THE DEGREE OF DIGITALIZATION OF THE INFORMATION OVERFLOW
A case study

Turo Kilpeläinen
Agora Center, University of Jyväskylä, P.O. Box 35 FIN-40014, University of Jyväskylä, Finland

Pasi Tyrväinen
Faculty of Information Technology, University of Jyväskylä, P.O. Box 35 FIN-40014, University of Jyväskylä, Finland

Keywords: Digital media, measuring organizational communication, genre theory, measures for digitalization

Abstract: The degree of digitalization in organizations has increased remarkably. This trend will continue if the so called natural laws of information technology are veracious. At the same time the format of communicated information has shifted from traditional face-to-face and analogue communication to digital communication forms, such as digital documents. Because information is increasingly available in digital form on one hand for example the duplication and forwarding of email messages and attachments gets easier, which may easily lead to information overflow. On the other hand digitalization increases productivity, improves quality and reduces costs. As the ability of humans to adopt information has not developed at the same pace as information and communication technology (ICT), it is interesting to see whether the degree of digital communication correlates with the total amount of communication in organization. In this paper we tested this hypothesis in an industrial organization using genre-based measurement method to gather data on communication flows. The results show, that the correlation between the degree of digital communication and the amount of total communication can be seen in some degree.

1 INTRODUCTION

The Big Bang of the digital world was in 1949, when the first business oriented information system, LEO I (cf. Aris, 2000; Land, 2000), was launched in Britain. After that the speed of development within the field of information systems, and later digital media, has been enormous. By the start of the 1970s, it had become impossible for any commercial enterprise of over 250 employees to function competitively in the United States without using computers (Cortada, 1996). Nowadays computers and applications have become a crucial part of organizations' infrastructure to survive in the hectic and turbulent business world (Morgan, 1988; Scott Morton, 1991). Manual punch cards and tabulation machines have been superseded by computers to automate human stages of work and accordingly human errors have diminished. In short, the presence of the computer plays a determinative part at the heart of personal and institutional economic activity (Cortada, 1997).

The pace of ICT development has been so immense that it has been a key target for research and different kinds of predictions. Maybe the most famous and tenable law was developed by Gordon Moore in 1965. He found out in his observations that the number of transistors per square inch on integrated circuits had doubled every year since the integrated circuit was invented (Moore, 1975). Moore also predicted that this trend would continue for the foreseeable future. In subsequent years the pace has slowed down a bit, but data density has doubled approximately every 18 months. This led to the current definition of Moore's Law. Most experts, including Moore himself, expect Moore's Law to hold until the year 2010 (Schaller, 1997).

At the same time, when business world has converted to the digital era, the tremendous changes occurring in the way current organizations communicate have been recognized (Yates and

Kilpeläinen T. and Tyrväinen P. (2004).
THE DEGREE OF DIGITALIZATION OF THE INFORMATION OVERFLOW - A case study.
In *Proceedings of the Sixth International Conference on Enterprise Information Systems*, pages 367-374
DOI: 10.5220/0002635403670374

Orlikowski, 1992). This has created many opportunities, which are not exploited in the best way in all cases, as well as many draw-backs for the ones adopting new technology without proper consideration of the impacts. For example the use of electronic and digital communication media is increasing the number of options for distributed development teams to coordinate their activities and to keep knowledge up-to-date (Sosa et al., 2002). A salient point is that the widespread use of information technology is reducing the traditional reliance on face-to-face communication (Sproull and Kiesler, 1991). Digital documents also enable increasing functionalities (Shepherd and Watters, 1999) as well as increased productivity, improved quality and reduced costs (Allen, 1991) and overall drastic improvements in organizational performance (Davenport, 1993; Scott Morton, 1991; Stalk et al. 1992), although the economical benefits of ICT investments are not fully incontrovertible (van Grembergen, 2002). Nonetheless digital convergence and benefits it brings are the most valuable assets that digitalization produces. For example, distinct publication channels and publication media are able to handle consistent technical infrastructure and to publish through different channels. In fact, convergence is the underlying enabler of the digitalization phenomenon. It is obvious that digital media has increasingly reshaped organizational communication (Yates, 1989) and will bring novel means for organizational communication (Tyrväinen and Päivärinta, 2003). Presumably the impact will go from strength to strength and even escalate in the future.

Nowadays the major proportion of organizational information resides in digital documents (Blair 2002; Päivärinta and Tyrväinen, 1998). This digitalization trend is not only adding productivity, but has also generated new problems and added to the impact of existing ones, such as the information overflow, which arises from effortless availability of all kinds of information. The ease of information distribution, for example by over-distributing or forwarding received mails to many people, can impair organizational communication by overloading the persons receiving the data with irrelevant or secondary information. This action does not increase communicators' knowledge in any way but it increases the amount of total communication in the organization.

Both the assumed productivity improvement achieved by digital communication and the ease of communication with digital media suggest that the degree of digital communication correlates with the total amount of communication. Further, as the digitalization of both seems to enable extra communication in organizations as well as provoke it, we hypothesise, that there exists a limit for the volume of communication, beyond which most of the additional communication takes place through digital media.

There exists little published research on this topic as well as little research on the communication volumes in organizations in general. However, comparing the figures from two recent case studies (Tyrväinen, 2003; Tyrväinen et al., 2003) supports the hypothesis. In those studies the total communication volume of two organizations were both measured in pages per person day, that was defined as the "amount of information equal to a view of the size of a visual letter / A4 page (Tyrväinen, 2003)". This same unit was applied for all communication including verbal, paper-based documents, and data stored in information systems. In a unit of a high-tech company using digital media for 56% of communication the employees communicated on an average 92 pages per day, while in a university faculty using digital media for 23% of communication the figure was about 20% smaller, i.e. 74 pages per day. Although comparing two very different kinds of organizations does not provide very strong evidence for our hypothesis, these two cases provide us with the means to approach the problem.

This paper reports a case study in which we test our hypothesis within a single organization. Instead of comparing organizations with hundred of employees, we compare communication of smaller groups of persons based on their roles within an organization. The remainder of the paper is organized as follows: section 2 describes the target organization, the research process, and the results of the case study in an industrial organization, where the amount of communication was measured. The relevancy of the results is discussed in section 3. Section 4 summarizes the results, draws conclusions and proposes paths for further research.

2 THE CASE STUDY

2.1 The Target Organization and the Research process

The case organization is an independent unit of a multinational corporation in process industry, involving thousands of employees worldwide. The target unit of the study was a specialized experimental unit within this organization evaluating the feasibility to produce newly designed products

with specific production line settings. The total head-count of employees in the target unit was about 20, including five roles or groups of persons.

The communication of the target organization was analyzed in summer 2003 as a part of another research activity. We followed the measurement method described in the two case studies referred to above (Tyrväinen 2003; Tyrväinen et al., 2003). The measurement method is a derivate of a genre-based method used for information systems planning and content analysis (cf. Päivärinta et al., 2001; Karjalainen et al., 2000). With the method, an information resource is conceptually structured by genres of organizational communication (Nunberg, 1997; cf. Yates and Orlikowski, 1992), i.e. typified, enacted and shared purposes and forms of documents or other communication occurring in recurrent situations.

The research process was as follows: After defining the scope of the target processes we invited the key persons to a genre identification session. Six persons participated in the identification session in one group in according to the method. Persons were selected to cover all roles (employee groups) in the target unit. As an employee group we refer to a role that an individual may play in the organization (cf. Scott, 2002). In the sessions we first collected the internal roles and external roles communicating with each other ending up with 14 roles, referred to as PUI entities in the method (Producers and Users of Information). Next, a total 64 genres were identified and named during the session using the diagonal matrix technique (Saaren-Seppälä, 1997). In this case a genre is considered as a unit of analysis by which all the organizational communication, both the elements of genres and social context of genre use, can be identified, including information flows through information systems, documents, verbal communications etc. (cf. Yoshioka et al., 2001). In this way the genre repertoire (Orlikowski and Yates, 1994) was defined.

After this three participants of the session were instructed to identify metadata for genres concerning his/her responsibilities. Thus each of the three persons identified metadata for about 20 genres in separate metadata collection sessions with the researcher during the couple of days following the session. The metadata collected included the metadata values for each genre, for the elements defined in the measurement method. I.e. for each genre, we collected:

- The number of unique communication instances (UI) per year.
- The number of copies of communication instances per instance.
- The volume or average size of the communication instances in "Pages".
- The communication forms / media used for the communication of these genre instances in according to a categorization of communication forms (CCF) separating verbal, analogue (i.e. paper), and digital communication as well as their sub-categories ending up with a total of 10 elementary categories (c.f. Tyrväinen 2003, or Tyrväinen and Päivärinta 2003 elaborating the process description further.)

In addition, we collected other metadata related to the overall research interests of the data mining research activity in the target unit adding up the total number of metadata elements to be filled in per genre to be 19.

After these steps the resulting data were analyzed from multiple viewpoints, in accordance with the main business processes, in general, and to meet the needs for other parts of research project. The amount of communication was analyzed per each employee role.

2.2 Overall Results

Table 1 represents the total values summed up according to the measurement method, i.e. number of annual unique instances, annual UI volume in pages, amount of copies, and annual total volume in pages for the 64 analyzed genres of the target unit.

Table 1. Organization's total Annual Volumes.

	Total Annual Volume (64 genres)
Annual Unique Instances	35 987
Annual UI Volume / Pages	84 711
Annual Copies	38 593
Annual Volume / Pages	143 909

The communication volumes per person per day are presented in Figure 1 for each group of employees (i.e. E1…E5 referring to Employee groups / roles).

Figure 2 represents the amount of communication for employee groups and the distribution of the communication in according to the communication media and formats used.

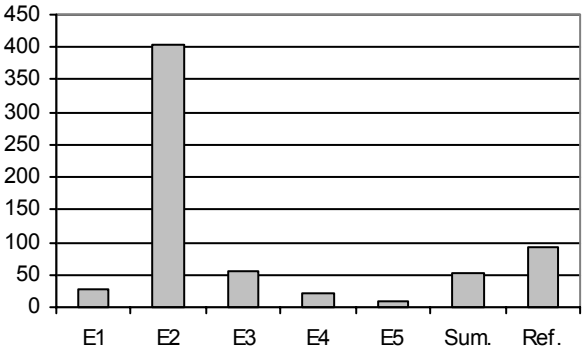

Figure 1. The amount of communication in pages per day per person (vertical axis) for the five employee roles (E1…E5), the average cross all the employees (Sum) and the reference value (Ref.) from (Tyrväinen 2003).

The employee groups on the x-axis are arranged in ascending order of total communication volume of the employees in the group. The Y-axis represents the total amount of communication of each employee in the group. Shares of distinct formats and media are represented by textures. Out of these, the three topmost categories (Encoded, Semi-structured, and Formal communication formats) are digital communication formats, Analogue represents mainly paper-based communication, Mediated refers to phone calls and similar technology enabled communication channels, and 1-to-1 represents verbal communication from one person to another. Figure 3 represents the same data with scaling to 100% for each group, i.e. distribution of the communication formats of each group.

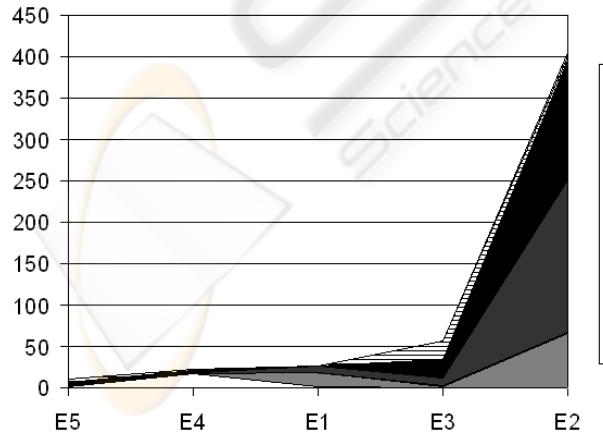

Figure 2. Categorized amounts of communication per employee for each group in ascending order of total communication volume per person day.

Role E2 on the right is the role with highest communication volume. According to Figures 2 and 3 this role communicates mainly using three media: digital documents, paper, and personal verbal communication. About 60 % of this communication takes place outside information systems. Furthermore from this 60 % the share of analogue (i.e. paper-based) communication is a little less than 70 %. The remaining 30 % of communication is face-to-face verbal communication. Digital communication accounting for about 40 % of all communication is almost completely in encoded form, i.e. in digital documents. The share of formal and semi-structured communication is altogether about 2 %.

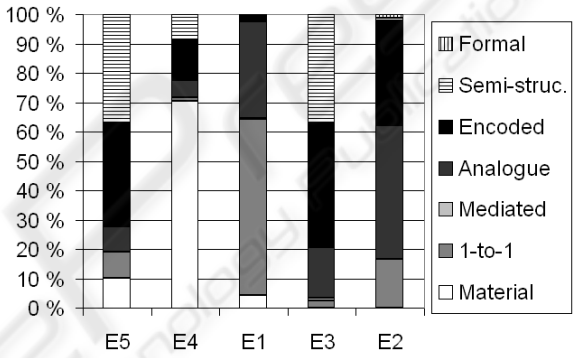

Figure 3. Shares of distinct formats by employee groups.

2.3 Result Analysis

Our hypothesis was that the degree of digital communication correlates with the total amount of communication. We can verify this hypothesis from multiple viewpoints: first, by comparing the degree of digitalization of the roles against the communication volumes of the roles, secondly, by comparing the total degree of digitalization with the reference values, and third, by analyzing in detail specific roles.

When comparing the individual roles, the hypothesis would lead to a situation, where the roles communicating more would have higher degrees of digitalization. Figure 4 represents the roles in a coordinate space with communication volumes on a logarithmic x-axis and the degree of digitalization in y-axis. No clear correlation is visible.

When comparing the sum values of the target unit (e.g. 40 % of digitalization and 52 pages of communication per person per day) with the two reference cases, the hypothesis seems to have some evidence. The three cases are presented in Figure 5 together with a linear approximation of their trend. Our case has the smallest average communication volume, while the degree of digitalization is in between the two reference values. The linear approximation of these three cases shows clear positive correlation in between total communication volume per person per day and the degree of digitalization. However, an approximation based on three data points does not yet provide very strong evidence for the hypothesis.

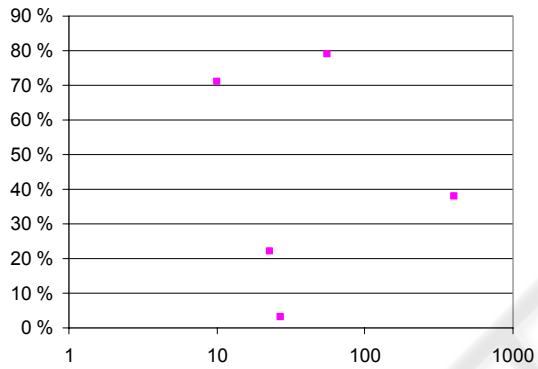

Figure 4. A comparison of three cases with the volume of communication per person on the x-axis and the degree of digitalization on the y-axis.

The third approach to analyze the results is to use qualitative analysis for individual roles. Role E2 with the highest communication volume (405 pages per person per day) seems to be a good candidate for this purpose.

According to the reference studies the volume of received communication is usually overwhelming the volume of produced unique communication due to the number of copies produced and distributed for each page of unique information created. If a person just reads a full day, s/he is able to read about 300 pages while writing 10 % of this is much more demanding. In the case of Tyrväinen et al. (2003) the lecturers produced 11 unique pages per day, but their total communication volume was reported to be about 370 pages per day due to the high number of copies, i.e. copies of lecture notes and high number of listeners of lectures. This number is close to the figure of 405 of role E2, which suggests that E2 is producing content distributed to a wider audience.

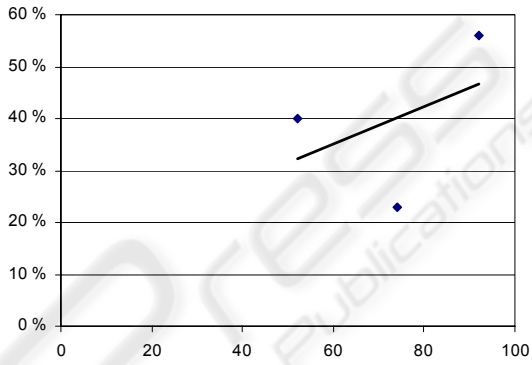

Figure 5. A comparison of three cases on a co-ordinate space of communication volume (x-axis) and degree of digitalization. The line represents linear approximation.

When analyzing the 21 communication genres of the role E2, the two genres with the highest volume accounted for 60% of the communication. Both of these genres can be characterized as reporting to external customers of the process (see table 2). The material used for this communication was produced using office tools, i.e., by personal computer. Half of this material was delivered as office documents and the other half was printed on paper. On average 10 copies of material were produced from each unique instance of these genres. Further, both of these genres were somewhat "hard", i.e. the structure and even contents of the instances repeated from one instance to another. This indicates that reuse of content takes place from one unique instance to another.

Table 2. Examples of genres collected in the case study.

Producer	**Genre**	**User**	**Categories of Communication Forms**	**Annual UI**	**Copies**	**Size in Pages**
E2	Report 1	Customer	Analogue, Encoded	Dozens	10	100
E2	Report 2	Customer	Analogue, Encoded	Dozens	10	40
Several employee groups	Results analysis	E2	Encoded, Semi-structured	Dozens	1	30
E2	Instructions	Several employee groups	1-to-1, Analogue	Thousands	1	6
E2	Objectives	Personnel	1-to-1, Analogue, Semi-structured	Thousands	1	6

The genres in fourth and fifth position accounted for 26% of the communication of the role E2. These included instructing other roles on tasks related to the operation of the main process of the unit. This communication took place either face-to-face or through paper, with distribution of 67 / 33 % and 50 / 50 % for verbal / paper communication for genres 3 and 4, respectively. As a rough generalization of the rest of the genres in the role 2, the trend of low volume with high number of unique instances was clearly seen in many communication situations, e.g. in verbal communication. Ad hoc -situations, e.g. unexpected problem situations where information must come across fast, act as a good example of this kind of genre.

3 DISCUSSION

It is easy to observe, that one of the roles (E2) is under a very high communication pressure as it is receiving and sending a total of over 400 pages of communication per an average day. However, the communication load of the other roles was very modest, none of which exceeded the reference values of 92 or 74 pages per day. This leads to a situation, where the average communication volume in the target organization is 52 pages per person day, i.e. below the reference values, but still a single role is under extremely high communication pressure. Even if the copies are excluded, the number of 137 unique pages of communication produced each day is enormous compared to any other role.

However, one should notice that the major portion of this high volume of communication is produced with the aid of digital computers although a major part of that communication is printed on paper for delivery. If the percentages of Figures 2 and 3 represented the medium or format used for creation of the data communicated instead of the delivery media of the messages, the percentage of digital communication of role E2 would be much higher. Changing just the data of the two topmost genres would raise the digitalization degree of E2 to close to 70 %. Figure 6 represents this situation by replacing the digitalization value of E2 (the right-most data point) to the value of 68 %. This change does not have an impact on the other data points as these two genres are communicated from E2 to external roles. On this basis also the correlation of communication volume and digitalization seems to be positive.

In general, the results of our research seem to be reasonable from the viewpoint that work in traditional industry is not as information intensive. The work consists of physical tasks such as reading the meters, collecting samples etc. 52 pages communicated per person in a day in our case is equivalent to about one sixth of his work time, which means 1 hour and 20 minutes spent on communicating each day. If the results would have been closer to the reference value, we would have questioned whether the people have time to do the actual tasks.

A worker in role 2 controls almost everything concerning the process. Without him, or rather without his knowledge, the process will not work in a way it should work. That is why his tacit knowledge should be expressed explicitly to become part of organizational information resource, to provide a contingency plan for problem situations like illnesses. Transformation of implicit knowledge to an explicit form increases also the transparency of the whole process to all interest groups. Common uncertainty avoidance is a thing which is possible to reject in this way. This is emphasized especially in alteration situations when the needed information does not always reach the one concerned.

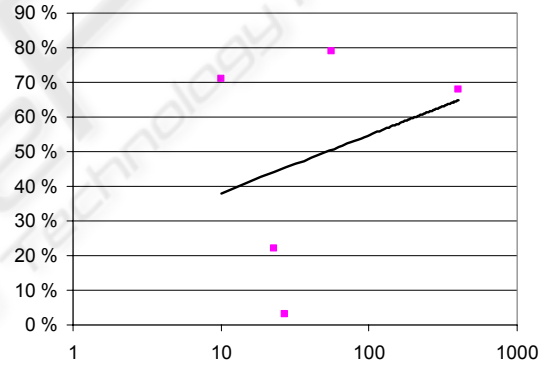

Figure 6. Volumes and digitalization degrees of the roles in the target unit when the degree of digitalization of role E2 is based on the medium of content production rather than content delivery.

The results mean in this case that the amount of E2's communication is so high that it is entitled to consider whether he is able to handle all the information he is working with. If his work cannot be reorganized, supportive applications for his work should be developed. The key point here is to transform the share of analogue communication to digital formats. In other words this requires reorganization of the communication channels used. Half of this can be achieved simply by not printing out the documents and presentation materials produced with office tools, as described above. Better exploitation of retrieval system used in the organization would provide a solution to a major part of the rest of the problem. An alternative way of

looking this problem is to develop an information system where all the information concerning the main process would be managed. This system should be accessible for all parties concerned to add, remove and update information depending on the rights tied to roles. From a broader perspective the use of this kind of information system would enable the utilization of data mining techniques and different kind of artificial intelligence and expert systems as well as computational methods. As a result massive amounts of information could be refined to better meet users' needs enabling effective exploitation of information.

In the beginning the two factors supporting our hypothesis were the assumed productivity improvement achieved by digital communication and the ease of communication with digital media. The qualitative analysis of role E2 supported both of these. Without this analysis the comparison of roles would not have supported the hypothesis, while now both the role comparison and the comparison of our case with the two other cases do support the hypothesis to a reasonable extent.

The qualitative analysis disclosed that it would be worthwhile to study the media and formats used for creation of new communication instead (or in addition to) the media used for delivering the message. For verbal communication there is no difference, but with digital documents delivered on paper this has major impact, as presented in this case.

In addition, the qualitative analysis disclosed a third potential factor increasing the correlation of digitalization and communication volume. In cases, such as role E2, where the throughput of a process depends on the performance of a single critical role, there exist good reasons to support that role with information systems. This means, that also high volumes of communication tend to increase the degree of digitalization, in order to improve organizational performance.

4 SUMMARY AND FURTHER RESEARCH

This paper discussed the degree of digital communication with relation to the amount of total communication. In the case study we used the genre-based measurement method to gather data on communication flows in an industrial target organization. The results show, that the amount of communication per person per day is less than the communication in reference cases and is highly independent from information systems, although they are used extensively in process control. The qualitative analysis of a key role disclosed that there seems to be some degree of correlation between the degree of digital communication and the amount of total communication, when comparing communication of the five roles in the target organization. Also when comparing these results with reference studies, there seems to be some evidence of positive correlation. Future research is needed to have statistically stronger evidence about this correlation.

REFERENCES

Allen, L., 1991. Using Organizational Communication to Manage Technical Change. In *IPPC '91. Proceedings. 'The Engineered Communication'., International, IEEE*, 1&2, 30 Oct.-1 Nov. 1991, 351-355 vol. 2.

Aris, J., 2000. Inventing Systems Engineering [LEO]. *Annals of the History of Computing, IEEE*, 22, 3, 4-15.

Blair, D.C., 2002. The Challenge of Commercial Document Retrieval, Part I: Manor Issues and a Framework Based on Search Exhaustivity, Determinacy of Representation and Document Collection Size. *Information Processing & Management*, 38 (3), 273-291.

Cortada, J.W., 1996. Commercial Applications of the Digital Computer in American Corporations, 1945-1995. *Annals of the History of Computing, IEEE*, 22, 3, 4-15.

Cortada, J.W., 1997. Economic Preconditions That Made Possible Application of Commercial Computing in the United States. *Annals of the History of Computing, IEEE*, 19, 3, 27-40.

Davenport, T.H., 1993. Process Innovation: Reengineering Work Through Information Technology. Ernst & Young, Center for Information Technology and Strategy, Harvard Business Press, Boston, Massachusetts.

Karjalainen, A., Päivärinta, T., Tyrväinen, P., and Rajala, J., 2000. Genre-Based Metadata for Enterprise Document Management, in *33rd Annual Hawaii International Conference on System Sciences*: Los Alamitos CA, IEEE Computer Society.

Land, F.F., 2000. The First Business Computer: A Case Study in User-Driven Innovation. *Annals of the History of Computing, IEEE*, 22, 3, 16-26.

Moore, G. E., 1975. Progress in Digital Integrated Electronics. IEDM.

Morgan, G., 1988. Riding the Waves of Change: Developing Managerial Competencies for a Turbulent World. Jossey-Bass Publishers, San Francisco CA.

Nunberg, G., 1997. Genres in Digital Documents: Introduction. In Proc. of *The 30th Annual Hawaii International Conference on System Sciences: Digital*

Documents, Vol VI, Los Alamitos CA: IEEE Computer Society Press, pp. 2.

Orlikowski, W.J. and Yates, J. 1994. Genre repertoire: The structuring of Communicative Practices in Organizations. *Administrative Science Quarterly*, 39, 4, 541-574.

Päivärinta, T., Halttunen, V., and Tyrväinen, P., 2001. A Genre-Based Method for Information Systems Planning, in *Information Modelling in the New Millennium*: Hershey PA, Idea Group, pp. 70-93.

Saaren-Seppälä, K., 1997. Seinätekniikka prosessien kehittämisessä. (Using the wall-chart technique in process development, in Finnish), Kari Saaren-Seppälä Ltd, Finland.

Schaller, R.R, 1997. Moore's Law: Past, Present and Future, *Spectrum, IEEE*, 34, 6, 52-59.

Scott, K. 2002. The Unified Process Explained. Pearson Education, Inc.

Scott Morton, M.S., 1991. (ed.) The Corporation of the 1990s: Information Technology and Organizational Transformation. Oxford University Press, New York, New York.

Sosa, M.E, Eppinger, S.D., Pich, M., McKendrick, D.G., Stout, S.K., 2002. Factors that Influence Technical Communication in Distributed Product Development: An Empirical Study in the Telecommunications Industry. *Engineering Management, IEEE Transactions*, 49, 1, 45-58.

Sproull, L., and Kiesler, S., 1991. Connections: New Ways of Working in the Networked Organization. Cambridge, MA: MIT Press.

Stalk, G., Evans, P., and Shulman, L.E., 1992. Competing on Capabilities: The New Rules of Corporate Strategy. Harvard Business Review, 57-69.

Tyrväinen, P., 2003, Estimating Applicability of New Mobile Content Formats to Organizational Use, In *HICSS-36, 36rd Annual Hawaii International Conference on System Sciences*, January 6-9 2003, IEEE Computer Society.

Tyrväinen, P., Järvenpää, M., and Sievänen, A., 2003. On Estimating the Amount of Learning Materials – A Case Study. *Proceedings of ICEIS 2003 – The Fifth Conference on Enterprise Information Systems*, April 22-26 2003, Angers, France, Vol. 4. pp. 127-135. ICEIS Press 2003.

Tyrväinen, P., and Päivärinta, T., 2003. How Digital is Communication in Your Organization? A Metrics and an Analysis Method. In Camp, 0., Filipe, J., Hammoudi, S., and Piattini, M. (eds.) Enterprise Information Systems V, Kluwer Academic Publishers, The Netherlands, November 2003.

van Grembergen, W., 2002. Information Systems Evaluation Management, IRM Press, 2002. 322 p.

Yates, J., 1989. Control Through Communication: The Rise of System in American Management. Baltimore, Johns Hopkins University Press.

Yates, J. and Orlikowski, W.J., 1992, Genres of Organizational Communication: A Structurational Approach to Studying Communication and Media, *Academy of Management Review*, 17, 2, 299-326.

Yoshioka, T., Herman, G., Yates, J., and Orlikowski, W., 2001. Genre Taxonomy: A Knowledge Repository of Communicative Actions, *ACM Transactions on Information Systems*, 19, 4, 431-456.