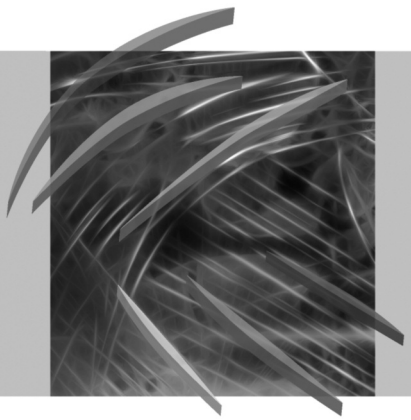
PRACE NAUKOWE
Uniwersytetu Ekonomicznego we Wrocławiu
RESEARCH PAPERS
of Wrocław University of Economics

205

Advanced Information Technologies for Management – AITM 2011

Information Systems in Business



edited by
Jerzy Korczak, Helena Dudycz, Mirosław Dyczkowski

Publishing House of Wrocław University of Economics
Wrocław 2011

Reviewers: Frederic Andres, Witold Chmielarz, Jacek Cypryjański, Beata Czarnacka-Chrobot, Bernard F. Kubiak, Wojciech Olejniczak, Celina M. Olszak, Marcin Sikorski, Ewa Ziemba

Copy-editing: Agnieszka Flasińska

Layout: Barbara Łopusiewicz

Proof-reading: Marcin Orszulak

Typesetting: Adam Dębski

Cover design: Beata Dębska

This publication is available at www.ibuk.pl

Abstracts of published papers are available in the international database The Central European Journal of Social Sciences and Humanities http://cejsh.icm.edu.pl
and in The Central and Eastern European Online Library www.ceeol.com

Information on submitting and reviewing papers is available on the Publishing House's website www.wydawnictwo.ue.wroc.pl

All rights reserved. No part of this book may be reproduced in any form or in any means without the prior written permission of the Publisher

© Copyright Wrocław University of Economics
Wrocław 2011

ISSN 1899-3192
ISBN 978-83-7695-178-2

The original version: printed

Printing: Printing House TOTEM

Contents

Streszczenia

PRACE NAUKOWE UNIWERSYTETU EKONOMICZNEGO WE WROCŁAWIU nr 205
RESEARCH PAPERS OF WROCŁAW UNIVERSITY OF ECONOMICS

Advanced Information Technologies for Management – AITM 2011 ISSN 1899-3192
Information Systems in Business

Radosław Wójtowicz
Wrocław University of Economics, Wrocław, Poland
e-mail: radoslaw.wojtowicz@ue.wroc.pl

THE CHOSEN ASPECTS OF REAL-TIME COLLABORATIVE EDITING OF ELECTRONIC DOCUMENTS

Abstract: The article presents a few problems of electronic documents collaborative editing. The main purpose of the article is to gather and explain the most significant notions in the described area, and to identify the major problems in the realm of creation and use of the software supporting collaborative creating and editing documents, as well as to present a synthetic overview of the modern market of such systems.

Keywords: document management, electronic document, electronic document management system, collaborative editing, real-time collaborative editing.

1. Introduction

Over the past few years, information technology systems whose task is to support common creation of various documents have become very popular, especially among individual users and small businesses. These systems are currently being intensively developed as the answer to the fast growing demand for software which will transfer the so far used groupware technologies to a cloud computing model. In our country the question concerning this kind of systems is still practically non-existent in business informatics literature, therefore spreading knowledge about that subject seems to be necessary.

Section 2 of the study concentrates on the key terminology. Section 3 describes chosen research and practical problems which are currently connected to the described area. Section 4 presents synthetic characteristics of the most popular software.

2. Key terminology

This part of the article presents some key terms in the documents management area.

Document is a recorded information or an object which can be treated as a unit. A document may be on paper, microform, magnetic or any other electronic medium. It may include any combination of text, data, graphics, sound, moving pictures or any other forms of information. A single document may consist of one or several components [http://www.moreq2.eu]. Documents represent basic media of open knowledge in organisations. For the most part (ca. 80%) this knowledge is represented by electronic documents which contain unstructured and semi-structured data [Gołuchowski 2005].

Electronic document is a collection of data, as a separate significative entity, ordered in a defined internal structure and recorded on a data carrier [Act 2005].

Electronic document management is a general term for paper documents, involves imaging, indexing/coding, and archiving of scanned documents/images, and thereafter electronically managing them during all life cycle phases. Electronic documents are likewise electronically managed from creation to archiving and all stages in between [http://www.amdoc.com/interior.php?sec=18]. Document management is related to the records management, but there is the subtle difference between these notions. With **records management** the main emphasis is on the organization and safeguarding of recorded information throughout its life cycle [Wiggins 2000, p. 66].

Electronic document management system (EDMS) is a computer system used to track and store electronic documents and/or images of paper documents. It is usually also capable of keeping track of different versions created by different users (history tracking) [http://en.wikipedia.org/wiki/Document_management_system]. These systems support work with documents by acquiring electronic form of documents, for example scanning and processing with the use of suitable software, managing the documents and work on them with keeping confidentiality policy [Misiak 2010, p. 435].

Collaborative editing is the practice of groups producing works together through individual contributions. Effective choices in group awareness, participation, and coordination are critical to successful collaborative writing outcomes. Most usually it is applied to textual documents or a programmatic source code. Generally, managing such work requires software – the most common tools for editing documents are wikis, and those for programming, version control systems. Most word processors are also capable of recording changes; this allows editors to work on the same document while automatically clearly labelling who contributed what changes [http://en.wikipedia.org/wiki/Collaborative_editing].

Collaborative editor is a form of collaborative software application that allows several people to edit a computer file using different computers. There are two types of collaborative editing: real-time and non-real-time.

Real-time collaborative editing (RTCE) is synchronous (simultaneous), meaning that users can edit the same file at the same time.

Non-real-time collaborative editing is asynchronous, which means that editors do not edit the same file at the same time (similar to revision control systems). Collaborative real-time editors generally permit both synchronous and asynchronous editing in any given instance [http://en.wikipedia.org/wiki/Collaborative_editor].

Real-time collaborative editors are usually one of the basic elements of online office suits, and in the next parts of the article we concentrate on online problems and solutions.

Online office suite is a type of office suite offered by websites in the form of software as a service. They can be accessed online from any Internet-enabled device running any operating system. This allows people to work together worldwide and at any time, thereby leading to international web-based collaboration and virtual teamwork. Usually, the basic versions are offered for free and for more advanced versions one is required to pay a nominal subscription fee [http://en.wikipedia.org/wiki/Online_office_suite]. Online office suits are dedicated for individual users and small and medium sized companies. They represent the modern trends on IT market: cloud computing and SaaS model, which seem to be an attractive alternative to the "traditional" IT-solution [Maciejewski 2010].

3. The most important problems in the real-time collaborative editing area

Generally, the problems that affect collaborative editors can be divided into research and utilitarian. Research problems relate mainly to system developers, while the utilitarian problems usually relate to the users.

One of the major research problems which still face the creator of real-time collaborative editors is to ensure proper operation of the software, despite the presence of communication lags in the network. In theory, if communication were instantaneous, then creating a real-time collaborative editor would be no more difficult than creating a single-user editor. Figure 1 illustrates this problem.

We have a text document with a string "abc" replicated at two collaborating sites and two concurrent operations: O1 = Insert[0, "x"] (to insert character "x" at position "0") and O2 = Delete[2, "c"] (to delete the character "c" at position "2") generated by two users at collaborating sites 1 and 2. Suppose the two operations are executed in the order of O1 and O2 (at site 1). After executing O1, the document becomes "xabc". To execute O2 after O1, O2 must be transformed against O1 to become: O2' = Delete[3, "c"], whose positional parameter is incremented by one due to the insertion of one character "x" by O1. Executing O2' on "xabc" shall delete the correct character "c" and the document becomes "xab". However, if O2 is executed without transformation, then it shall incorrectly delete character "b" rather than "c".

This kind of transformation is called **operational transformation (OT)**. The basic idea of OT for consistency maintenance is to transform an editing operation

into a new form according to the effects of previously executed concurrent operations so that the transformed operation can achieve the correct effect and ensure that replicated documents are identical.

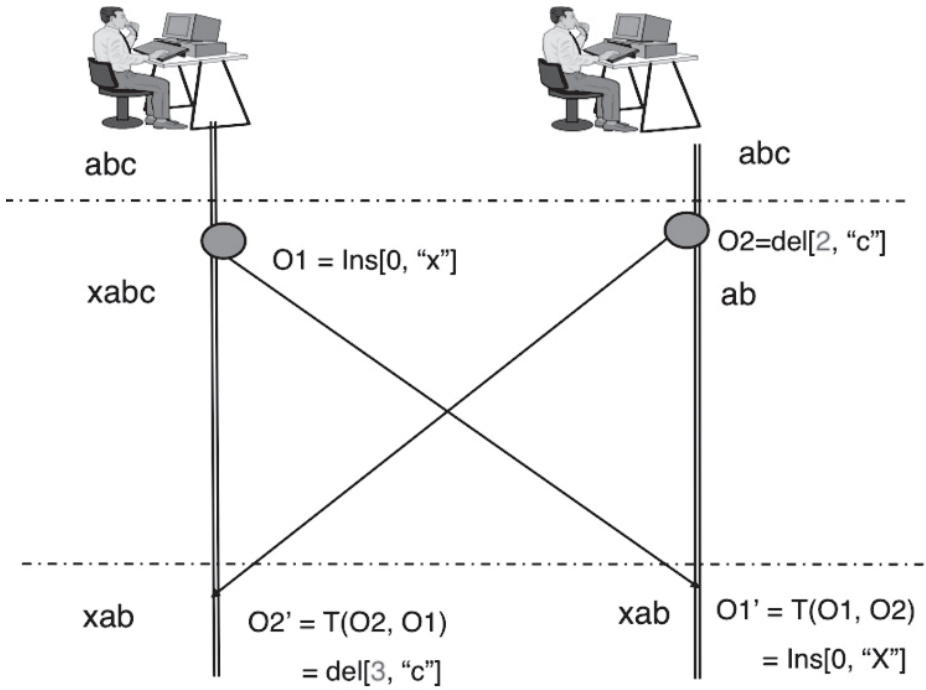


Figure 1. An example of the real-time collaborative editing

Source: [http://en.wikipedia.org/wiki/File:Basicot.png].

It is a technology for supporting a range of collaboration functionalities in advanced groupware systems. OT was originally invented for consistency maintenance and concurrency control in collaborative editing of plain text documents. For the first time the operational transformation model was proposed in 1989 by C.A. Ellis and S.J. Gibbs and used in the GROVE system (GRoup Outline Viewing Edit) [Ellis, Gibbs 1989, pp. 399–407].

Over two decades of research has extended its capabilities and expanded its applications to include group undo, locking, conflict resolution, operation notification and compression, group-awareness, HTML/XML and tree-structured document editing, collaborative office productivity tools, application-sharing, and collaborative computer-aided media design tools. Recently, OT has been adopted as a core technique behind the collaboration features in Google Wave and Google Docs, which are taking OT to a new range of web-based applications [http://en.wikipedia.org/wiki/Operational_transformation].

One established strategy of structuring an OT system is to separate the high-level transformation control (or integration) algorithms from the low-level transformation functions and specify the relationships (responsibilities and constraints) between these two layers as transformation properties and conditions in the middle of the OT system, as shown in Figure 2.

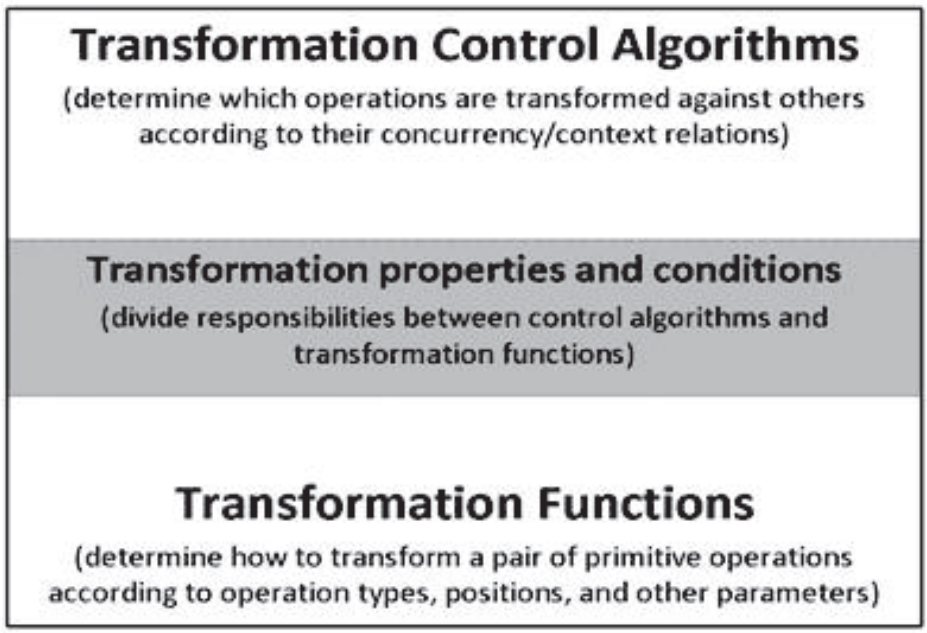


Figure 2. A layered structure of an OT system

Source: http://cooffice.ntu.edu.sg/otfaq/.

It should be emphasized that the use of the operational transformation solves the problem of the syntactic (formal) consistency of the document, but does not solve the difficulties concerning the so-called semantic coherence of the document, which means preserving the proper and correct meaning of the text changed by many users. In this case, it is necessary to connect the operational transformation with the respective mechanisms to lock the documents.

The issue of operational transformation is still being developed, both theoretically and practically. In recent years several extended models of OT have taken into account all the possible operations performed on documents. Researchers also created a lot improved and more efficient algorithms for operational transformation, but their detailed description is beyond the scope of this paper [Li, Li 2010, pp. 1–43].

The most important utilitarian problems in the real-time collaborative editing area are:

– Access to the documents requires connectivity. If the remote server or network is unavailable, the content will also be unavailable.
– There are speed and accessibility issues. Most of the available online office suites require a high speed Internet connection.
– Online office suites lack the more advanced features available on their offline counterparts.
– There may be a subscription charge to use the service. In that case, in the long run, the ongoing subscription cost may be more expensive than purchasing offline software.
– Potential low security level of the data and documents which have to be stored on the remote central servers.
– The lack of trust of small entrepreneurs in big companies. This is the common problem that affects the whole IT market in the cloud computing era. There are also questions as to how businesses will be affected by storing all of their documents in online environments [http://en.wikipedia.org/wiki/Online_office_suite].

The subject matter of the collaborative editing is a very wide area and the author of the article will continue his research on it.

4. The chosen software supporting real-time collaborative editing

In this part of the article we propose a generic classification and synthetic characteristic of the most popular software which support real-time collaborative editing. We concentrate on these online office suits that have been available on the market for longer than one year and are intensively developed by the authors.

We make the classification of the office software using two essential criteria: the cost of the basic version and the kind of the access to the functionality of the software. Due to the costs of using of the basic version, we can divide the described software into payable and free. Due to the kind of the access to the functionality we can divide the software into locally installed applications (offline) and applications based on the Internet browser (online, web-based). Table 1 shows the classification of the selected software that incorporates the described criteria.

Table 1. The generic classification of the chosen office software

	Payable	Free
Offline	• MS Office (with SharePoint Workspace) • ThinkFree Office • iWork (Apple)	• OpenOffice • LibreOffice • IBM Lotus Symphony • Calligra Suite
Fully online	• ShareOffice • ZCubes • Feng Sky	• Google Docs • MS Office Web Apps • Zoho Office Suite

Source: own study based on [http://en.wikipedia.org/wiki/Office_suite].

Currently, more and more popular are free online office suits, among which the most known is Google Docs. In view of the fact that its functionality, can be considered as representative for other software, now we present a brief description of this product.

Google Docs is recommended for personal files, work in progress, and generally temporary usage. Users can create documents, spreadsheets, presentations, forms and drawings online. It is also possible to upload existing files because Google Docs accepts most popular file formats, including DOC, XLS, ODT, ODS, RTF, CSV, PPT, etc. The main page of the Google Docs and types of the new files are presented in Figure 3.

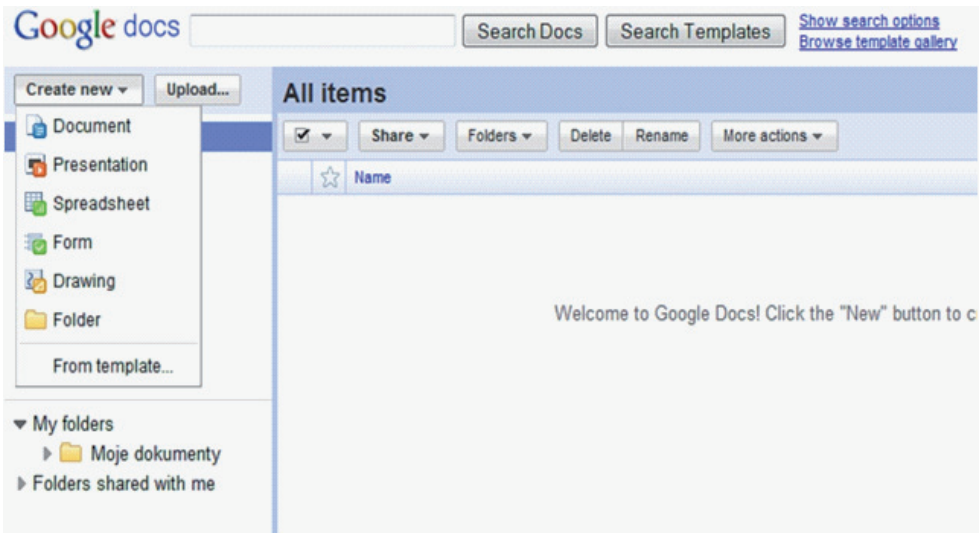


Figure 3. Types of the new files in Google Docs

Source: own study based on [http://docs.google.com].

One of the most important advantages of all online office suits is very easy documents sharing. If the persons with whom we want to share a given document have a Google Account, we just enter the e-mail addresses and send them an invitation. Anyone we have invited to either edit or view our document, spreadsheet or presentation, can access it as soon as they sign in. Multiple people can view and make changes at the same time. There is an on-screen chat window for spreadsheets, and document revisions showing exactly who changed what, and when. The window for sharing document is presented in Figure 4.

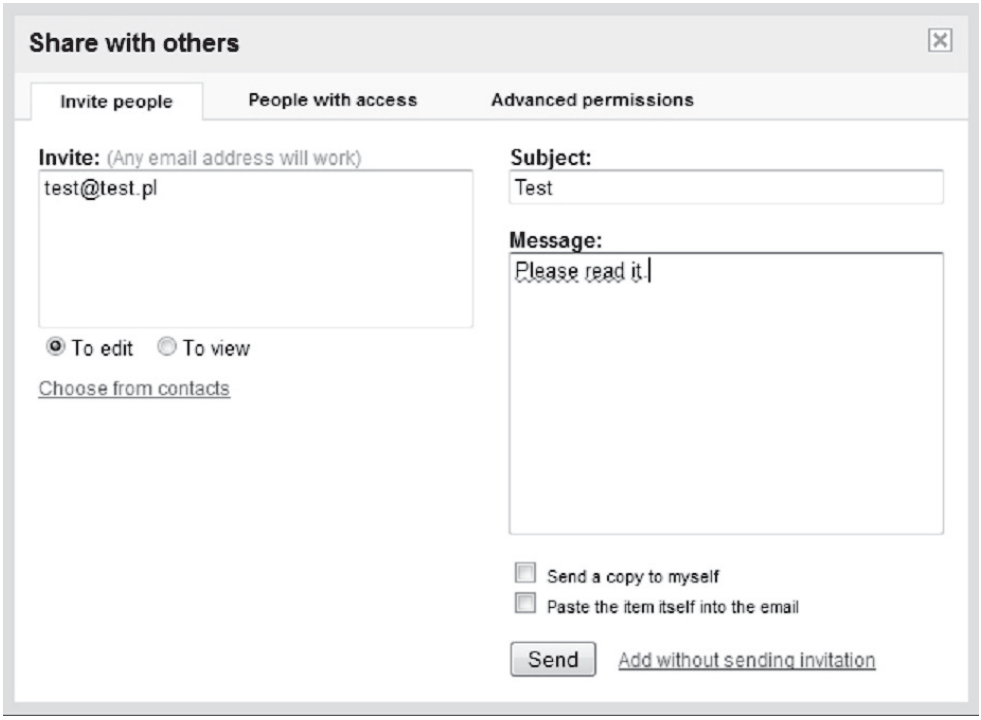


Figure 4. Window "Share with others" in Google Docs

Source: own study based on [http://docs.google.com].

There are also some disadvantages of Google Docs. The most important is that Google Docs does not permit having a company folder structure, thus it is not possible for all users to see documents in a specific folder structure. Another problem is sharing defined on a per user basis, thus will be cumbersome when dealing with changing team members and distributed teams.

It is worth mentioning another innovative project of the company from Mountain View, namely Apache Wave, originally developed by Google as Google Wave. **Apache Wave** is a web-based computing platform and communications protocol, designed to merge key features of media like e-mail, instant messaging, wikis, and social networking. Google Wave works like previous messaging systems such as e-mail and Usenet, but instead of sending a message along with its entire thread of previous messages, or requiring all responses to be stored in each user's inbox for context, message documents (referred to as *waves*) that contain complete threads of multimedia messages (blips) are perpetually stored on a central server. Waves are shared with collaborators who can be added or removed from the wave at any point during a wave's existence [http://en.wikipedia.org/wiki/Apache_Wave].

Only in 2010 Microsoft offered for all users a new online office suite called **Office Web Apps.** It includes the web-based versions of Microsoft Word, Microsoft Excel, Microsoft PowerPoint, and Microsoft OneNote. The web apps allow users to access their documents directly from anywhere within a web browser as well as share files and collaborate with other users online. Supported web browsers include Internet Explorer 7 or, later, Mozilla Firefox 3.5 or, later, Google Chrome and Safari 4. The main ribbon of the Microsoft Word Web App is presented in Figure 5.

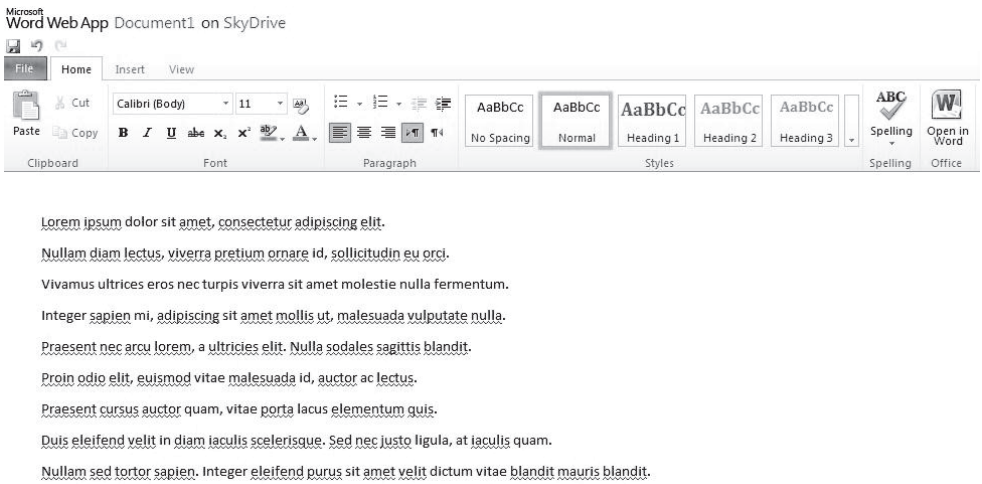


Figure 5. A screenshot of Microsoft Word Web App

Source: own study based on [http://office.microsoft.com/en-us/web-apps].

An interesting alternative to Google and Microsoft solutions for collaborative editing seems to be a software called CodoxWord (formerly CoWord). **CodoxWord** is a free software add-on to Microsoft Word to enable multiple users to edit the same document over the Internet with MS Word. The concurrency control technology used by CoWord is, described earlier, operational transformation.

ZOHO is a suite of online applications (services) that users can access from a website. The office applications are free for individuals and some have a subscription fee for organizations. ZOHO Corporation's vision is to provide customers (individuals, students, educators, non-profits, small and medium sized businesses) with the most comprehensive set of applications available anywhere.

ZOHO applications are divided into two sets: Productivity & Collaboration Apps and Business Apps. The most important office applications are Zoho Writer, Zoho Sheet, Zoho Show and Zoho Docs. Like in Google Docs users can create documents, spreadsheets, presentations and upload own files. The main page of the ZOHO suite is presented in Figure 6.

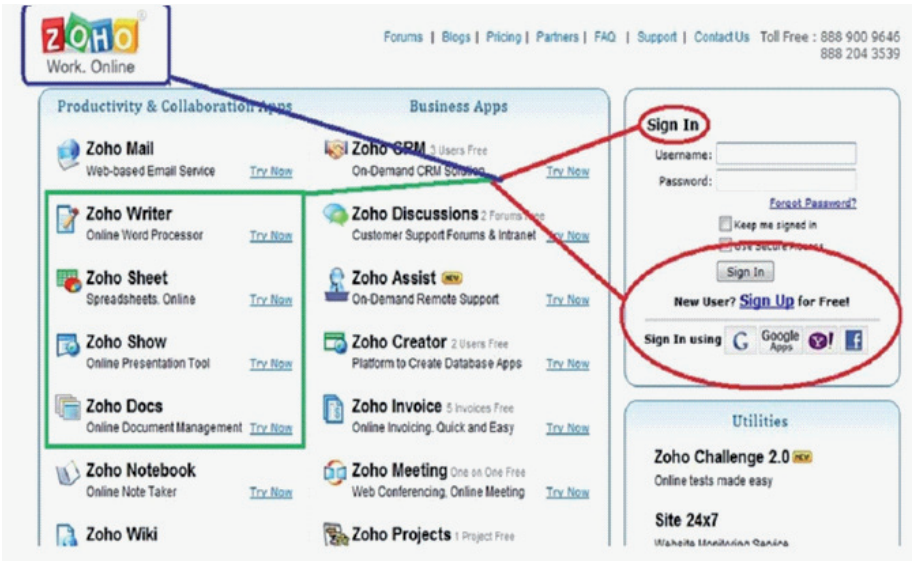


Figure 6. The main page of the ZOHO suite

Source: own study based on [http://www.zoho.com].

ZOHO suit is bringing together a wide range of online applications. Extra applications, which are not available in Google Docs, are for example project management, database or CRM. However, the interface and usability of this suite may be complicated for inexperienced users.

5. Conclusions

Modern software used for real-time collaborative editing of electronic documents is a practical implementation of recently popular idea of cloud computing which is becoming more accessible for individual users and small and medium businesses. It seems that one of the positive results of launching this type of solutions on the market by many IT potentates will be a decrease in "traditional" office software prices.

Another visible trend on the office software market is putting a greater emphasis on the question of ensuring a high level of security for users' documents by the systems' producers. This seems to be one of the key questions which may decide if a cloud computing model will achieve a real success on the market.

References

Act (2005), Ustawa z dnia 17 lutego 2005 r. o informatyzacji działalności podmiotów realizujących zadania publiczne, *Dziennik Ustaw* [Journal of Laws] Nr 64, poz. 565.
Ellis C.A., Gibbs S.J. (1989), Concurrency control in groupware systems, *ACM SIGMOD Record*, Vol. 18, No. 2.
Gołuchowski J. (2005), *Technologie informatyczne w zarządzaniu wiedzą*, Prace Naukowe Akademii Ekonomicznej w Katowicach, Wydawnictwo Akademii Ekonomicznej, Katowice.
Li D., Li R.(2010), An admissibility-based operational transformation framework for collaborative editing systems, *Computer Supported Cooperative Work*, Vol. 19, No. 1.
Maciejewski A. (2010), Rewolucja przyjdzie z chmury, *Computerworld*, nr 34, Wydawnictwo IDG, Warszawa.
Misiak Z. (2010), *Systemy elektronicznego obiegu dokumentów WFM*, [in:] S. Wrycza (Ed.), *Informatyka ekonomiczna*, PWE, Warszawa.
Wiggins B. (2000), *Effective Document Management*, Gower Publishing Limited, Hampshire.

Websites

http://cooffice.ntu.edu.sg/otfaq/ (accessed: 1.07.2011).
http://docs.google.com (accessed: 30.06.2011).
http://en.wikipedia.org/wiki/Apache_Wave (access: 1.07.2011).
http://en.wikipedia.org/wiki/Collaborative_editing (accessed: 1.07.2011).
http://en.wikipedia.org/wiki/Collaborative_editor (accessed: 1.07.2011).
http://en.wikipedia.org/wiki/Document_management_system (accessed: 1.07.2011).
http://en.wikipedia.org/wiki/File:Basicot.png (accessed: 1.07.2011)
http://en.wikipedia.org/wiki/Office_suite (accessed: 1.07.2011).
http://en.wikipedia.org/wiki/Online_office_suite (accessed: 1.07.2011).
http://en.wikipedia.org/wiki/Operational_transformation (access: 1.07.2011).
http://office.microsoft.com/en-us/web-apps (accessed: 28.06.2011).
http://www.amdoc.com/interior.php?sec=18 (accessed: 29.06.2011).
http://www.moreq2.eu (accessed: 30.06.2011).
http://www.zoho.com (accessed: 30.06.2011).

WYBRANE ASPEKTY GRUPOWEGO REDAGOWANIA DOKUMENTÓW ELEKTRONICZNYCH W CZASIE RZECZYWISTYM

Streszczenie: Artykuł prezentuje wybrane aspekty zastosowania systemów do grupowej edycji dokumentów elektronicznych. Głównym celem artykułu jest zebranie i wyjaśnienie najważniejszych pojęć dotyczących opisywanego obszaru, identyfikacja podstawowych problemów w dziedzinie tworzenia i użytkowania systemów wspomagających grupowe edytowanie dokumentów, a także dokonanie syntetycznego przeglądu współczesnego rynku tego rodzaju systemów.