a much more challenging endeavor. The primary precondition for this data mining task is evidently that the IETF data is publicly available. The second important precondition is that all the data is in machine readable format. The documents and mailing lists are in the ASCII text format, which can successfully be processed using e.g. PHP scripts. Inside the IETF there has been ongoing discussion about modernizing the format of the IETF documents and mailing lists, but no changes have yet been made. However, as Bradner concludes, even the first RFCs from 1969 are still readable, how many other document formats than ASCII text can achieve that [12].

Our primary source of data are the IETF working documents. Table 4 stated the data types we are to extract from the documents. The bulk of that data resides in the content of the IETF working documents. Hence, our approach is to create a computer program that crawls the document repositories and inserts the documents, as they are, into the project database, thereby enabling us to do further processing on the documents locally. The information we are interested in lies in the working documents' *Author's Address* sections. The section is required in each document as specified in RFC 2223 [116]. In practice the information and presentation on the information in the *Author's Address* section varies considerably. The section is of course written in natural language and is not always structured very consistently, thus it is quite difficult to parse it using computer programs. In Figure 7 an example of a typical *Author's Address* section is presented.

```
Authors' Addresses

  Junghoon Jee (editor)
  ETRI
  161 Gajeong-dong Yuseong-gu
  Daejeon  305-700
  Korea

  Phone: +82 42 860 5126
  EMail: jhjee@etri.re.kr


  Syam Madanapalli
  Ordyn Technologies
  1st Floor, Creator Building, ITPL
  Bangalore - 560066
  India

  EMail: smadanapalli@gmail.com


  Jeff Mandin
  Runcom

  EMail: j_mandin@yahoo.com
```

Figure 7: RFC 5154. Author's Address Section

Additionally we are also gathering data on IETF meeting attendees. Table 6 stated the data types we are to collect about attendees. This information is picked up from the *IETF Online Proceedings* pages [76]. The *IETF Online Proceedings* include lists of meeting attendees. The Proceedings from the IETF 29 meeting onwards are available in electronic form. The most recent attendee lists contain attendee names an affiliations, whereas older attendee lists contain a.o. attendee names and e-mail addresses. The computer script we designed to collect attendee information with, effectively has to cope with five different types of attendee lists to get everything inserted in to the project database. In some cases post processing is needed e.g. to establish the attendees' organizational affiliations using the domain from their e-mail address.

The last repository to be worked through is the IETF mailing list archive. The mass of the IETF work takes form as discussion on the mailing lists. The earliest discussions archived are from 1991. Data types to be collected from the mailing list archives are stated in Table 5. The actual messages lies grouped in separate folders per mailing list, where the messages for each list are grouped by year and month into files conforming to the *Internet Message Format* specified in RFC 822, updated by RFC 2822 and RFC 5322 [31] [118] [119]. To collect the desired data we designed a computer script that processes the archives for both the active and the concluded mailing lists. The script parses the 23199 distinct files containing messages and extracts data from the header fields in each message. Determining the organizational affiliation for the individuals participating in the discussion in done as a separate task using primarily the address and the domain part of the address the mail is sent from.

As a final task we collect some additional data related to the RFCs. Primary sources for this data are the RFCs themselves, the Google Search web page, the Google Scholar web page and the United States Patent and Trademark Office (USPTO) database.

The first aspect of our task to give each RFC an importance factor is to scan through all the RFCs and establish their cross-references. This is done using an automated computer script and the results are inserted into the project database. Before the final analysis, some manual adjustments also have to be made to the cross-reference data. IETF process documents such as RFC 2119 titled *Key words for use in RFCs to Indicate Requirement Levels* are referenced in almost every other RFC, but will nevertheless not be considered more important from the technical standardization point of view we in this thesis are interested in.

The USPTO patent database is used to collect information on both patents and patent applications referencing RFCs. This is accomplished using the USPTO Patent Full-Text database. The full text of all patents issued since 1976, and all patent applications published since March 2001 are searchable in the database [141]. We queried both the issued patents and the patent application databases for each RFC using an automated computer script. The queries were constructed to search for the following types of strings:

```
RFC1112 OR "RFC 1112"
```

for each RFC. The amount of patents and patent applications matching our search string anywhere in the patent text was then inserted into the project database to enable for later analysis.

To add a dimension to our RFC importance classification we also used two Google services to gather additional information about the RFCs. First we queried the Google Search web page with exactly the same type of search string used to query the USPTO patent database. By doing this, we are able to establish a general popularity reading for the RFCs. The Google Scholar service on the other hand enables us to gather data about the amount of times a specific RFC is mentioned in academic literature. Google Scholar documentation states that Google Scholar gives the user a way to search many academic disciplines and sources. Google Scholar material include peer-reviewed papers, theses, books, abstracts and articles, from academic publishers, professional societies, pre-print repositories, universities and other scholarly organizations [53]. The Google scholar service is also queried the same way we searched the USPTO patents database and the amount of search hits are inserted into the project database.

Some further information about a chosen set of organizations participating in IETF work was also collected, thereby enabling us to establish and compare certain aspects of their success on the market. Financial data about the selected organizations was collected primarily from their annual reports. The chosen organizations have their annual reports archived on their web servers. Reports from 1994 to 2007 were collected where applicable. Some of the chosen organizations have been founded post 1994 so the financial data consequently does not reach as far as 1994. The data from the financial statements was again inserted into the project database to enable for further analysis. In addition to data from the organizations' financial statements, we also collected stock performance data. Stock performance data was collected using the *MSN money* service. Stock price history was gathered from 1994 to 2007 and inserted into the project database.

Technical details on the database design, the tools, and the programming are documented in Appendix A.

4.3 Limitations and Reliability of the Data

Establishing the limitations and reliability of the data we have collected is perhaps not entirely trivial. Data is collected from several, quite different, sources and some of the data is embedded in the actual text of IETF working documents.

There are several aspects of the data we have collected that are uncertain. The data we have collected was not originally intended for, nor set up with the goal that it would be easy to insert into a relational database and be analyzed. While building the project database we have had to make several manual enhancements and exceptions to the algorithms parsing the IETF data. Mostly unstructured natural language is, as mentioned, challenging to parse using computer scripts. Therefore the integrity, validity, consistency, and accuracy of the data might also be hard to establish exactly.

The IETF working documents are the primary source of data we explored. They are also the most challenging in respect to establishing an exact quality of the data collected. The data we collect reside in the *Author's Address* section of the documents. The specification for the *Author's Address* section can be found in RFC 2223, and it reads as follows: Each RFC must have at the very end a section giving the author's address, including the name and postal address, the telephone number, (optional: a FAX number) and the Internet email address [116]. RFC 2223 does however not specify specifically who to mention in the *Author's Address* section. This very issue was also discussed in the Document Life cycle Training session on the 73^{rd} IETF meeting. The conclusion in the training session was that currently the IETF has somewhat imprecise rules on defining and documenting exactly who should be documented as authors and contributors in the *Author's Address* section of the document. Currently the data reflects what the authors have chosen to write in the documents and thus significant contributors can be missing or others can be mentioned as contributors even though they have not contributed at all.

In this case, the amount of IETF documents in the project database permits us to do quite extensive manual evaluation of the parsed results. We have 5334 RFCs and 53911 Internet-Drafts in the project database. The parsed *Author's Address* sections of all RFCs have been checked and errors have been corrected. We also created tools that enable us to make and store custom rules and exceptions to enhance the accuracy of the *Author's Address* section parser algorithm. The special rules and exceptions database table contain 3648 strings we know that are author names, 2465 strings we know to be organization names, 613 strings we know that certainly are not author names, and 1089 lines that we order the parser to skip altogether. These special rules greatly enhance the parser algorithm and in many cases enable us to correctly parse several *Author's Address* sections that without these rules would be quite impossible to parse.

Meeting attendee lists for each IETF meeting are published as a part of the meeting proceedings. These list also reflect what the attendees have chosen to document about themselves when registering for the meeting. 31 of the 45 IETF meeting attendee lists we have inserted into our database contain the most relevant data for our analysis, namely the attendee names and their organizational affiliation. 14 of the attendee lists lack information about the attendees' organizational affiliation, however those lists do contain the attendees' email addresses. Cross referencing meeting attendee lists and author data from the working documents help us to establish the missing information. Email address domain names are also helpful in establishing the organizational affiliation in many cases.

To collect data from the IETF mailing lists should be technically a fairly straight forward task. However, the mailing list message files, that should confirm to the *Internet Message Format*, does not always do that. There are several erroneous files that contain incorrect messages in the archive. The reasons for the errors are difficult to present. Certainly there are email client software that do not produce messages conforming to the prevalent standard. The bulk of the messages are parseable and our parser script can successfully extract the specified data from the messages. In addition to technical obstacles, there is also the issue of spam messages on the mailing lists. The IETF mailing lists are however managed in accordance to an IESG statement that spam control must be provided on each IETF mailing list [73]. However, there are a substantial amount of spam messages on the IETF mailing lists interfering with our analysis. The parser script does however match the *X-Spam-Score* header field and insert the value into the project database, thus enabling us to sort out the bulk of the spam messages later on.

The additional data is mostly collected from external sources and stored, as is, into the project database. Financial data about the selected set of companies is picked up from the financial statements in the annual reports of the companies by hand. Stock performance data is gathered from the MSN Money service, also using a manual process. In the scope of this thesis we thus consider the aforementioned data to be reliable and correct. The organizations for which we have collected financial data are presented in Section 5.4 in Table 11.

The USPTO database was queried using the full-text quick search tools provided on the web-page [142] [140]. The full-text search returns the patents or patent applications that contain our search string anywhere in the writing. We did not limit the query to match our search string only in *References* section. Using this approach we believe that we will get a more generic salience figure for a specific RFC-document in the patent domain.

Similarly the Google Search and the Google Scholar academic literature Search gives us generic figures on the importance of the RFCs. Both databases were queried in a full-text search manner, thereby the number of hits reflect documents where the search string may be anywhere in the writing. Regarding the Google queries, it is also challenging to establish the definite quality dimensions and limitations of the data. The data will therefore be used bearing these limitations in mind. The Google Search hits reflects a general popularity of the documents within the Internet community and the Google Scholar hits will provide us with a general understanding of how each RFC have impacted the academic community.

5 Analysis

In this chapter we will perform certain analyses on the data we have constructed into our database. Primarily we will search for a meter measuring success in Internet standardization. We will also take a closer look at the accomplishments of the Finnish ICT cluster within the IETF. Ultimately we will also establish the correlation between the efforts put into Internet standardization by IETF participants and their success on the marketplace.

5.1 Methodology

As mentioned, our goal is to quantify the results and take different analytical views on IETF standardization work. We have collected data measuring several different aspects of participants' activities within the IETF. In addition to that, we also have data that enables us to look at the participants' performance outside the standardization scene and thereby compare their standardization activities to their performance in their actual line of business.

Deciding how to measure performance and success in standardization work, proved to be challenging. Public academic research on this topic seems to be quite scarce. There are some articles focusing on e.g. measuring the performance of SDOs, other articles looking at the impact of the actual standards produced, and also some articles analyzing the co-operative and competitive aspects of standardization work [128], [125], [100]. Research on the inner workings of SDOs appears absent, also data on the inner workings of SDOs is not generally available. Hence, in the scope of this thesis we have had to put quite an effort in constructing the IETF data to be analyzed. The aforementioned issues consequently have an effect on our choice of analysis methodology.

A primary goal will be to establish what the benchmark for success regarding IETF standardization work is. This part of the analysis will be done from the IETF's point of view. Exemplary participants best serving the IETF's mission will be isolated. IETF data covering a longer time-span will be used to enable us to establish trends and changes over time.

In addition to taking a look at the inner workings of the IETF we will also establish the correlation between the level of activity in Internet standardization and participants' success on the marketplace. Figures such as Net Sales and R&D expenditure will be compared to quantities in IETF participation.

Finally we will also analyze the accomplishments of the Finnish ICT cluster in the IETF. The Finnish participants will be extracted from the data and similar analyzes will be made on them, as on the IETF general population. The Finnish company SSH Communications Security will be given special attention, as their technology and their success has a special relationship to standards development work done in the IETF.

Our analysis methodology also includes a system to classify IETF documents by importance, thereby giving participant contributions an extra dimension of quality. This document importance classification can be made only for the RFCs, as there by definition are no references to Internet-Drafts. The IETF process document states that under no circumstances should an Internet-Draft be referenced by any paper, report, or Request-for-Proposal, nor should a vendor claim compliance with an Internet-Draft [13]. The RFCs are however given an importance dimension measured using the variables listed in Table 8.

Table 8: RFC importance factors.

Variable	**Amount of**	**Source**
RFCcr	RFC cross-references	IETF RFCs
RFCip	Issued patents	PATFT: USPTO
RFCpa	Patent applications	AppFT: USPTO
RFCar	Ref. in academic literature	Google Scholar
RFCgh	Hits using Google search	Google Search

Using all five importance variables will enable us to calculate a unified importance value, $RFCi$, for each RFC as in Equation 1.

$$RFCi = (RFCcr)/100 + (RFCip)2 + (RFCpa)1 + (RFCar)/100 + (RFCgh)/1000 \quad (1)$$

Equation 1 will be used in Section 5.2 to classify all RFCs by their impact and as a continuum to that, in Section 5.3, we will determine the participants who have actively contributed to the most influential RFCs. First we will however take a look at some general parameters of IETF work.

5.2 General Measures of IETF Performance

In this section we will take a look at some general performance measures of the IETF. We will focus on analyzing aspects of standardization that have been covered in articles referenced earlier in this thesis, however adding more recent data and some other fresh views.

In Section 2.7 we looked at the time it takes SDOs to complete a standard. There was a general concern that the delays in standardization are increasing meanwhile the product and market life-cycles are decreasing. Simcoe concluded that over time it has become harder for the IETF to produce rough consensus in a timely fashion [127]. Simcoe analyzed the delays in Internet standardization from 1992 to 2000. As a response to his work, we will here look at the development in IETF delays from 1995 to 2008 and thereby see if the slowdown Simcoe noticed has endured. We will also measure the growth of the IETF and analyze the trend in the complexity of IETF standards.

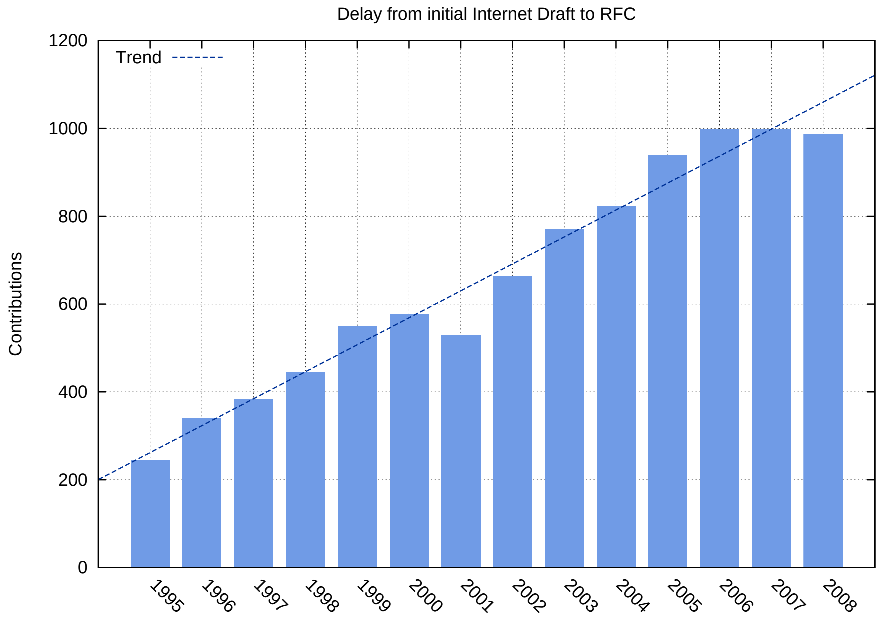


Figure 8: Delays from publication if initial I-D to the publication of the RFC

The delays in IETF standardization can bee seen in Figure 8. Columns represent the time, in days, it takes to create an RFC, measured from the initial publication of the first version of an Internet Draft to the publication of the RFC. The trend line is calculated using an implementation of the nonlinear least-squares Marquardt-Levenberg algorithm [107]. We clearly see that the delay in publishing an RFC has been growing since 1995, however from 2006 to 2008 there seems to be a deceleration in the growth of the delay. At the same time as the delay in RFC publishing has risen, the IETF itself has also grown in many aspects. The amount of documents published per year, the number of attendees at the meetings, and the amount of posts on the mailing lists have increased.

In relation to the increased delay in the IETF standards development process we also take a look at how the document length of the published RFCs have evolved. Again, we look at all RFCs from 1995 to 2008. In this context Simcoe used the length of the RFC as a proxy for its technical complexity [127]. Shah's and Kesan's results also indicated that the length of a standard and the number of references to it, were significant measures of its impact [125]. Using the data in our database shows us that the average length of an IETF RFC has grown from 50.000 characters in 1995 to 65.000 characters in 2008. However, IETF's goal is to keep standards as concise as possible. Shah and Kesan thus suggest that longer standards cover more technical terrain and are therefore more complex and also take longer to develop. The data in our database suggest that this is true. Analyzing the data gives a factor of 0.67 for the linear correlation coefficient between the average time it took to develop an RFC and the average length of an IETF RFCs from 1995 to 2008.

Furthermore, as described in the previous section, we use certain variables to give each RFC an importance value. Using Equation 1, we calculate a unified importance value for each RFC. Table 9 lists the top 10 RFCs by our calculated importance rating. The top three RFCs are RFC 1889 *RTP: A Transport Protocol for Real-Time Applications*, RFC 3261 *SIP: Session Initiation Protocol*, and RFC 2616 *Hypertext Transfer Protocol – HTTP/1.1*. In Section 5.3 we will extend the analysis to see which IETF participants have been involved in authoring the most influential IETF documents.

Table 9: RFCs by importance.

RFC	**RFCcr**	**RFCip**	**RFCpa**	**RFCar**	**RFCgh**	**RFCi**
RFC 1889	1524	451	621	71600	5410	1663.94
RFC 3261	7801	109	1016	297000	5360	1662.61
RFC 2616	7104	253	557	320000	4730	1501.34
RFC 2543	438	332	557	63200	2590	1314.48
RFC 2002	585	369	344	29400	3700	1154.25
RFC 2119	52888	5	21	424000	8290	1066.78
RFC 1866	329	319	360	27900	995	1039.14
RFC 2131	4185	256	298	89500	1930	960.65
RFC 2205	4853	253	320	42200	4350	960.23
RFC 1321	2989	206	377	96100	3890	953.89

Technical details on how the RFC importance data was constructed can be found in Appendix A.

Finally, we will take a look at the activities and performance of the IETF working groups. Using the importance measure for the RFCs and the overall amount of documents affiliated to the Working Groups allows us to establish the most influential Working Groups. Classifying the RFCs by importance using Equation 1, then grouping the documents by working group and adding up all the importance values gives us one view on the most influential IETF Working Groups. The result is that the *Point-to-Point Protocol Extensions* (pppext) Working Group is the most influential when counting in all RFC importance values. The pppext Working Group is tightly followed by the *Audio/Video Transport* (avt) and *IP Security Protocol* (ipsec) Working Groups. The *HyperText Transfer Protocol* (http) and the *IP Version 6* (ipv6) Working groups are also among the top five most influential Working Groups. The avt, ipv6, and pppext Working Groups are also the most productive with 95, 80, and 75 RFCs published respectively. The http Working Group has published only 12 RFCs, but the sum of importance values still entitles it the fourth most influential Working Group title.

5.3 Benchmark For Success

Here we will take a closer look at the outcome of the IETF participants' efforts. Exemplary participants and success factors will be established using data from our database. In the scope of this thesis we will define exemplary participants in IETF work, as those who best serve the goal of the IETF as defined in its mission statement.

The mission of the IETF is to produce high quality, relevant technical and engineering documents that influence the way people design, use, and manage the Internet in such a way as to make the Internet work better [2].

Serving the mission of the IETF can be done by attending IETF meetings, participating in the discussion on IETF mailing lists, and authoring documents. In addition to the aforementioned, participants can also contribute by serving as chair for one of the Working Groups or possibly work as an IETF Area Director. We have not compiled data on IETF chairs and ADs, therefore why we will focus our analysis on document authors, meeting attendees and mailing list activity, which as well form the bulk of IETF day-to-day work.

First we will establish the most active IETF participants. This is done using data on IETF meeting attendance, mailing list activity, and RFC authorship. Combining the top 70 most active participants from each category gave us a list of 80 distinct organizations. The top 20 most active IETF participants are presented in Figure 9. Meeting attendance is presented in 10^s and mailing list contributions are presented in 100^s. The earliest data on meeting attendance is from 1994, so activities in all categories were limited to cover the time from 1994 to 2008.

Looking at all three performance measurements makes it evident that Cisco is in a class of its own. Cisco is a U.S. based company that was founded in 1984 by computer scientists Len Bosack and Sandy Lerner from Stanford University. Cisco's corporate overview states that Cisco's offerings are used to create the Internet solutions that make networks possible. Cisco engineers have also been leaders in the development of Internet Protocol (IP) based networking technologies since the company's birth. [26] Cisco's mission statement and the scope of the IETF clearly converge quite well. The academic background of Cisco's founders can also be explanatory factors for Cisco's success in Internet standardization. IETF participants representing Cisco have authored more RFCs than the second, third, fourth, and fifth largest contributors combined. Cisco's RFC contributions represent c. 6% of all RFC contributions.

Another interesting aspect of IETF participation is the trends and changes over time in stakeholder activities. As a part of this thesis we created a computer script that enables us to establish the trend-line slope for each organization that has contributed to the IETF working documents. The trend-line is calculated using a linear regression model. The procedure will fit a straight line of the form:

```
y = b + mx
```

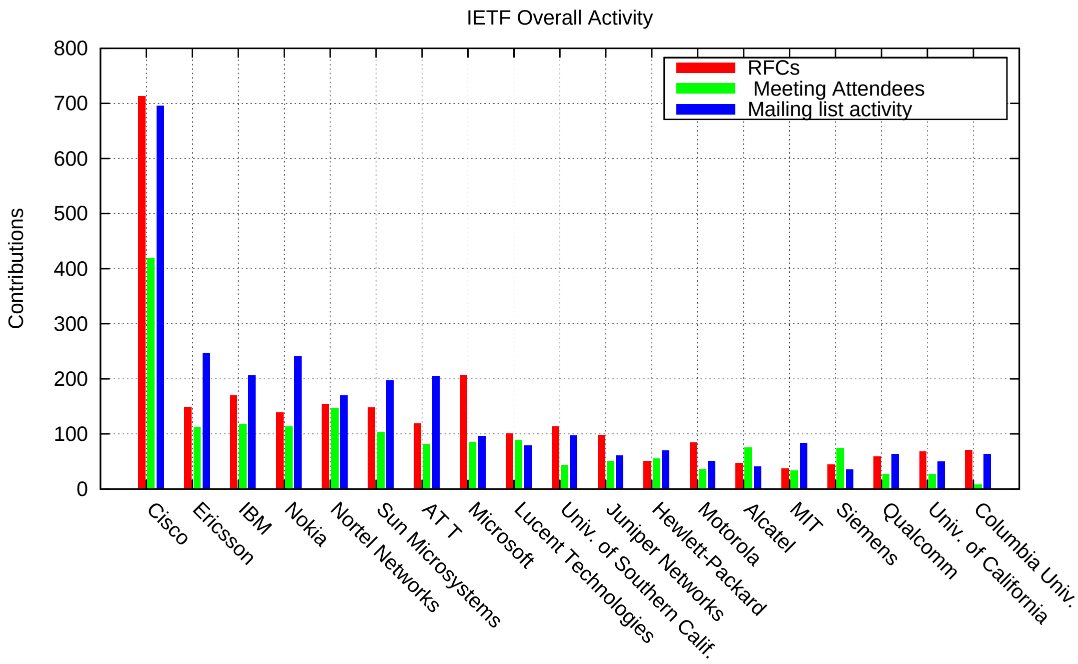


Figure 9: IETF Overall Activity

to the data. The top and bottom five participants with most vs. least accretion in IETF contributions are listed in Table 10. We have calculated two distinct 36 month trends ending in 1998 and 2008. There is an apparent difference in the data from 1998 and 2008. E.g. IBM and Microsoft were top risers in the late 90ies, on the other hand in 2006-2008 they were among the organizations reducing their IETF contributions the most.

Table 10: Participant risers and fallers.

36m trend 1998	**Slope**	**36m trend 2008**	**Slope**
Microsoft	.89	Huawei Tech.	.44
Cisco	.76	British Telecom	.36
IBM	.52	AVAYA	.18
Sun Microsystems	.50	China Mobile	.16
Lucent Tech.	.45	Ericsson	.16
Carnegie Mellon Univ.	-.02	Intel	-.13
Isode	-.02	IBM	-.14
Univ. of Washington	-.02	Samsung	-.16
Motorola	-.03	Microsoft	-.17
DEC	-.04	Motorola	-.21

Huawei Technologies stand out as the organization increasing their IETF efforts the most in the same time period. The calculations behind the figures in Table 10 included both RFC and Internet-Draft contributions. The distinct increase in Huawei's contributions can yet be a reflection of the evolving Internet-Drafts process within the IETF. Between 1990 and 1999 the Internet-Draft-to-RFC ratio was 6.8, however considering documents published between 2000 and 2008, that same ratio measured 16.0, when accounting for the top 50 Internet-Draft and RFC contributors. Cisco lies very close to the average in both periods. Huawei on the other hand has an I-D-to-RFC ratio of 95.6 in the second period, that is 671 published Drafts versus 7 Published RFCs. The University of Southern California represents the other extreme, with an I-D-to-RFC ratio of 1.9 in the former period and 7.4 in the second equivalent.

As a final measure of success, we will analyze participant contributions in relation to the impact of the IETF standards. The importance factor given to each RFC, as presented earlier in Table 9, is used to establish the participants affiliated to the most influential RFCs. The most influential RFCs can certainly be said to serve the IETF's mission to make the Internet work better. Limiting the set of documents to be processed to the 100 most influential RFCs, and extracting the affiliated organizations from those documents, gives us a list of 115 organizations. The 15 most active contributors from this set is presented in Figure 10.

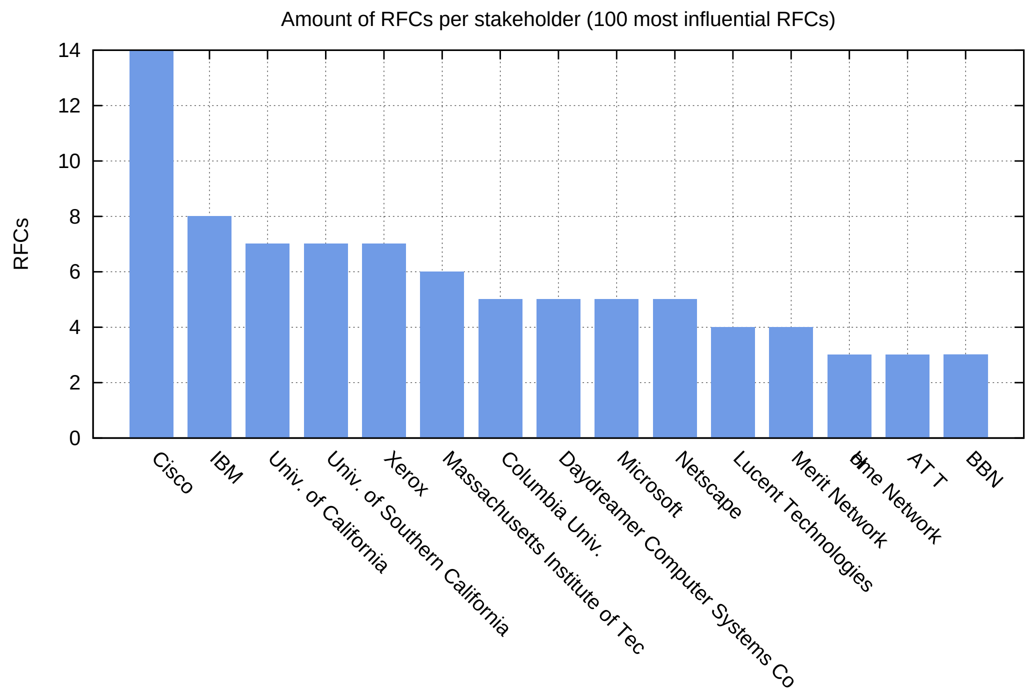


Figure 10: Organizations affiliated to the 100 most influential RFCs

Again, Cisco stands out as the exemplary participant. In contrast to the most active contributors overall, presented in Figure 9, we can see that the organizations in Figure 10 includes stronger representation from the academic community. The overall top ten RFC contributors include only one academic participant, whereas the top ten contributors to the 100 most influential RFCs include four academic participants.

5.4 Success in Standardization and Success on the Marketplace

In this section we will take a look at how the participants in IETF standards development have succeeded on their respective markets. We will try to establish if activity and success in standardization work will imply success on the marketplace.

We have stated some of the general benefits of standardization in Chapter 2.2 and also gone through reasons for different stakeholder groups to participate in standards development in Chapter 2.6. However, now we set out to look more specifically on the economic aspects of stakeholder participation. But, as Bolin concluded, attempting e.g. to determine the return on investment of standards development is about as easy as measuring the ROI of education or training [10]. Relevant papers dealing with the economics of standardization are The DIN report on Economic benefits of standardization, the DTI paper on The Empirical Economics of Standards, the more theoretic approach by Farrell and Saloner titled Standardization, Compatibility and Innovation, and finally we want to mention the article on The Economics of Compatibility Standards: An Introduction to recent research by David and Greenstein [38] [136] [45] [34].

Several of the papers discuss macroeconomic aspects of standardization. E.g. between 1981 and 2004 in Canada, a 17% growth rate in labor productivity and a 9% growth rate in GDP has been linked to standardization [56]. In the United Kingdom, standards have contributed to a 2.5 billion annual growth of the GDP, that is an 75 billion overall increase since 1948. UK standards have also contributed to 13% of labor productivity growth over the period from 1948 to 2002. [104] Finally, in Germany the estimated national economic benefits of standardization is approximately 1% of gross national product [38].

More detailed data on specific organizations' and companies' economic success in relation to their standardization activities seem to be rare. However, as a part of this thesis we have collected a set of financial parameters concerning certain organizations that are active in the IETF. We will compare Net Sales, R&D expenditure, and Stock performance to the extent of participation in the IETF. Financial data have been collected for the organizations listed in Table 11.

First we will take a look at Cisco and Juniper Networks and their achievements in the Internet router segment. Cisco represents the largest company on the Internet router market. Cisco is also the overall most active IETF participant. Juniper Networks, on the other hand, founded in 1996, is a newer company on the router market. Juniper networks is the ninth most active RFC contributor, looking at

Table 11: Financial data.

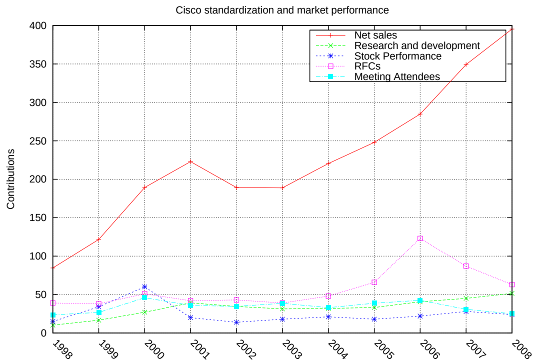
Organization	**Years**
Cisco	1993-2008
Ericsson	1993-2008
Juniper Networks	1998-2008
Microsoft	1993-2008
Nokia	1993-2008
NTT	1993-2008
SSH Communications Security	1997-2008

RFCs from 1999 to 2008, with close to 100 documents with at least one affiliated author. Cisco is respectively affiliated to about 600 RFCs. Cisco's and Juniper's IETF participation and financial figures are presented in Figure 11.

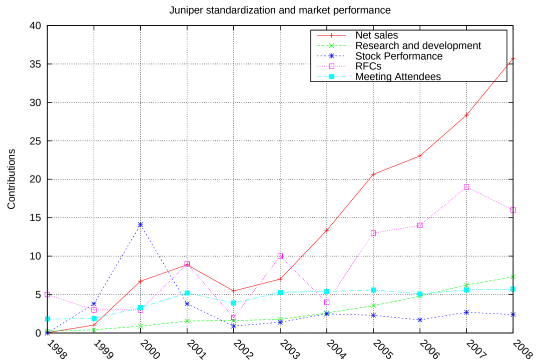
The Internet router market is at the time of writing largely dominated by Cisco and Juniper Networks. Barron's estimate the Edge Internet Protocol/MultiProtocol Label Switching (IP/MPLS) market to be about $7.8 billion annually and Core IP/MPLS segment to be about $3.2 billion market annually in 2008 [91]. Barron's further suggest that in 2008 Juniper Networks is gaining the most share and Cisco Systems losing the most market share. Market shares in the Core Router segment estimated by the Dell'Oro Group in the third quarter of 2008 gives Cisco a market share of 61%, down from 65% in the previous quarter, Juniper Networks had 18 percent, up from 17%, Alcatel-Lucent's share grew to 7% from 6%, and finally Huawei Technologies' share increased to 6% from 5%. [126] A Reuters article states that most analysts estimate Cisco's 2008 market share in core routers to be around 60% and Juniper Networks' to be around 30% [3].

The *Routing Area* standardization efforts of the above-mentioned companies looks as follows. There are 256 IETF RFCs published by working groups in the *Routing Area* between 1999 and 2008. Cisco has contributed to 47% (121) of the RFCs, Juniper Networks has contributed to 26% (66) RFCs, Alcatel/Alcatel-Lucent to 14% (35), and finally Huawei Technologies has contributed to 0.4% (1) of the IETF Routing Area RFCs. Measuring the standardization activity puts the companies in the same order as above where market shares were compared.

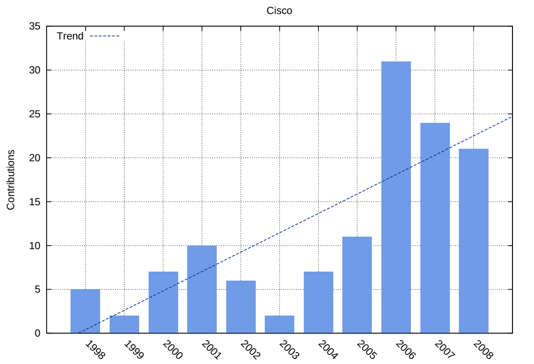
Furthermore we will also take a closer look at correlations between some of the financial variables and standardization activity parameters. The correlation is calculated using the equation for a Pearson correlation coefficient presented in Equation 2. The results are presented in Table 12. The data will not fully determine the causality, but we will still get a measure of how the financial data and the standardization activity data is related.



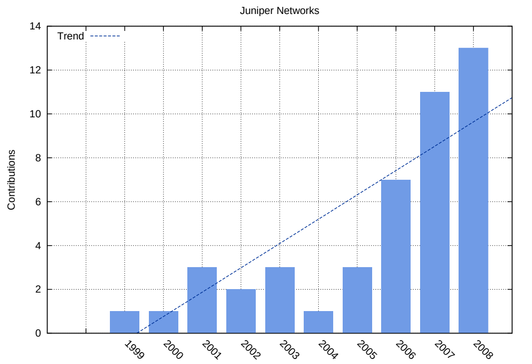
(a)



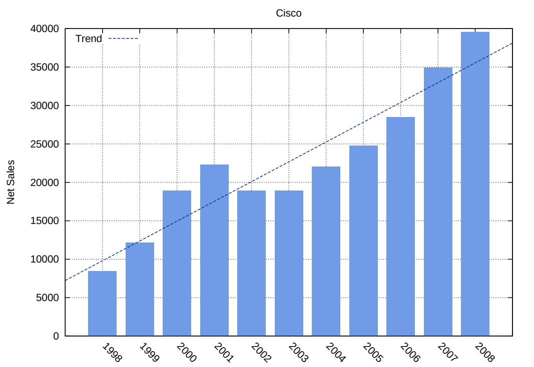
(b)



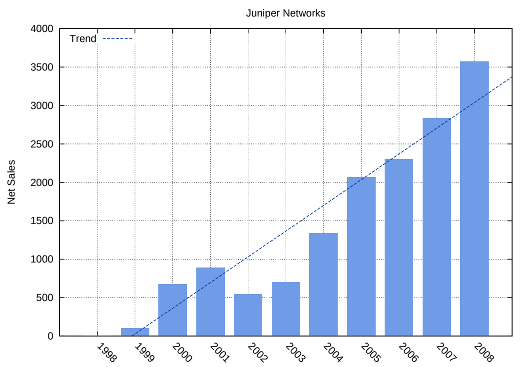
(c)



(d)



(e)



(f)

Figure 11: Financial charts

$$r = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \quad (2)$$

Table 12: Pearson correlation coefficients.

Company	**Years**	**Sales RFCs**	**Sales M.Att**	**R&D RFCs**	**R&D M.Att**	**Stock RFCs**	**Stock M.Att**
Cisco	93-08	0.94	0.71	0.96	0.78	0.45	0.73
Ericsson	93-08	0.37	0.87	0.45	0.88	-0.09	0.69
Juniper	98-08	0.61	0.75	0.61	0.72	-0.38	-0.17
Microsoft	93-08	0.56	-0.27	0.67	-0.26	0.89	0.22
Nokia	93-08	0.78	0.59	0.81	0.53	0.49	0.81
NTT	93-08	0.65	0.74	-0.31	0.04	-0.7	-0.74
SSH	97-08	0.45	0.41	0.54	0.48	0.32	0.51

The correlation coefficients describe a strong relationship between both Cisco's Net Sales and RFC output and between Cisco's R&D expenditure and RFCs output. The relationships between Cisco's IETF meeting attendance and financial parameters are not quite as strong. Nokia also seem to have a fairy strong relationship between its Net Sales/R&D expenditure and its level of IETF activity. The data on the correlation between stock prices and IETF activity does not enable us to draw any sane conclusions. Ericsson's data suggest that they have a fairly strong relationship between their R&D expenditure and IETF meeting attendance.

Our goal is to establish if success in standardization work will imply success on the marketplace. It might be hard to definitively settle using the data we have collected. However, there clearly seem to be a relation between the level of participation in standardization and the size of the organization and it's market share. The causality is on the other hand difficult to determine. Further analysis can be meaningful to carry out with market data that accurately reflects some specific working area within the IETF. Then the IETF data can be used to establish the relationship between market shares in a specific market linked to a architectural segment of the Internet and the contributions affiliated to the standards of that segment of the Internet.

5.5 IETF Performance of the Finnish ICT Cluster

The Finnish ICT cluster has been active in the IETF for a significant amount of time. Both commercial companies and Finnish academia are represented. In this section we will take a closer look at these participants and their contributions to the IETF standardization work.

Internet research and development in Finland is very much in action at the time of writing. The strategic centre for science, technology and innovation in the field of ICT (ICT SHOK) functions under the auspices of Tekes, the Finnish Funding Agency for Technology and Innovation, and is operated by TIVIT Oy. The goal of the ICT SHOK is to is to advance the progress of know-how in the Finnish ICT sector and increase the pace at which research results and innovations are embraced by the business community. [135] One of the four programmes running under the ICT SHOK is the Future Internet programme. The goal of the Future Internet programme is to bring together the key research resources to develop future Internet networking technologies and to create new global ICT based business ecosystems [65].

Creating new ICT business ecosystems include international collaboration as well as influencing and taking part in Internet standardization. The program plan mentions active interaction with international SDOs. Short and medium term activities thus include participating in applicable areas of both the IETF and the IRTF. The Future Internet Programme Plan also identifies the dissemination of results and international collaboration as a key parts of the work. Creating and maintaining presence in international SDOs, particularly in the IETF and thereby enabling active dissemination of programme results and constructs is considered essential. [66] The programme Research Agenda furthermore states that, strong links to key research groups and researchers in the industry and academia worldwide established. Maintaining close liaisons with key standardization activities in IETF, W3C, and other SDOs are needed for the group to reach its intended impact [67]. Dr. Andrei Gurtov, representing the Helsinki Institute for Information Technology (HIIT), sets a veritable example by working as co-chair for the IRTF Host Identity Protocol (HIP) Research Group.

The ICT SHOK Future Internet programme includes the following industrial partners: Ericsson, Nokia, Nokia Siemens Networks, TeliaSonera, CSC, and the following research partners: HIIT, Tampere University of Technology, TKK, University of Helsinki, University of Jyväskylä, and VTT. The Future Internet Programme Plan was written in Apr. 2008 so there are not likely to be many tangible results showing in the IETF yet. However, a majority of the programme partners have been, and are, active in the IETF anyhow and therefore we will focus on their accomplishments in the IETF so far. That is, analyze the success of the Finnish ICT cluster, both industrial partners and academic entities, in the IETF.

The Finnish ICT cluster will be outlined from the IETF general population using some purpose built algorithms and rules. The baseline is that the contribution has to originate from Finland in order for it to be accepted as a Finnish piece of work. We use the IETF working documents, both RFCs and Internet-Drafts, as a starting point knowing that the *Author's Address Section* most often include a line directly stating the author's country or some other address information by which it is possible to determine the country of origin. Email addresses with the Finnish top-level domain *.fi* also classifies a contribution as Finnish. All Finnish contributions are extracted using the Perl compatible regular expression documented in Figure 12.

```
"/(FI\-[0-9]|\.fi[^a-zA-Z0-9_]|\b(FI|Suomi|Finland|Helsinki|Tampere|Espoo|Oulu|Jorvas|Vantaa|Turku)\b)/"
```

Figure 12: Regular expression for extracting Finnish contributions

From the result-set returned by the regular expression we manually isolated companies and academic participants that operate only from Finland and we are thereby able to decide that all documents, meeting attendees, and mailing list activity affiliated the those participants are counted in as Finnish contributions. Some large commercial organizations like Ericsson, Nokia, and Nokia Siemens Networks which do operate in Finland, are multinational by nature and all their contributions can not be considered Finnish. In these cases we resort to extracting both the author and organization names from the IETF working documents. We then use that value-pair to determine which meeting attendees and mailing list entries belonging to the multinational organizations should be considered a part of the Finnish ICT cluster.

By using the tools we have designed and the project database, we can establish that there are approximately 3900 IETF documents, both RFCs and Internet-Drafts with at least one Finnish author. The trend-line, when measuring the amount of documents with at least one Finnish author, from 2000 to 2008 is rising from approx. 230 to 485 documents per year. However, if we limit the set to documents from 2002 to 2008, the trend line is rising much more moderately from approx. 380 to 430 documents per year. The Finnish authors are affiliated to over 40 different organizations.

Figure 13 shows the overall achievements of the Finnish ICT cluster in the IETF. Meeting attendance is shown in 10^s and mailing list activity in 100^s. Ericsson's mailing list activity column have been cropped to fit the chart, the correct values is 106.

Ericsson and Nokia are by far the most active IETF contributors from Finland. Nokia and Ericsson are also among the largest research organizations in Finland. Teknikka & Talous listed Nokia as the largest, and Ericsson as the fifth largest research organization in 2008 [134]. Nokia's global R&D expenditure was 5968 million € and Ericsson's expenditure in Finland 92 million € in 2008.

The most recent efforts of the Finnish cluster can be established by looking at the active Internet-Drafts. At the time of writing we have eight Finnish organizations with active IETF I-Ds. See Figure 14 for details. Ericsson and Nokia are not surprisingly again among the largest contributors. Nokia Siemens Networks is however the most active organizations with active I-Ds. It is also worth noting that four of the Finnish I-D contributors are academic entities.

Analyzing the most recent Finnish IETF contributions from a individual perspective shows that the set is quite focused around two individuals. Hannes Tschofenig from Nokia Siemens Networks and Gonzalo Camarillo from Ericsson combined are affiliated to over 35% of all the Finnish I-D contributions. Mr. Pekka Savola and

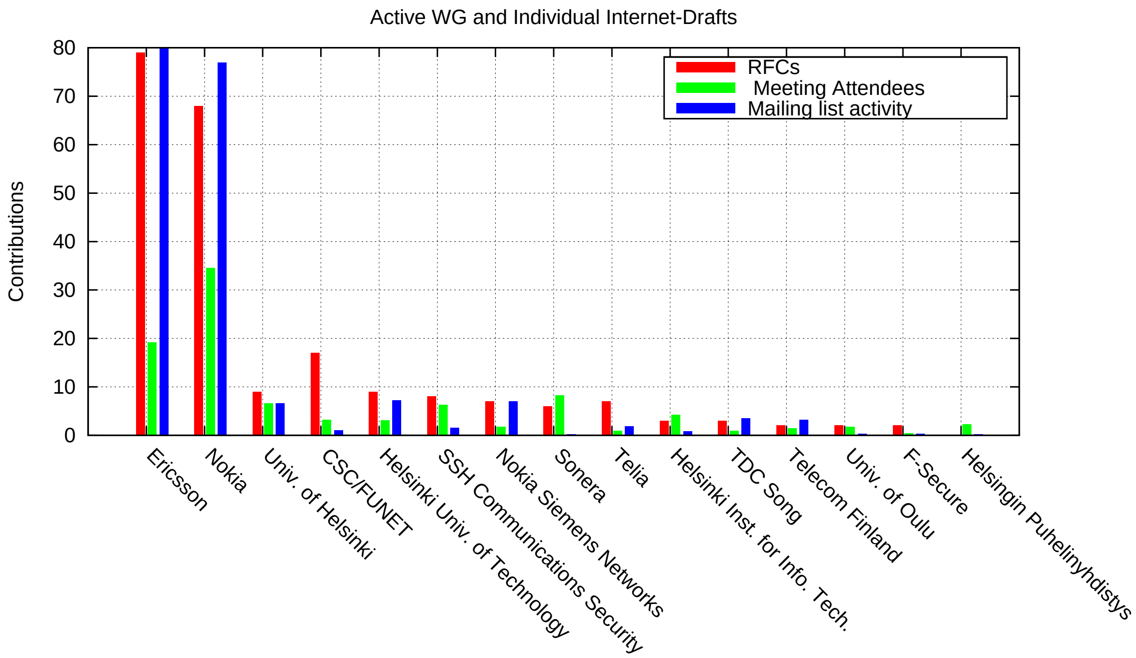


Figure 13: Overall activity of the Finnish stakeholders

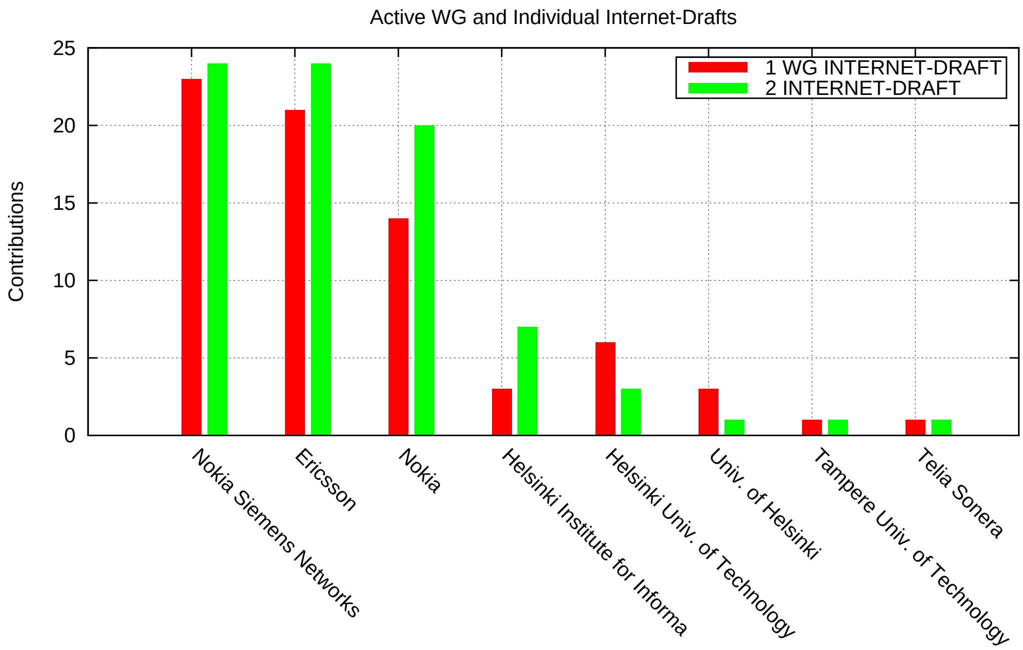


Figure 14: Recent activity of the Finnish stakeholders

Mr. Jari Arkko are also among the overall most active individuals on the IETF mailing lists with over 5000 and over 3000 mails sent respectively. All in all the Finnish IETF community totals to approximately 40 individuals when looking at the currently active Internet-Drafts.

To give the Finnish IETF contributions a further dimension of relative quality we choose to compare them to the contributions from a number of other key countries active within the IETF. In this comparison we include the U.S., the overall most active IETF country, Germany, Sweden, and Finland from the E.U. and China and Japan from Asia. We will look at the amount of companies active in each country and furthermore also analyze the amount of contributions originating from the selected countries. In this context each author mentioned in an RFC or Internet Draft will give credit to one contribution point. In addition to IETF working document related activities we will also collect the sum of meeting attendees per country from IETF meetings 72, 73, and 74.

Table 13: IETF activity by country. Amount of organizations contributing.

Country	**RFCs**	**I-Ds**	**Active I-Ds**
U.S.A.	790	1760	244
Germany	57	168	46
Japan	55	134	32
Sweden	30	72	10
Finland	20	34	8
China	8	79	28

The amount of contributing organizations from each country is presented in Table 13. The total amount of organizations the Finnish authors are affiliated to is relatively small in comparison to the other countries analyzed. However, when looking at the actual amount of contributions from the Finnish authors a somewhat different light is shed on the situation. The amount of RFC contributions from Finnish authors is not far behind Germany and Japan and the amount of Finnish contributions supersedes the amount of both the Swedish and the Chinese contributions. The figures are presented in Table 14.

Furthermore we can also see that there are only 8 organizations in Finland having active Internet Drafts at the time of writing. Here Finland is roughly at the same level as Sweden. However, counting the amount of active Internet Draft contributions leaves Sweden considerably behind Finland with only 64 contributions versus 173 contributions from Finland. Analyzing the actual amount of active I-D documents lets us establish that Finnish authors are affiliated to 133 active I-Ds where the Swedish authors are affiliated to 46 active I-Ds. The amount of active Internet Draft documents affiliated to the other countries are as follows: 932 documents with U.S. authors, 171 with German authors, 145 with Chinese, and 115 documents with Japanese authors. The U.S. is clearly the largest and most active IETF contributor

Table 14: IETF activity by country. Amounts of document contributions and meeting attendees.

Country	**RFC**	**I-D**	**Active I-D**	**Meeting att.**
U.S.A.	5799	69022	1723	1672
Germany	268	5908	225	180
Japan	255	5456	189	311
Finland	215	3800	173	110
Sweden	167	2622	64	86
China	30	1926	255	255

in every category analyzed. Altogether there are authors from over 40 countries that have authored or co-authored IETF RFCs. Overall Finland is placed at a fifth position behind the U.S., the UK, Canada, and Germany when counting in all RFCs with at least one author from the respective countries.

In addition to looking at the overall situation, we separately also examined RFCs published in the five most recent years. This gives us a somewhat different list of most active IETF countries. The U.S. and the UK are still the two largest countries, but Finland distinguish itself as the third most active country. The eight most active countries, when taking RFCs published from 2005 to 2009 into account, are listed in Table 15. It is worth noting that the overall top eight country list and the top eight countries in the last five years are identical except of Finland progressing from a fifth to a third position. When regarding year 2008 exclusively, we can also establish that Finland was the second most active country with 43 RFCs having at least one Finnish author. The figures in Table 15 are constructed so that an RFC with at least one author from the respective country gives one credit. Figures for 2009 are not directly comparable to the earlier years as the project database contain only RFCs from January to mid August 2009 at the time of writing. We also made an equivalent comparison where each author mentioned in the RFCs gave one credit to the respective country. This did however affect the situation only slightly. The top eight countries remained the same, except for Canada and Japan changing positions between each other.

As a final part of our analysis on the Finnish ICT cluster's performance in the IETF, we will compile a small-scale case study on the achievements of the Helsinki based organization, SSH Communications Security.

SSH Communications is a company whose primary competence lies in developing end-to-end communications security solutions [129]. Their products use technologies such as Secure Shell (SSH), PKI, and secure authentication systems. The fact that makes SSH Communications particularly interesting in this context is that all core technologies they base their business on have been brought to the IETF for standardization.

Table 15: IETF activity by country. RFCs from 2005 to 2009.

Country	**2005**	**2006**	**2007**	**2008**	**2009**	**Tot**
U.S.A.	235	293	226	197	97	**1048**
United Kingdom	31	48	37	38	30	**184**
Finland	28	42	20	43	16	**149**
Germany	22	27	25	34	19	**127**
Canada	22	33	11	19	13	**98**
Japan	25	19	14	21	10	**89**
France	15	12	21	12	21	**81**
Sweden	15	19	16	5	4	**59**

SSH1 and the SSH-1 protocol were developed in 1995 by Mr. Tatu Ylönen, at the time a researcher at the Helsinki University of Technology. Ylönen's work gained attention almost immediately and he realized that his invention could be put to wider use. SSH1 was released in 1995 as an open source program. By the end of the year 1995, there was an estimated 20000 SSH users globally and in response to the rising demand for support Ylönen founded SSH Communications to maintain, commercialize, and further develop the SSH technology [5].

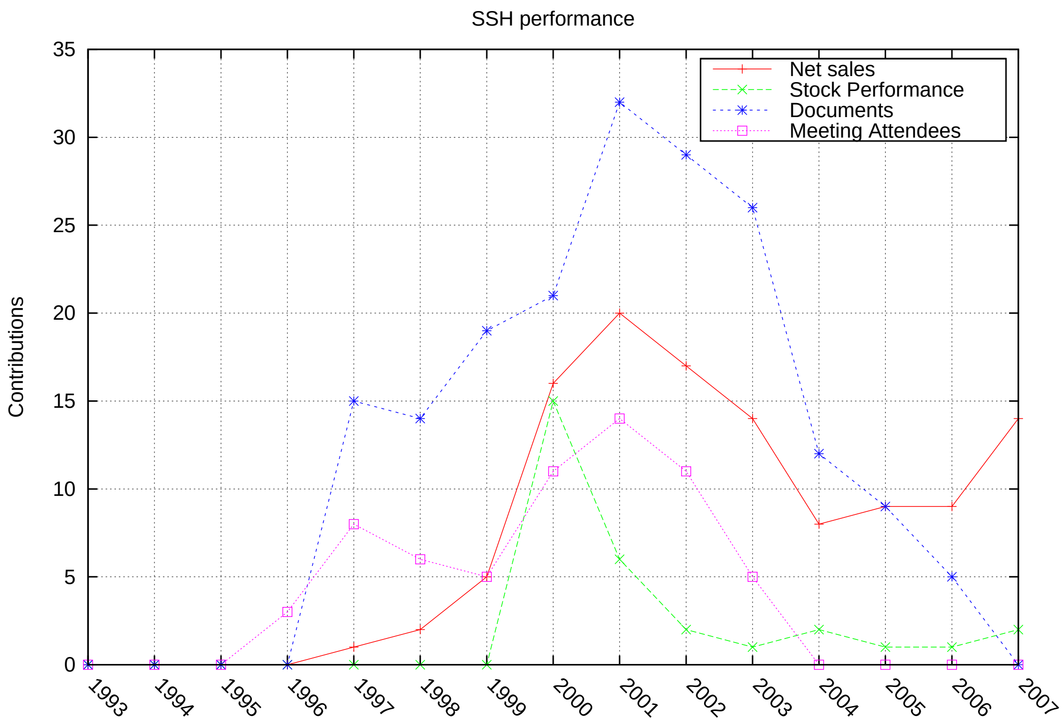


Figure 15: SSH IETF activity and market performance

IETF activity began almost immediately after the founding of the company, as Mr. Ylönen published his first Internet-Draft, titled *The SSH (Secure Shell) Remote Login Protocol*, in March 1996. Since that, 12 people, including Mr. Ylönen, representing SSH Communications have worked on a total of 174 Internet-Drafts and 8 RFCs. The amount of work made by SSH Communications employees can be seen in Figure 15.

The technical work on SSH security is considered a success. Mr. Tatu Ylönen was also granted the *Finnish Engineering Award* in 2000 for his work on the SSH data security solution [132]. Matters as very high development potential and fast financial success were mentioned in the award report. In retrospect SSH's financial success have however been weak. Teknikka & Talous mentions that SSH's revenues have been marching in place and the company has mostly showed negative profits [133]. The R&D expenditure of SSH Communications was 4.0 million € in 2008 [134]. SSH Communications have however not worked on any IETF documents since January 2006 when the 5 *Secure Shell* RFCs were published.

6 Discussion and Results

In the final chapter of this thesis we will look at the results of our analysis. Some further evaluation of the reliability of the results is presented and we will also look at some aspects of how the results could be utilized. Finally we will also document some topics for further research that might be interesting within this domain.

6.1 Results

Academic research on technical standardization is widespread. There are several articles covering different aspects of standards development. In this thesis we focused on ICT and Internet standardization. The reviewed literature allows us conclude that well organized technical standardization presents several benefits. Standards benefit the economy on both micro- and macro-levels. We see enhanced quality, reduced variety, more efficiently diffused information, and improved cost effectiveness. Standardization also lay the grounds for compatibility and interoperability between technologies. It can be argued that the exponential growth and success of the Internet would not have been possible without good quality open standards.

Standards development also present many benefits to the participating organizations. Standardization is a knowledge intensive activity thereby stimulating innovation. The vast amount of information shared within the SDOs give participants both time advantages and important insider knowledge related to new technologies. Furthermore participants are also able to shape their operating environment by their voting and thereby influencing the content of technical standards.

We also covered the special relationship between standardization and R&D. We conclude that organizations not only reduce the economic risk of their research and development by participating in standardization, but the corresponding costs can also be lowered.

However, none of the aforementioned benefits can be considered self-evident. The ICT standardization field is often characterized as fragmented and complex. There is a myriad of SDOs, consortia, and fora with overlapping activities. Added complexity in standards development complicates the task of creating timely, relevant, and robust standards. Better communication of the benefits and opportunities of standardization should hence take place. Greater collaboration and harmonization between the different entities in the ICT standardization universe should therefore also be targeted.

Internet standardization presented itself as an interesting research topic. Internet standards development is strongly focused around a single entity, the IETF, which is often also described as the paragon consortium for developing open standards. As a part of this thesis we designed a database and created a set of tools that enabled us to analyze the inner workings of the IETF. With this data we were able to quantify several aspects of the IETF day-to-day work. Our general analysis on the IETF data showed that the time from initial Internet-Draft to RFC publication has approximately doubled in the last ten years. At the same time the complexity measured by the length of the IETF RFCs has also grown. Since 2006 the increase in RFC development delay has slowed down and in 2008 the delay was actually smaller than in 2007. The Internet-Draft to RFC ratio has also changed significantly in the last ten years. In 1998 approximately seven I-Ds were published for every RFC, but in 2008 there were sixteen I-Ds per RFC. We could also see an increase in the share of commercial participants in contrast to the stark initial representation from the academic community. However, the academic participants are still well represented when looking at the most influential RFCs.

The first task in this thesis was to establish *what the benchmark for success in IETF Internet standards development work* is. Our analysis on IETF participants clearly showed Cisco as the exemplary contributor. Cisco was by all our measures the most active Internet standards developer. Cisco also has market share leadership in several domains where the technology is standardized within the IETF. Therefore we want to establish Cisco's IETF achievements as the benchmark for success when focusing on Internet standardization. Authors affiliated to Cisco have authored or co-authored approx. 800 RFCs whereas authors from the second most active participant Microsoft have authored a little over 200 RFCs. Regarding the Internet Drafts, authors affiliated to Cisco have submitted a total of almost 10000 I-Ds while the second most active participant measured by I-D submissions, namely Nortel Networks, has produced approx. 2800 I-Ds. The total number of IETF Meeting attendees, counted from meetings 29 to 74, affiliated to Cisco are also almost three times as many as the attendees from the second most active IETF meeting participant Nortel Networks.

Special attention was also given to the Finnish IETF participants and we focused on establishing *the success of the Finnish ICT cluster in IETF standardization work.* The academic community was exceptionally well represented when analyzing the most recent contributions from the Finnish participants. Four of the eight organizations with active Internet-Drafts were academic entities. Ericsson and Nokia were the two largest Finnish contributors. The bulk of Finnish contributions also originated from a relatively small amount of individuals. The two most active persons accounted for over 35% of the Finnish contributions. Being successful in Internet standardization can accordingly be much dependent on getting the right person to work for your company. We also compared the activities of the Finnish IETF participants to the achievements of participants from other countries. The results showed that the sheer amount of organizations contributing from Finland is relatively small, somewhat over 30 organizations all in all. On the other hand we were able to establish that the Finnish participants have been comparably productive. Finland is overall the fifth most active country, of the over 40 countries with IETF participants, when measuring RFC contributions. Thereby we conclude that the Finnish ICT cluster has been successful within the IETF.

We finally also included an analysis on the *IETF participants' success in standardization and their corresponding merits on the marketplace.* Macroeconomic benefits of standardization have been discussed in several papers, but in our analysis we however focused at specific organizations' standardization activities and their success on the market. We established that the most active contributors to the IETF *Routing Area* standardization also were the ones with the largest market shares in the Internet router market segment. Especially Cisco also showed strong correlation between both the development of their Net. Sales and their IETF activities and their R&D expenditure and IETF activities. The causality of these matters are hard to establish, but our results suggest that by actively taking part in standardization work and thereby shaping your operating environment you can positively affect the development of your market share and your success on the marketplace.

6.2 Assessment of Results

The results of our analysis and the conclusions we have drawn from the results regarding the IETF standardization work relies much on the quality and reliability of the data in the project database. We dealt with the technical limitations and the reliability of the data in Chapter 4.3 and here we further want to shed some light on the overall quality of our results.

We carried out the data mining task as an analytic process with the goal to explore and analyze the IETF data in search of consistent patterns and systematic relationships. We believe that the computer scripts and the tools used in this task certainly are accurate enough to give us data that enables us to measure and demonstrate the positions of the different participants within the IETF. Analyses on trends and other dynamic measures of IETF activities are also accurate despite a small number of parse errors when processing the working documents. The amount of contributions from the more active IETF participants are also large enough to eliminate the impact of some documents being unprocessable.

The computer scripts gathering and parsing the data are not perfect. The original source data is written natural language and thus challenging to parse using computer scripts. Many of the IETF working documents also provide only limited author information. Many errors can be detected by the scripts and subsequently corrected, but the amount of non trivial errors can be confirmed only by manually checking the parsed data for every IETF document. Tools enabling manual inspection were created as a part of this thesis and the data parsed from the RFCs was mostly reviewed. The majority of the more than 50000 Internet Drafts in the project database were not manually reviewed, but with proper resources it would be possible to do so.

In this thesis we also chose to define all the persons mentioned in the *Author's Address Section* as equal contributors to the document. In practice the list of authors can contain persons with very different contributions to the final document. The IETF rules on who to mention in the *Author's Address Section* may also not be completely clear and unambiguous.

6.3 Exploitation of Results

The results of this thesis can be utilized from basically two different angles. We are able to produce data and measure the performance of the standards development organization itself and also look at several measures on the achievements of the participants within the IETF.

Especially the data on the performance of the IETF itself can successfully be compared to previous research in the domain. In chapter 5.2 we e.g. compared different aspects of the IETF standardization work to analyses presented in articles cited in this thesis. These results can also be used to generally measure the time it takes to develop an Internet standard, the complexity of the standards and the relative importance of the documents published by the IETF. The results also shed light on the accomplishments of the working groups within the IETF.

The other part of our results are focused on the stakeholders participating in the IETF standards development. These results gives an overview of the organizations active in Internet standardization and also enables us to establish the absolute and relative dimensions of their contributions. These results can also be used from a participants point of view in order to perhaps establish if their level of activity is in balance with their competitors' standardization efforts. New entrants and other changes over time can also successfully be isolated as we have collected data covering a large time span.

6.4 Further Research

Data mining and academic analysis of standards development organizations' inner workings is sparse. The open principles employed by the IETF made our data mining task interesting as we were able to collect data on almost all the working processes in the IETF. The IETF is however not the only SDO with their working documents and other data openly available. This type of data mining could therefore successfully be expanded to cover also other standard development organizations as well. The project database is designed so that data from other SDOs can readily be inserted into it.

Cross-analyses and comparison of the data from different SDOs might also produce interesting results. Having data from several SDOs in the project database would enable us to effectively analyze the performance of both the SDOs and the participants' achievements in the different standardization venues.

Another aspect of participating in standards development is the co-operational schemes implemented by the participants. Further analysis on our data could thus include a network view on the document authors and their affiliations. IETF documents are on average co-authored by 2.6 authors and more than 60% of the documents are authored by more than one person. Applying network analysis on this data can therefore also clarify the positions and interrelations of the participants.

Acquisitions and mergers are commonplace among the organizations active with

Internet technology. Future work could thus also include a view on the impact of acquisitions and mergers on the standardization work. E.g. the Cisco Corporate Overview contains a summary over acquisitions and it lists well over 100 acquired companies since 1993 [27].

References

[1] **Ali-Vehmas, T.** *Service Diffusion Strategies for Push to Talk Over Cellular*, 2004

[2] **Alvestrand, H.** *RFC3935: A Mission Statement for the IETF*, 2004
Available at: http://www.ietf.org/rfc/rfc3935.txt

[3] **Ando, R.** *Juniper seen regaining core router share in 2008*, 2007
Available at: http://www.reuters.com/article/ companyNewsAndPR/idUSN1333533920070613

[4] **Austein, R., Wijnen, B.** *RFC4071: Structure of the IETF Administrative Support Activity (IASA)*, 2005
Available at: http://www.ietf.org/rfc/rfc4071.txt

[5] **Barrett, D. J., Silverman, R. E., Byrnes, R. G.** *SSH: The Secure Shell The Definitive Guide*, 2005

[6] **Besen, S. M., Farrell, J.** *Choosing How to Compete: Strategies and Tactics in Standardization*, 1994
The Journal of Economic Perspectives, Vol. 8, No. 2 (Spring, 1994)

[7] **Blind, K.** *Explanatory Factors For Participation In Formal Standardisation Processes: Empirical Evidence At Firm Level*, 2006
Economics of Innovation and New Technology, 2006, Vol. 15(2), March

[8] **Blind, K., Gauch, S.** *Trends in ICT Standards in European Standardisation Bodies and Standards Consortia*, 2005

[9] **Blum, U** *Lessons from the Past: Public Standardization in the Spotlight*, 2005
International Journal of IT Standards & Standardization Research; Jan-Jun 2005

[10] **Bolin, S.** *The Nature and future of ICT standardisation*, 2003

[11] **Braden, B.** *Overview of the IETF*, 1999

[12] **Bradner, S.** *IETF Structure and Internet Standards Process*, 2008
Available at: http://tools.ietf.org/group/edu/attachment/wiki/IETF72/72newcomers.ppt?format=raw

[13] **Bradner, S.** *RFC2026: The Internet Standards Process – Revision 3*, 1996
Available at: http://www.ietf.org/rfc/rfc2026.txt

[14] **Bradner, S.** *RFC2418: IETF Working Group Guidelines and Procedures*, 1998
Available at: http://www.ietf.org/rfc/rfc2418.txt

[15] **Bradner, S.** *RFC4748: RFC 3978 Update to Recognize the IETF Trust*, 2006
Available at: http://www.ietf.org/rfc/rfc4748.txt

[16] **Brusse, B.** *COPRAS: Improving the Interface Between Research and Standardization*, 2005
International Journal of IT Standards & Standardization Research; Jan-Jun 2005

[17] **Cargill, C.** *Standardization And The Information Technology Industry*, 2001

[18] **Cargill, C.** *Standardization: The Necessary Luxury*, 2004

[19] **Cargill, C., Bolin, S.** *Standardization: A Failing Paradigm*, 2004

[20] **Carpenter, B.** *The IETF Process: an Informal Guide*, 2008

[21] **Carpenter, B.** *RFC2850: Charter of the Internet Architecture Board (IAB)*, 2000
Available at: http://www.ietf.org/rfc/rfc2850.txt

[22] **Carpenter, B., Baker, F., Roberts, M.** *RFC2860: Memorandum of Understanding Concerning the Technical Work of the Internet Assigned Numbers Authority*, 2000
Available at: http://www.ietf.org/rfc/rfc2860.txt

[23] **Cerf, V. G.** *RFC1160: The Internet Activities Board*, 1990
Available at: http://www.ietf.org/rfc/rfc1160.txt

[24] **Cerf, V. G., Kahn, R. E.** *A Protocol for Packet Network Intercommunication*, 1974

[25] **Chiao, B., Lerner, J., Tirole, J.** *The Rules of Standard Setting Organizations: An Empirical Analysis*, 2005

[26] **Cisco** *Cisco: Corporate Overview*, 2009
Available at: http://newsroom.cisco.com/dlls/ corpinfo/corporate_overview.html

[27] **Cisco** *Cisco's Corporate Overview Presentation*, 2009
Available at: http://newsroom.cisco.com/dlls/2009/ ekits/PublicCorporateOverview_BlackBackground_Eng_2009Q3.pdf

[28] **Cohen, W. M., Levinthal, D. A.** *Innovation and Learning: The Two Faces of R&D*, 1989

[29] **COPRAS** *Standardization Guidelines for IST research projects interfacing with ICT standards organizations*, 2007

[30] **Crocker, D.** *Making Standards the IETF way*, 1993
StandardView Vol. 1, No. 1, September 1993

[31] **Crocker, D. H.** *RFC0822: STANDARD FOR THE FORMAT OF ARPA INTERNET TEXT MESSAGES*, 1982
Available at: http://www.ietf.org/rfc/rfc822.txt

[32] **Crocker, S.** *RFC0001: Host Software*, 1969
Available at: http://www.ietf.org/rfc/rfc1.txt

[33] **Daigle, L.** *RFC4052: IAB Processes for Management of IETF Liaison Relation*, 2005
Available at: http://www.ietf.org/rfc/rfc4052.txt

[34] **David, P. A., Greenstein, S.** *The Economics of Compatibility Standards: An Introduction to Recent Research*, 1990
Economics of Innovation and New Technologies, 1990, Vol. 1

[35] **David, P. A., Shurmer, M.** *Formal standards-setting for global telecommunications and information services*, 1996
Telecommunication Policy Vol. 20 1996

[36] **de Vries, H. J., Verheul, H., Willemse, H.** *Stakeholder Identification In It Standardization Processes*, 2003

[37] **Denning, P. J., Hearn, A., Kern, C., W.** *HISTORY AND OVERVIEW OF CSNET*, 1983

[38] **DIN German Institute for Standardization** *Economic benefits of standardization. Summary of results. Final report and practical examples*, 2000

[39] **Egyedi, T.** *Why Java Was - Not - Standardized Twice*, 2001
Proceedings of the 34th Hawaii International Conference on System Sciences - 2001

[40] **EICTA Members** *EICTA Interoperability White Paper*, 2004

[41] **EU** *European ICT standardisation policy at a crossroads: A new direction for global success. The Way Forward. A discussion document for the Open meeting on 12 February 2008*, 2008

[42] **European Committee for Standardization (CEN)** *ICT Standards Consortia List, 14th edition, October 2008*, 2008
Available at: http://www.cen.eu/cenorm/sectors/ sectors/isss/consortia/index.asp

[43] **European Committee for Standardization (CEN)** *Standardisation & Research (STAR) Executive Summary*, 2004

[44] **Farrell, J.** *Choosing the Rules for Formal Standardization*, 1996

[45] **Farrell, J., Saloner, G.** *Standardization, compatibility, and innovation*, 1985
The RAND Journal of Economics, Vol. 16, No. 1, (Spring, 1985)

[46] **Farrell, J., Saloner, G.** *Coordination Through Committees and Markets*, 1988
The RAND Journal of Economics, Vol. 19, No. 2, (Summer, 1988)

[47] **Fedora** *Objectives of Fedora*, 2008
Available at: http://fedoraproject.org/wiki/Objectives

[48] **Freeman, R. E.** *Strategic Management: A Stakeholder Approach*, 1984

[49] **Funk, J. L.** *Competition between regional standards and the success and failure of firms in the world-wide mobile communication market*, 1998
Telecommunications Policy Vol. 22

[50] **Funk, J. L., Methe, D. T.** *Market- and committee-based mechanisms in the creation and diffusion of global industry standards: the case of mobile communication*, 2001

[51] **Gauch, S.** *INTEREST: integrating research and standardisation*, 2007
J. of IT Standards & Standardization Research, 5(2), 50-52, July-December 2007

[52] **Gaynor, M., Bradner, S.** *The Real Options Approach to Standardization*, 2001

[53] **Google Scholar** *About Google Scholar*, 2009
Available at: http://scholar.google.com/intl/en/scholar/about.html

[54] **Greenstein, S.** *Markets, Standards, and the information infrastructure*, 1993

[55] **Gross, P.** *Proceedings of the 1st IETF*, 1986
Available at: http://www.ietf.org/proceedings/prior29/IETF01.pdf

[56] **Haimowitz, J., Warren, J.** *Economic Value of Standardization*, 2007

[57] **Harris, S.** *RFC3160: The Tao of IETF - A Novice's Guide to the Internet Engineering Task Force*, 2001
Available at: http://www.ietf.org/rfc/rfc3160.txt

[58] **Hauben, M.** *History of ARPANET*, 1994

[59] **Hoe, N. S.** *Free/Open Source Software - Open Standards*, 2006

[60] **Hoffman, P., Harris, S.** *RFC4677: The Tao of IETF: A Novice's Guide to the Internet Engineering Task Force*, 2006
Available at: http://www.ietf.org/rfc/rfc4677.txt

[61] **Hovey, R., Bradner, S.** *RFC2028: The Organizations Involved in the IETF Standards Process*, 1996
Available at: http://www.ietf.org/rfc/rfc2028.txt

[62] **Huitema, C., Postel, J., Crocker, S.** *RFC1796: Not All RFCs are Standards*, 1995
Available at: http://www.ietf.org/rfc/rfc1796.txt

[63] **Hurd, J., Isaak, J.** *IT Standardization: The Billion Dollar Strategy*, 2005
International Journal of IT Standards & Standardization Research; Jan-Jun 2005

[64] **IANA** *IANA - About the Internet Assigned Numbers Authority*, 2009
Available at: http://www.iana.org/about/

[65] **ICT SHOK** *ICT SHOK Future Internet Programme of Finland*, 2009
Available at: http://www.futureinternet.fi/

[66] **ICT SHOK** *ICT SHOK Future Internet Programme: Programme Plan*, 2008

[67] **ICT SHOK** *ICT SHOK Future Internet Programme: Research Agenda*, 2007

[68] **ICT Standards Board member organisations** *ICTSB Critical issues in ICT Standardization*, 2005

[69] **ICTSB** *ICT Standardization*, 2008
Available at: http://www.ictsb.org/ICT_standardization.htm

[70] **IEC, ISO** *ISO/IEC Directives, Part 1. Procedures for the technical work*, 2008

[71] **IEC, ISO, ITU-T and UN/ECE** *Proposal from IEC, ISO, ITU-T and UN-ECE on International Standards for the Information Society at WSIS 2003*, 2003

[72] **IEEE** *IEEE Membership*, 2009
Available at: http://www.ieee.org/web/membership/home/index.html

[73] **IESG** *IESG Statement on Spam Control on IETF Mailing Lists*, 2008
Available at: http://www.ietf.org/IESG/STATEMENTS/spam-control-policy.html

[74] **IETF** *IETF Administrative Support Activity (IASA)*, 2009
Available at: http://iaoc.ietf.org/

[75] **IETF** *The Internet Engineering Steering Group (IESG)*, 2009
Available at: http://www.ietf.org/glossary.html#IESG

[76] **IETF** *IETF Proceedings Directory*, 2009
Available at: http://www.ietf.org/proceedings_directory.html

[77] **Information Society Technologies Advisory Group** *ISTAG Report on Shaping Europe's Future Through ICT*, 2006

[78] **Integrating Research and Standardisation (INTEREST)** *A Guide to Standardisation for R&D Organisations and Researchers*, 2007

[79] **International Telecommunication Union (ITU)** *CCITT - 50 YEARS OF EXCELLENCE: 1956-2006*, 2006

[80] **IRTF** *Internet Research Task Force*, 2009
Available at: http://www.irtf.org/

[81] **ISO/IEC JTC 1** *The Delta Value of JTC 1 as an IS Provider*, 2002

[82] **ISO/IEC JTC 1** *JTC 1 Long Term Business Plan (LTBP)*, 2002

[83] **ISOC** *ISOC Annual report 2007*, 2007

[84] **Iversen, E. J.** *Raising Standards: Innovation and the emerging global standardization environment for ICT*, 2000

[85] **Jakobs, K.** *A Closer Look At The Internet's Standards Setting Process*, 2003

[86] **Jakobs, K.** *Information Technology Standards, Standards Setting and Standards Research*, 2003

[87] **Jakobs, K.** *ICT Standardization Co-ordinating the diversity*, 2008

[88] **Jakobs, K., Procter, R., Williams, R.** *The Making of Standards: Looking Inside the Work Groups*, 2001
IEEE Communications Magazine April 2001

[89] **Kantor, A.** *Technology standards increase competition, innovation*, 2006
Available at: http://www.usatoday.com/tech/columnist/ andrewkantor/2006-03-02-standards_x.htm

[90] **Katz, M. L., Shapiro, C.** *Network Externalities, Competition, and Compatibility*, 1985
The American Economic Review, Vol. 75, No. 3, (Jun., 1985)

[91] **Keegan, M.** *Rout in the Router Market*, 2008
Available at: http://online.barrons.com/article/SB122643998617818539.html

[92] **Kolkman, O.** *IETF response to UN enhanced cooperation*, 2008

[93] **Kolkman, O.** *Words from the IAB Chair*, 2009
IETF Journal Volume 4, Issue 3 February 2009

[94] **Kolodziej, S.** *Egos, Infighting & Politics: Standards Progress Bogged Down*, 1988
Computerworld; Sep 7, 1988; 22, 36;

[95] **Konidaris, S., Johnston, P.** *Final Report on ACTS (1994-1998)*, 2000

[96] **Kristiansen, E. G.** *R&D in the Presence of Network Externalities: Timing and Compatibility*, 1998
The RAND Journal of Economics, Vol. 29, No. 3 (Autumn, 1998)

[97] **Kroes, N.** *Being open about standards*, 2088

[98] **Leiner, B. M., et al.** *A Brief History of the Internet*, 2003

[99] **Leiponen, A.** *Clubs And Standards: The Role Of Industry Consortia In Standardization Of Wireless Telecommunications*, 2005

[100] **Leiponen, A.** *Competing Through Co-operation: Standard-Setting in Wireless Telecommunications*, 2006

[101] **Lerner, J., Tirole, J.** *A Model of Forum Shopping, with Special Reference to Standard Setting Organizations*, 2004

[102] **Liebowitz, S. J., Margolis, S. E.** *Network Externalities (Effects)*, 1998

[103] **Liebowitz, S. J., Margolis, S. E.** *Path Dependence, Lock-In, and History*, 1995
Journal of Law Economics & Organization 1995; 11: 205-226

[104] **Low, M.** *BSI British Standards: The Benefits of Standardization and the NSSF*, 2005

[105] **Maher, M.** *An Analysis of Internet Standardization*, 1998
Virginia Journal of Law and Technology Volume 3 - 1998

[106] **McPherson, D., Huston, G.** *Defining the Role and Function of IETF Protocol Parameter Registry Operators*, 2008
Available at: http://tools.ietf.org/id/draft-iab-iana-04.txt

[107] **Merritt, E.** *The Gnuplot Fit Command Manual*, 2009
Available at: http://www.gnuplot.info/docs/node82.html#fit

[108] **Mitchell, R. K., Agle, B. R., Wood, D. J.** *Toward A Theory Of Stakeholder Identification And Salience: Defining The Principle Of Who And What Really Counts*, 1997

[109] **MySQL** *Why MySQL?*, 2008
Available at: http://mysql.com/why-mysql/

[110] **Narten, T.** *Bringing New Work to the IETF*, 2007

[111] **Narten, T.** *RFC5434: Considerations for Having a Successful Birds-of-a-Feather (BOF) Session*, 2009
Available at: http://www.ietf.org/rfc/rfc5434.txt

[112] **National Science Foundation (NSF)** *The Internet changing the way we communicate*, 2001

[113] **Organisation for Economic Co-operation and Development (OECD)** *ICT STANDARISATION IN THE NEW GLOBAL CONTEXT*, 1996

[114] **Pelkmans, J.** *The GSM standard: explaining a success story*, 2001
Journal of European Public Policy 8:3 Special Issue

[115] **Postel, J.** *RFC1591: Domain Name System Structure and Delegation*, 1994
Available at: http://www.ietf.org/rfc/rfc1591.txt

[116] **Postel, J., Reynolds, J.** *RFC2223: Instructions to RFC Authors*, 1997
Available at: http://www.ietf.org/rfc/rfc2223.txt

[117] **Research in Advanced Communications in Europe (RACE)** *The Race Programme*, 1991

[118] **Resnick, P.** *RFC2822: Internet Message Format*, 2001
Available at: http://www.ietf.org/rfc/rfc2822.txt

[119] **Resnick, P.** *RFC5322: Internet Message Format*, 2008
Available at: http://www.ietf.org/rfc/rfc5322.txt

[120] **Reynolds, J., Postel, J.** *RFC1000: THE REQUEST FOR COMMENTS REFERENCE GUIDE*, 1987
Available at: http://www.ietf.org/rfc/rfc1000.txt

[121] **RFC Editor** *XML schema to describe the complete RFC Index*, 2008
Available at ftp://ftp.isi.edu/in-notes/rfc-index.xsd

[122] **Rutkowski, A. M.** *Today's Co-operative Competitive Standards Environment For Open Information and Telecommunication Networks and the Internet Standards-Making Model*, 1994

[123] **Sanders, T. R. B.** *THE AIMS AND PRINCIPLES OF STANDARDIZATION*, 1972

[124] **Schulzrinne, H.** *A Brief History of the Internet Advisory / Activities / Architecture Board*, 2002

[125] **Shah, R., Kesan, J. P.** *An Empirical Examination of Open Standards Development*, 2008

[126] **Shankar, V.** *Cisco's Router Market Share Drops Most Since 2001*, 2008
Available at: http://www.bloomberg.com/apps/news ?pid=20601103&sid=a7byJNgbW5hQ&refer=us

[127] **Simcoe, T.** *Delays and de Jure Standards: What Caused the Slowdown in Internet Standards Development*, 2004

[128] **Simcoe, T. S., Rysman, M.** *Measuring the performance of standard setting organizations*, 2006

[129] **SSH** *SSH Annual Report 2007*, 2007

[130] **Swann, G. M. P.** *The Economics of Standardization*, 2000

[131] **Techapalokul, S., Alleman, J.H., Chen, Y.** *Economics of standard: a survey and framework*, 2001

[132] **TEK** *The Finish Engineering Award 1981-2008*, 2008
Available at: http://www.tek.fi/ci/pdf/teknologia/SITP2008english.pdf

[133] **Tekniikka & Talous** *Nanoluokan diagnostiikka toi Insinöörityöpalkinnon*, 29.5.2009

[134] **Tekniikka & Talous** *Suomen Suurimmat Tutkijat*, 12.6.2009

[135] **TIVIT** *Producing Innovations for the ICT Industry*, 2009
Available at: http://www.tivit.fi/en/

[136] **UK Department of Trade and Industry (DTI) and British Standards Institution (BSI)** *The Empirical Economics of Standards*, 2005
DTI Economics Paper No. 12, 2005

[137] **Updegrove, A.** *STANDARDS STAKEHOLDERS: WHO SHOULD, AND WHO DOES, SET STANDARDS?*, 2006

[138] **Updegrove, A.** *ICT STANDARD SETTING TODAY: A SYSTEM UNDER STRESS*, 2007

[139] **Updegrove, A.** *The Essential Guide to Standards*, 2007

[140] **USPTO** *US Published Application Full-Text Database Boolean Search*, 2009
Available at: http://appft1.uspto.gov/netahtml/PTO/search-bool.html

[141] **USPTO** *Patent Full-Text and Full-Page Image Database*, 2009
Available at: http://patft.uspto.gov/

[142] **USPTO** *US Patent Full-Text Database Boolean Search*, 2009
Available at: http://patft.uspto.gov/netahtml/PTO/search-bool.html

[143] **Van Eecke, P., Fonseca, P. P., Egyedi, T.** *EU Study on the specific policy needs for ICT standardisation*, 2007

[144] **van Wegberg, M.** *Standardization and Competing Consortia: The Trade-Off Between Speed and Compatibility*, 2004
International Journal of IT Standards & Standardization Research; Jul-Dec 2004

[145] **Vercoulen, F., van Wegberg, M.** *Standard Selection Modes in Dynamic, Complex Industries: Creating Hybrids between Market Selection and Negotiated Selection of Standards*, 1998

[146] **Vinaja, R., Raisinghani, M. S.** *A review of the standards making process in the telecommunications industry: challenges and potential solutions*, 2001

[147] **Weinrib, A., Postel, J.** *RFC2014: IRTF Research Group Guidelines and Procedures*, 1996
Available at: http://www.ietf.org/rfc/rfc2014.txt

[148] **Weiss, M. B. H., Toyofuku, R. T.** *Free-Ridership in the Standards-Setting Process: The Case of 10BaseT*, 1996
StandardView Vol. 4, No. 4, December/1996

Appendix A

This appendix describes the technical details of the database design and the computer programming work done as a part of this thesis. We look at the development environment and development tools, the database structure, and the way the data is constructed and processed using the PHP scripts.

A Computer Programming and Database Design

A.1 Development Environment

In order to set up the project database we need to make a number of technical choices. We have to fix our development environment and a number of programs that enable us to do the actual programming and testing. This being an academic research task, we outline the solution stack to cover only such open-source software that does not impose a economic burden on our work.

Derived from the goals of our task, we need at least the following components to set up the database, extract the data and create an user interface to the database:

- Operating system
- Database management system
- Document repository
- Programming environment
- User interface

To solve the issue we use a bundle of free software, often referred to as the LAMP solution stack. LAMP being an acronym for the bundle of the Linux operating system, The Apache HTTP server, the MySQL database management system, and, in our case the PHP programming language.

Fedora Core Linux was chosen as the operating system for our development environment. The Fedora Core distribution comes exclusively with free and open source software. It is also a distribution that rapidly adopts new releases of the packages it is composed of. [47] Hereby, this distribution is ideal for our development environment as we are able to use current versions of all components.

MySQL was chosen as the DBMS because it has the features needed for this task, it is easy to use and it performs well. Especially Full-Text indexing and Full-Text search functions in addition to regular expression capabilities provide nice functionality in our database. MySQL is also the world's most popular open source database and it runs on a large number of different platforms. [109]

All the computer scripts needed to create the project database are written using the PHP language. The MySQL database interface is solid and file processing is fluent in PHP. SimpleXML functions provided in the PHP environment are also useful, as parsing XML files is a part of the task. Perl-Compatible Regular Expression Functions which we use when extracting certain parts of the RFCs and I-Ds are also needed.

The graphical user interface for the project database will be developed using the above-mentioned components. We use the PHP programming language, filled in with HTML mark-up and JavaScript scripting, to build a dynamic web user interface to the database. The MySQL client works as the specified command line user interface.

All the charts and histograms presented in this thesis are also produced using purpose made computer scripts and readily available open source tools. The data is extracted from the database using custom queries and then fed into the command-line driven data plotting utility Gnuplot. With this configuration we can make a broad variety of queries into the database and get the graphical output effectively in real-time.

A.2 Necessary Documents and Tools

To reach our set goal for extracting meta-data from the IETF working document there are certain documents and tools that prove to be useful for our task. The *RFC Editor* provides us with the *RFC index* file and through the *Internet-Drafts Database Interface* we are able to obtain the *Internet-Drafts Status Summary* and the *Index of Current Internet Draft Abstracts*.

As a starting point for the RFC meta-data extraction we choose the RFC index document. The RFC index exists in two formats, a Text version and an XML version. We chose to use the XML version whereas it is superior to the Text version both in structure and information value. The XML structure is defined using an XML Schema Definition. [121] The data available in the RFC index is collected and the body of the RFCs are also pulled from the IETF repository and inserted into the project database.

The Internet-Drafts are officially not archived the same way the RFCs are. However, using the *IETF Documents* listing on the IETF Tools pages *http://tools.ietf.org /html/recent-drafts* allowed us to gather a much larger number of I-Ds than currently are available in the Internet-Drafts repository. The *IETF Documents* listing only provides us with the file-name of the I-D and a URL to the body of the document. The IETF data-tracker *https://datatracker.ietf.org/drafts* is therefore use as an additional aid to collect further meta-data regarding the draft documents. The IETF Tools pages allows us to access over 52000 active and expired I-Ds. The I-Ds are also picked out from the IETF server and inserted as such into the database enabling us to process the further later on.

A.3 Database Structure

The goal is to design a database generic enough to include RFCs, I-Ds, and possibly, in the future, also documents and their meta-data from other ICT standardization bodies.

We were able to use the XML Schema Definition for the RFC index as an initial illustration for the database structure. The RFC index covers our problem space fairly well and is also quite a comprehensive source of meta-data itself. The XML Schema Definition documents both the structure and data types of the RFC index.

Based on the RFC index XML Schema Definition and concept prototyping, we opted for a relational database model that is made up of five primary tables containing document meta-data and the actual documents, see Figure A1. In addition to this we have several tables containing data covering other aspects of the IETF work, see Figure A2. The database is not normalized as it will be used for data mining and there will not be any updates nor transactions.

The main table is the document_header table. It stores the two most important data items of each document, the standardization body that published the document (organization) and the name the standardization body gave that specific document (document_id). These two data items makes up the set that uniquely identifies each document. The *organization,document_id* unique key in the document_header table signifies this. Each row in the document_header table is also given a numeric key, the column *id*, that functions as the primary key for the relations.

The *author*, and *document_content* tables are all linked to the *document_header* table's primary key using their respective foreign key *doc_id*.

Extra information about the authors are stored in the *author_info* table. The *author_info* table is linked to the *document_header* table's primary key using the foreign key *author_id*. The actual IETF documents are stored as is in the *document_content_body* table.

A.4 Programming

Aggregating all the meta-data needed for the analyzes in this thesis is carried out with PHP scripts. The main task is to find and parse author meta-data from the *Author's Address Section*. This section is required to be included at the very end of each RFC. It should contain at least the following information about the authors: name, postal address, the telephone number, and the email address. The meta-data extraction is done with the *Author-info parser*, created as a part of this research task. The *Author-info parser* relies much on a set of Regular Expressions and a number of custom rules and exception stored in the *parser_help* table.

From a technical point of view the main task lies in finding the beginning and end of the *Author's Address Section*, then breaking it into subsections per author, so that one author's information will not be mixed with another author's. Establishing the

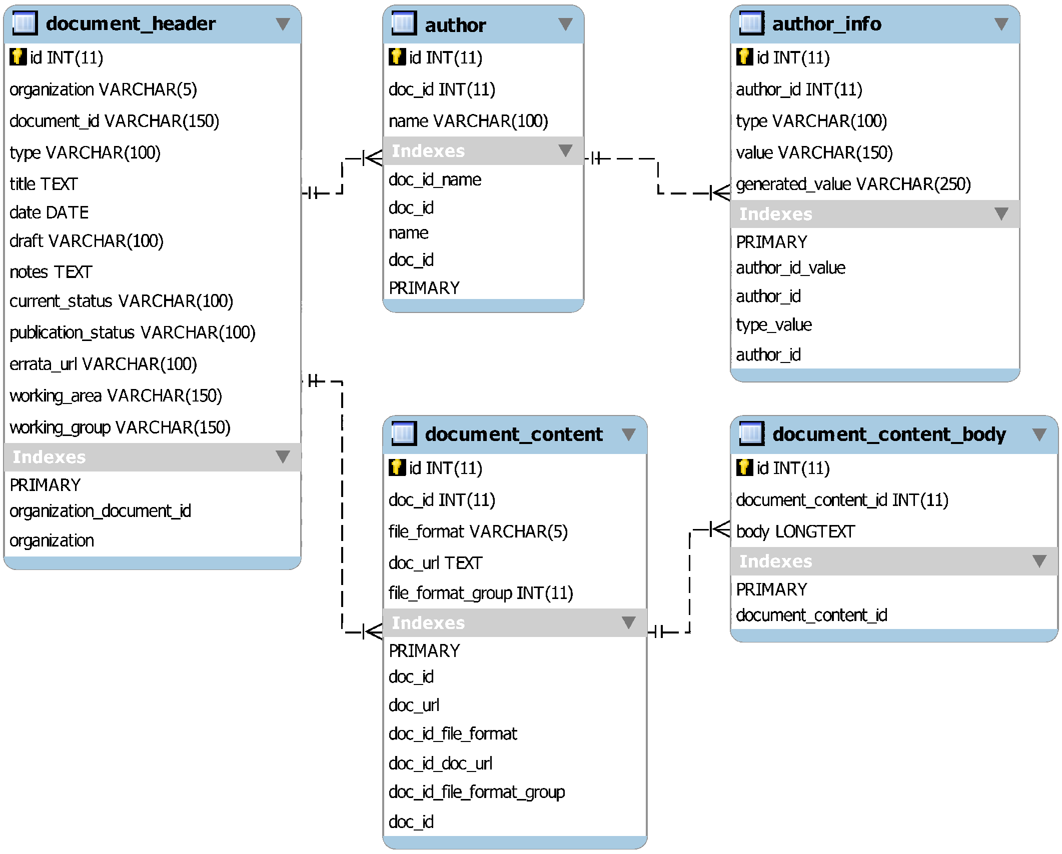


Figure A1: Database structure

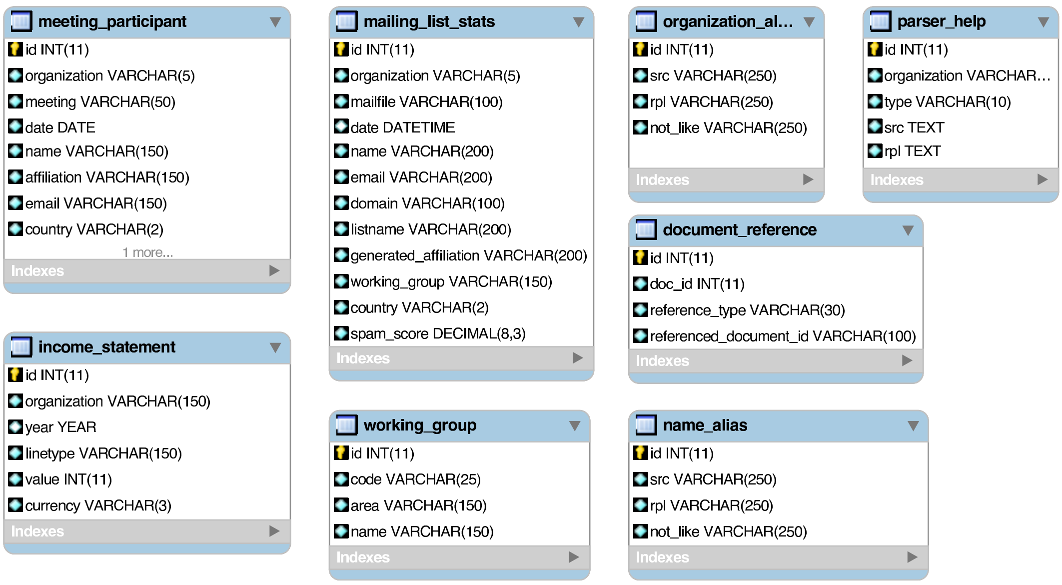


Figure A2: Database structure

subsections is done using Regular Expressions to match certain patterns. This is not trivial, as the information is in natural language and not completely structured.

All the rows of parsed author meta-data also need to be labeled so that their meaning is clarified. The labeling is also done using regular expression pattern matches. The following labels are used.

- o NAME
- o ORGANIZATION
- o EMAIL
- o X.400
- o URI
- o PHONE
- o FAX
- o COUNTRY

The final task regarding meta-data extraction is analyzing non complete sets of author meta-data and deciding if they can be enhanced. Methods of enhancement can be e.g. to copy missing parts of a certain meta-data-set from another more complete set of the same author's data. Copying parts of an author's meta-data between documents involves the risk of copying incorrect information in the case where the author's address information actually has changed. By specifying time frames from how far meta-data can be copied the risks of copying incorrect meta-data can be minimized.

Id addition to the IETF working documents we also gather data on meeting attendance and mailing list activities. Meeting attendee lists are available from the IETF 29 meeting onward. There are however five different types of attendee lists so there are again some additional matters to be considered. The IETF meeting attendee data is inserted into the *meeting_participant* table. Finally all the emails from the IETF mailing lists are also processed and inserted in to the *mailing_list_stats* table.

As a part of the data mining process we also collected additional data regarding the RFCs. We gathered data on RFC cross-references, USPTO patent applications, USPTO issued patents, Google Scholar academic references, and also Google Search hits. The USPTO database is queried with the following types of search strings for every RFC:

```
http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF
&u=/netahtml/PTO/search-adv.htm&r=0&p=1&f=S&l=50
&Query="RFC+1234"+OR+RFC1234&d=PTXT
```

```
http://appft1.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF
&u=/netahtml/PTO/search-adv.html&r=0&p=1&f=S&l=50
&Query="RFC+1234"+OR+RFC1234&d=PG01
```

The Google scholar and the Google search pages are it the same way queried with the below-mentioned search strings for every RFC:

```
http://scholar.google.com/scholar?hl=en&lr=&q="RFC+1234"+OR+RFC1234
```

```
http://www.google.com/search?hl=en&safe=off&q="RFC+1234"+OR+RFC1234
```

Finally, to be able to analyze the data in the project database we needed to standardize some of the values in the data. Especially the author names and organization names need to be processed. E.g. the organization name *Motorola* was documented by RFC and I-D authors in 36 different ways. This presents and hindrance to successfully group IETF contributions per affiliated organizations. The matter was however solved using custom rules removing prefixes and suffixes and thereby standardizing the writing of organization names.

As a result of our data mining task we acquired a total of 59245 IETF working documents into the database. That is, 53911 Internet-Drafts and 5334 RFCs. We parsed 148119 lines of author names and altogether 897856 lines of author information inserted in to the *author* and *author_info* tables respectively. The total size of the *document_content_body* table containing the actual bulk of the working documents adds up to roughly three gigabytes of data. In addition to the working document related data, we also collected 67772 lines of meeting attendee data and 1401882 lines of IETF mailing list data.

The total amount of PHP scripts created as a part of this thesis work is 21. The scripts are divided in to groups with distinct functions. Some scripts are responsible for fetching the documents into the database whereas others are put in place to parse author data from the documents. The total sum of lines of code is approximately 9700.