Internet-Based Corporate Reporting  
- Filling the Standardization Gap

Falk v. Westarp*, Dieter Ordelheide**, Michael Stubenrath**, Peter Buxmann* and Wolfgang König*

*Institut für Wirtschaftsinformatik  
Johann Wolfgang Goethe Universität  
Frankfurt/Main, Germany  
westarp@wiwi.uni-frankfurt.de

**Professur für Internat. Rechnungswesen  
Johann Wolfgang Goethe Universität  
Frankfurt/Main, Germany  
stubenrath@wiwi.uni-frankfurt.de

Abstract
In addition to printed annual reports, a growing number of corporations use the Internet for corporate reporting. Current practice shows that neither the amount of data supplied nor the presentation modes of digital annual reports are standardized. Authenticity and integrity problems are obvious. In order to close this standardization gap we discuss standards concerning content, presentation, authenticity and integrity of data supplied. Furthermore, we evaluate the potential application of the new emerging data format XML which could add significant benefits to digital reports.

1. Introduction
The annual report is one of the most important sources of information for capital market participants. It provides detailed financial information about a corporation in a more or less standardized way. In addition to printed annual reports, a growing number of corporations use the Internet for corporate reporting. A recent empirical study [6] shows that large US, UK and German corporations already use the Internet to distribute annual reports. Nevertheless the acceptance of these digital reports and the professional use of the data provided seems to be limited due to the following problems. First, many of the analyzed digital annual reports on homepages provide less information than the printed reports. Second, the user is confronted with various presentation formats and not all of them enable further automated processing of the data. Finally, the authenticity and integrity of the data is not secured, which is a serious problem of electronic versions of annual reports [18]. Taking all this into consideration, a standardization gap within Internet-based corporate reporting becomes obvious.

In this article we address current deficiencies concerning the scope of financial accounting standards with respect to different media. Furthermore, we explicitly outline the authenticity and integrity problem of annual reports on corporations’ homepages. We will give concrete recommendations as to how the problem can be solved by the application of digital signatures. In addition, we analyze different presentation standards currently in use and compare them with the new XML standard, which might reduce information costs, e.g. the costs of data processing, significantly.

The paper is structured as follows: In section 2 the communication network of capital markets is described and the significance of the annual report within the network is discussed. Based on a recent empirical study, section 3 analyzes the status quo of current Internet-based corporate reporting. In section 4 we discuss the deficiencies of standards used for today’s Internet-based corporate reporting. In this context we examine standards concerning the content of digital reports, presentation formats and data security. Furthermore, we describe the newly emerging Internet standard XML and analyze to what extent the use of this standard for Internet-based corporate reporting might overcome the current problems. The final section summarizes the results.

2. Communication Networks and the Annual Report

Coordination on capital markets is based on complex communication networks between companies and investors. Information intermediaries serve as middlemen between these key actors. One example of information intermediaries on capital markets are financial analysts, who act as agents of investors by analyzing potential investments. Investors make their capital allocation deci-
sions based on a specific set of information. This set consists of a variety of information from different sources, including advice of financial analysts.

Annual reports (as well as interim reports) are part of this complex communication network on capital markets, especially with regard to the communication between companies and other capital market participants like investors and financial analysts. The communication strategy of companies regarding potential and current investors, including intermediaries like analysts and the financial press, is called investor relations [8]. The typical instruments of investor relations are annual reports, interim reports, annual general meetings, analyst conferences, roundtable and one-to-one discussions as well as conference calls with company representatives [9: 404-410]. Of these instruments the annual report is one of the most important sources of information in practice [1, 2, 19].

Financial reporting communicates financial information to investors, analysts and other interested groups. Financial information can thereby be described as information about business transactions and status in terms of money. To be decision-useful, financial information is classified and summarized. One kind of a classified summary is the annual report, which informs in words and numbers about the financial position of an enterprise [10: 2-4]. The main financial statements within the annual report are the balance sheet, the profit and loss account and the cash flow statement. Figure 1 illustrates the characteristic structure of annual reports by showing an example for the consolidated statements of operations (equivalent to the profit and loss account).

### CONSOLIDATED STATEMENTS OF OPERATIONS

Northern Telecom Ltd., year ended December 31 (millions of U.S. dollars except per share figures)

<table>
<thead>
<tr>
<th></th>
<th>1997</th>
<th>1996</th>
<th>1995</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Revenues</strong></td>
<td>15,442</td>
<td>12,847</td>
<td>10,672</td>
</tr>
<tr>
<td><strong>Cost of revenues</strong></td>
<td>9,111</td>
<td>7,714</td>
<td>6,379</td>
</tr>
<tr>
<td><strong>Gross profit</strong></td>
<td>6,331</td>
<td>5,133</td>
<td>4,303</td>
</tr>
<tr>
<td>Selling, general and administrative expense</td>
<td>2,714</td>
<td>2,193</td>
<td>1,923</td>
</tr>
<tr>
<td>Research and development expense (note 2)</td>
<td>2,147</td>
<td>1,813</td>
<td>1,579</td>
</tr>
<tr>
<td>Goodwill amortization (note 3)</td>
<td>45</td>
<td>45</td>
<td>90</td>
</tr>
<tr>
<td>Special charges (note 5)</td>
<td>95</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Operating earnings</strong></td>
<td>1,334</td>
<td>1,080</td>
<td>752</td>
</tr>
<tr>
<td>Equity in net earnings of associated companies</td>
<td>14</td>
<td>(3)</td>
<td>7</td>
</tr>
<tr>
<td>Investment and other income (expense) - net (note 5)</td>
<td>(16)</td>
<td>47</td>
<td>107</td>
</tr>
<tr>
<td>Interest expense</td>
<td>(131)</td>
<td>(122)</td>
<td>(126)</td>
</tr>
<tr>
<td>Long-term debt</td>
<td>(38)</td>
<td>(35)</td>
<td>(26)</td>
</tr>
<tr>
<td>Other</td>
<td>102</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gain on sale of business (note 5)</td>
<td>102</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Earnings before income taxes</strong></td>
<td>1,267</td>
<td>944</td>
<td>736</td>
</tr>
<tr>
<td>Income tax provision (note 7)</td>
<td>438</td>
<td>321</td>
<td>233</td>
</tr>
<tr>
<td><strong>Net earnings</strong></td>
<td>829</td>
<td>623</td>
<td>473</td>
</tr>
<tr>
<td>Dividends on preferred shares</td>
<td>17</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td><strong>Net earnings applicable to common shares</strong></td>
<td>$812</td>
<td>$619</td>
<td>$469</td>
</tr>
<tr>
<td><strong>Earnings per common share</strong></td>
<td>$1.56</td>
<td>$1.20</td>
<td>$0.92</td>
</tr>
<tr>
<td>Dividends declared per common share</td>
<td>$1.59</td>
<td>$1.23</td>
<td>$0.92</td>
</tr>
<tr>
<td>Weighted average number of common shares outstanding (millions)</td>
<td>522</td>
<td>516</td>
<td>507</td>
</tr>
</tbody>
</table>

Figure 1: Consolidated statements of operations (http://www.nt.com/home/about/97arpt/consolid.html).
Financial statements are generally complemented by notes giving additional information about the figures. The standards concerning the content of annual reports and the calculation of numbers in the statements are nationally defined by standard setters and enforced by law (e.g. in Germany) or by a securities and exchange commission (e.g. in the USA). Financial reporting information is used to evaluate the financial situation of corporations. In order to give recommendations to buy, hold or sell the shares of corporations, financial analysts often use earnings discount models for their evaluation or calculate representative ratios based on accounting data. Therefore data has to be imported into a data processing system. Using a printed annual report, data import has to be done manually. As described above, corporations increasingly distribute digital versions of their annual reports via the Internet. The main advantages of digital annual reports are [for a comprehensive overview see: 16]:

- low cost distribution in comparison to printed annual reports,
- easy and fast worldwide access to the data,
- automated data processing (dependent on the presentation standard) and
- hyperlinking of information and linking to additional external information sources.

The Internet may change the whole communication network within capital markets. On the investor relations sites of companies all information relevant to investors, e.g. financial information, press releases and current share price information can be pooled and hyperlinked. Furthermore, direct communication between investors and the investor relations department via e-mail, mailing lists concerning frequently asked questions and videos from analyst conferences can be provided [6]. In the future online annual meetings and analyst conferences may take place [11].

Due to the potential advantages outlined, it seems reasonable for corporations to use the Internet extensively for corporate reporting. To confront this hypothesis with current practice, the results of a recent empirical study concerning actual practice of Internet-based corporate reporting are outlined in the next section.

3. Internet-Based Annual Reports – The Status Quo

In order to analyze the future role of Internet-based annual reports, it is necessary to identify factors or conditions that influence the potential emergence of such a new genre on the basis of an empirical description of the status quo [21].

The study of Deller, Stubenrath and Weber [6] analyzes the web sites of the 100 US, 100 UK and 100 German corporations representing each country's relevant stock market 100 index (S&P, FTSE and DAX). The study shows that 95 per cent of the US, 85 per cent of the UK and 76 per cent of the German corporations actually are present in the Internet with a homepage. The results concerning corporate reporting via the World Wide Web are summarized in table 1 (calculations by the authors based on data given in the study).

### Table 1. Contents of homepages (relative to the whole national sample)

<table>
<thead>
<tr>
<th></th>
<th>USA</th>
<th>UK</th>
<th>GER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balance Sheet Data</td>
<td>86%</td>
<td>53%</td>
<td>46%</td>
</tr>
<tr>
<td>(summaries and full accounts)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Profit &amp; Loss Account Data</td>
<td>86%</td>
<td>54%</td>
<td>45%</td>
</tr>
<tr>
<td>(summaries and full accounts)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Notes to the Accounts</td>
<td>84%</td>
<td>43%</td>
<td>30%</td>
</tr>
<tr>
<td>(summaries and full notes)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Balance sheet data is supplied by 86 per cent of the US, by 53 per cent of the UK and by 46 per cent of the German corporations. The results concerning the annual report items 'profit & loss account data' and 'notes to the accounts' are as well heterogeneous across the countries. One reason for these international differences might be network externalities due to the much higher level of Internet dissemination in the USA compared to the UK and Germany. But differences with respect to the content of homepages also occur nationally. In all three countries there are corporations using the Internet to present their annual reports (or parts of it) and those who do not. Within the group of corporations that use the Internet for this purpose, the amount of data supplied varies from summary accounts to full accounts and the notes to the accounts are missing in part. The sole presentation of summary accounts on homepages often occurs in Germany and in the UK, while in the USA it represents an exception [see for Germany: 5].

Focusing on the presentation format of financial information on companies’ homepages, the mentioned study comes to the following results. In all three countries the majority of information is given in the HTML format. In addition, some corporations offer downloadable files in other formats like Adobe Acrobat (PDF), Microsoft Excel or ASCII. Table 2 summarizes the results. It shows that PDF is the dominant format for downloadable annual reports.
4. Analysis of Content, Security and Presentation Standards

The study mentioned above shows that Internet-based financial reporting is already common. The question arises whether this evidence is enough to characterize this practice as a new genre in the field of corporate reporting. From an institutional perspective, a new or modified genre can evolve within a communication network like the capital market if a new or modified set of characteristics better supports the needs of all or at least a significant group of the participants. In the context of annual reports we generally refer to this set of characteristics as content, security and presentation standards. To characterize a new genre, these standards have to be stable and sufficiently adopted [21]. In the following we analyze in detail current and potential content, security and presentation standards of Internet-based annual reports.

4.1. Analysis of Content Standards

As described in section 2, the content of printed annual reports is nationally standardized by the respective accounting standard setters. Corporations have the obligation to follow these standards. However, companies select their own content standards by choosing a certain set of information within the range of the given restrictions. Focusing on digital annual reports, the empirical study mentioned above shows that the amount of financial information presented on companies’ web sites often differs from the particular printed annual reports. These deficiencies can easily be solved. Since content standards for corporate reporting are set traditionally by the respective accounting standard setters, these standard setters have to extend the content standards to digital annual reports. In addition, Internet-specific rules, such as those concerning interlinking of information, have to be developed.

4.2. Analysis of Security Standards

Even if companies provide a digital copy with the full information content of the printed report, the digital copy cannot yet be seen as equivalent, since authenticity and integrity problems arise for Internet documents [see for an overview of the security problems: 4]. Integrity in this context means that, to avoid illegal manipulation, any change within digital documents must be discernible. Authenticity means that the receiver of a document must be able to identify the true sender, without any possibility of fraud.

Paper is a static medium. Therefore the printed version of an annual report with an auditor’s report and/or opinion that is signed by the auditor, generates a guarantee of authenticity and integrity of the information given. This does not apply to digital documents. Actual practice shows that auditors’ reports and signatures are digitized and copied to web pages. But Internet sites can change without noise. How can investors and financial analysts be sure that the digital version of an annual report found on the WWW is identical to the printed version of the report and that the auditor’s report and/or opinion refers to the sites currently accessed by users? One may argue that printed annual reports as well can be manipulated after the audit and before the print, but manipulation of digital data is by far easier, and even the homepage and the World Wide Web address (domain name) could be provided by someone else. Security standards in practice are missing at this point of time. Secured electronic versions of auditor’s reports and/or opinions, for example using digital signatures, cannot be found.

In practice, corporations in the USA try to make financial information more reliable via external links to the EDGAR database of the Securities and Exchange Commission [http://www.sec.gov/edaux/wedgar.htm], where text versions of financial information have to be filed electronically. In the UK and in Germany some corporations try to signal reliability by stating on their homepages that the annual report on the Internet sites is identical to the printed version. Others state that only the printed report is legally binding [6]. These approaches do not solve the security problem.

The solution for the authenticity and integrity problem of Internet-based annual reports is the technology of digital signatures [see also: 4]. The corporation’s auditor could easily add a digital signature to an annual report by encrypting the file with his private key. After downloading the document, the receiver can verify its integrity by using the public key of the respective auditor. Manipulation by the management of the corporation or third parties are impossible, any change would be observable by the users. The public key of the accountant can be provided on the homepage of the particular corporation or in any other location. To prove the authenticity of the document, which means that the sender is the actual owner of the public key, the public key and the identity of the owner can be certified by independent certification authorities. In many countries, e.g. in Germany, a reliable infrastructure to make digital signatures legally binding already exists [for the German Digital Signature Act see: 20].

<table>
<thead>
<tr>
<th>Table 2: Presentation modes besides HTML</th>
<th>USA</th>
<th>UK</th>
<th>GER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial Data in Processable Format (e.g. Excel or ASCII)</td>
<td>12 %</td>
<td>4 %</td>
<td>5 %</td>
</tr>
<tr>
<td>PDF Format (Adobe’s Acrobat Format)</td>
<td>32 %</td>
<td>25 %</td>
<td>12 %</td>
</tr>
</tbody>
</table>
Although the technology and legal infrastructure for digital signatures already exist, this kind of security standard is missing in the area of Internet-based corporate reporting. Digital annual reports encrypted with the auditor’s digital signature could not be found on any of the examined homepages. As long as the authenticity and integrity problem is not solved, participants of capital markets will not accept digital versions of annual reports as substitutes for the printed versions. Therefore, accounting standard setters that extend current standards by specifying rules for digital annual reports must also consider the technology of digital signatures.

4.3. Analysis of Presentation Standards Currently Used

Assuming that it is easy to print information provided on the Internet, a digital version of an annual report is a perfect substitute for the printed version, as long as it contains the same information and authenticity and integrity is secured. It is rather irrelevant which presentation format is used. The presentation format gains importance if the focus is on further advantages like easy access to data via the Internet and the facilitation of information retrieval and further procession. All potentially result in cost reductions for providers as well as users of the information, thus giving added value to corporate reporting. The results of section 3 show that currently different data formats are used. This diversity is a general disadvantage for users. Professionals that want to satisfy their information demand by using the Internet must deal with different software to download, display and process the data. Furthermore, the formats presently in use cannot be sufficient. The reason for this is that the different formats only provide a part of the possible benefits of digital data distribution via the Internet.

Most of the financial information offered on web sites of corporations is presented by using the HTML (Hyper-text Markup Language) standard, which currently is the most common standard to present information on the WWW. Figure 1 (in section 2) displays part of the digital version of the Northern Telecom 1997 annual report [http://www.nt.com/home/about/97arpt/consolid.html]. This document, like all HTML documents, consists of text, graphics and formatting information (so called tags). There are also hyperlinks that point to other documents for further, more detailed information. The example below displays a part of the HTML code of the consolidated statement of operations. For reasons of simplification we eliminated some of the HTML tags that define the table display.

```
<table>
  <tr><td><font size=-1>$10,672</font></td><td><font size=-1>$12,847</font></td><td><font size=-1>Revenues</font></td></tr>
  <tr><td><font size=-1>$15,449</font></td><td><font size=-1>$10,672</font></td><td><font size=-1>Cost of revenues</font></td></tr>
  <tr><td><font size=-1>$16,379</font></td><td><font size=-1>6,379</font></td><td><font size=-1></font></td></tr>
</table>
```

Although the HTML format - especially with its hypertext structure - has great advantages, the example of the Northern Telecom consolidated statement of operations shows the deficit of HTML concerning the exchange of structured data. Besides hyperlinks, the tags only define the display format of the text like fonts, size and how the table is displayed. The data is not at all content encoded, i.e. structured or indexed, so that it could be processed by the receiver’s computer system in an automated way. The receiving computer would, for example, not be able to identify the revenues of the Northern Telecom in 1997 automatically. HTML simply does not provide the possibility of structuring data by specifying semantic elements. Using HTML to present financial data via the Internet enables all the advantages of hypertext structured financial information on the WWW, but it cannot enable the automated processing of financial data by the computer system of the receiver.

Many corporations offer downloadable files of financial information. One of the most common formats in this context is the Portable Document Format (PDF). PDF documents are independent formats written by Adobe’s Acrobat software, which retain original fonts, colors, formatting and images on multiple computer platforms [http://www.acrobat.com]. To view, navigate, and print a PDF document the Adobe Acrobat Reader is necessary, which is downloadable from Adobe Systems at no charge. It is also possible to view and print PDF documents with certain WWW-browsers by using a plug-in. The PDF format allows the provision of digital annual reports identical to the printed version, since all the fonts, colors, images and graphics of the document are embedded. Potential time and cost savings are obvious. But other conceivable fundamental advantages of the Internet-based distri-
bution of annual reports are missing. Although new versions of PDF do allow the hyperlinking of text, it is not very common to do so. None of the PDF formatted annual reports that were found on the examined homepages of corporations were hyperlinked. Furthermore and even more important, PDF does not enable further automated data processing.

Other formats of downloadable documents actually offer the option of further processing. Some companies structure the financial information in Microsoft Excel spreadsheets or comma-delimited text files that can be imported into the database of the receiver’s system for further processing. These formats imply great potential benefits especially for financial analysts that transfer financial data into their own financial indicators. However, those attributes that characterize information on the WWW like hypertext and user friendly visualization of data are currently neglected.

Table 3: Benefits of data presentation formats

<table>
<thead>
<tr>
<th></th>
<th>Hypertext</th>
<th>Information retrieval</th>
<th>Further automated processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDF</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excel, ASCII</td>
<td>●</td>
<td></td>
<td>●</td>
</tr>
<tr>
<td>HTML</td>
<td>●</td>
<td>●</td>
<td></td>
</tr>
<tr>
<td>XML?</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

Table 3 summarizes the results of our examination. Possible benefits of the digitalization of reports are only partly realized by the different data formats. The potential advantage of combining the HTML format with a format that allows automated further processing of the data is not yet possible.

It is obvious that even if standard setting institutions successfully extend content standards and develop security related standards for Internet-based annual reports, the emergence of a new or modified genre is not guaranteed, because of the lack of sufficient presentation standards.

There are some efforts to solve the problem by developing Internet solutions for corporate data. In the US, for example, the National Science Foundation (NSF) sponsored a project starting in 1993 to provide the corporate disclosure documents of the EDGAR database (Electronic Data Gathering Archiving and Retrieval system of the U.S. Securities and Exchange Commission) over the Internet. The solution enables the Internet-based data transfer (data provided as ASCII with limited SGML type tags), but has significant shortcomings concerning the effective automatic data extraction [12]. In the final part of this work we introduce a new Internet data format called XML, which could close this gap.

4.4. XML: The Future of Internet-Based Reporting?

XML (Extensible Markup Language) is a new Internet presentation standard, which was developed by the XML Working Group of the World Wide Web Consortium (W3C) [For further information see: http://www.oasis-open.org/cover/xml.html]. It is a simplified version of SGML, which represents structured data in the World Wide Web. It is completely described in the XML Specification, of which the first version was released on W3C Recommendation February 10, 1998 [http://www.w3.org/TR/REC-xml]. XML is a markup language. Like HTML it can be commonly used to display and hyperlink documents in the World Wide Web. Unlike HTML, which contains only a fixed set of format-oriented tags that do not allow content related encoding [7], XML supplies the means to enrich the data with semantic information.

The benefits of XML are in summary the following [17]:

- XML provides a structural representation of data, which can for example be used for data transmission between client and server, for sharing data between applications or for persistence data storage.
- Data displayed on the desktop can be manipulated locally and can be processed by local applications.
- Authors can create self-describing XML documents by enclosing either Document Type Definitions (DTD) or alternatively create XML-Data schemas.
- XML data is independent from any presentation format, which means that it can be displayed in various formats (including HTML) on the desktop.

XML has been developed recently and only few XML applications can be found at the moment, most of which have exemplary character only. Nevertheless XML has a promising future. Large software companies like Microsoft and Netscape are currently integrating XML in their WWW-browsers. In addition, they have started to develop complementary standards (like XML-Data [15]) and applications. In the following, we analyze how XML can lead to improvements in the area of Internet-based corporate reporting.
Figure 2 shows that the information structure of a typical annual report is a tree structure. It is easy to identify elements like the consolidated statement of operations and subelements like the cost of revenues. The elements used in annual reports in part vary across industries and/or companies. Nevertheless, a closer look reveals - especially within industries - a typical set of elements (the majority of them are explicitly required by accounting standards). The tree structure makes it very simple to structure annual reports using XML. XML does not have predefined elements or element relations. The author declares his own document types and the element definitions in the so-called Document Type Definition (DTD). Using XML, the typical tree structure of annual reports can be incorporated in standardized DTDs. Clients can configure according to the this and enable client-side processing of the information [13]. Using XML, annual reports could be distributed over the Internet and then, as a result of the self-describing data structure, be easily imported into applications of analysts and investors. The following code shows exemplary and simplified how the DTD of an annual report could look like.

```xml
<?xml version="1.0"?>
<!DOCTYPE ANNUAL REPORT [  
<!ELEMENT ANNUAL REPORT (CONSOLIDATED STATEMENT OF OPERATIONS, CONSOLIDATED BALANCE SHEET)>  
<!ELEMENT CONSOLIDATED STATEMENT OF OPERATIONS (REVENUES, COST OF REVENUES, GROSS PROFIT, ANNOTATION*)>  
   <!ELEMENT REVENUES (#PCDATA)>  
   <!ELEMENT COST OF REVENUES (#PCDATA)>  
   <!ELEMENT GROSS PROFIT (#PCDATA)>  
   <!ELEMENT ANNOTATION (#PCDATA)>  
<!ELEMENT CONSOLIDATED BALANCE SHEET (ASSETS, LIABILITIES)>  
   <!ELEMENT ASSETS (CURRENT ASSETS, LONG-TERM RECEIVABLES, INVESTMENTS)>  
      <!ELEMENT CURRENT ASSETS (INVENTORIES, PREPAID EXPENSES)>  
         <!ELEMENT INVENTORIES (#PCDATA)>  
         <!ELEMENT PREPAID EXPENSES (#PCDATA)>  
   <!ELEMENT LONG-TERM RECEIVABLES (#PCDATA)>  
   <!ELEMENT INVESTMENTS (#PCDATA)>  
   <!ELEMENT LIABILITIES (...)>  
]>
```
A DTD defines the elements of a XML document, how elements might be nested and what rules have to be obeyed. In annual reports we find, for example, elements like the consolidated statement of operations or the consolidated balance sheet. The elements can be nested. In this example, the element CONSOLIDATED STATEMENT OF OPERATIONS contains the elements REVENUES, COST OF REVENUES, GROSS PROFIT, and ANNOTATION.

<!ELEMENT CONSOLIDATED STATEMENT OF OPERATIONS (REVENUES, COST OF REVENUES, GROSS PROFIT, ANNOTATION*)>

The tree structure is obvious. The expression #PCDATA stands for parsed character data or text. Elements can be further specified, for example, by data elements quantifiers. In our example the * symbol means that the element ANNOTATION is not obligatory, but that there might be one or more than one occurrence of this element. In the context of annual reports, such specifications can be used to indicate that certain elements may or may not, depending on the individual corporation, occur in the report. Therefore annual reports of different corporations, which are structured similarly, but contain different elements, can be described by the same DTD. This could be a great advantage in the context of digital annual reports. Different corporations often use different denotations for the same information. As an example the term “consolidated statement of operations” in our example is equivalent to the term “consolidated income statement” or “profit and loss account”. With XML it is possible to use the same DTD for elements with the same structure, but with different names. It does not matter how corporations name their information as long as it is structured the same way.

The DTD can be defined within a document or referenced and accessed externally. This means that a common DTD for annual reports located somewhere in the World Wide Web could be used to define elements and rules for individual annual reports. If such a common DTD contains all the possible figures that corporations might use in their reports, digital versions of reports referring to that DTD could always be further processed by any computer system. Since financial accounting information is standardized by regulations to a relatively high degree, the development of such a DTD seems to be achievable.

The following code exemplary shows how some data of the Northern Telecom annual report structured by a common DTD for annual reports could look.

```xml
<?XML Version="1.0">
<!DOCTYPE ANNUAL REPORT SYSTEM ".../DTD/report.dtd">

<ANNUAL REPORT>

<CONSOLIDATED STATEMENT OF OPERATIONS>
  <REVENUES>15,449</REVENUES>
  <COST OF REVENUES>9,111</COST OF REVENUES>
  <GROSS PROFIT>6,338</GROSS PROFIT>
</CONSOLIDATED STATEMENT OF OPERATIONS>

<CONSOLIDATED BALANCE SHEET>
  <ASSETS>
    <CURRENT ASSETS>
      <INVENTORIES>1,765</INVENTORIES>
      <PREPAID EXPENSES>155</PREPAID EXPENSES>
    </CURRENT ASSETS>
  </ASSETS>
  <LONG-TERM RECEIVABLES>334</LONG-TERM RECEIVABLES>
  <INVESTMENTS>285</INVESTMENTS>
</CONSOLIDATED BALANCE SHEET>

</ANNUAL REPORT>
```
The data is structured using the external DTD report.dtd, which might be stored at the local computer or elsewhere, e.g. for the USA, at the Financial Accounting Standards Board’s homepage. Using this DTD, a local XML processor of the receiving computer system can automatically verify and process the data embedded in a XML file.

How can the use of XML lead to a greater acceptance of Internet-based annual reports? How can it provide the added value, which could lead to the emergence of a new genre? XML provides a standard to structure data independent from any presentation format. Since XML is able to represent a range from simple data structures like database records to complex structures like directed graphs [14], it can be used as a general storage format. To display data, it can easily be transformed into various formats. The W3C group currently discusses the specifications of XSL (“Extensible Style Language” or “Extensible Stylesheet Language”), which will be used to display structured information like XML-DATA. XSL is a stylesheet language, which is based on DSSSL (Document Style Semantics and Specification Language) and which is designed for the Web community [for more information see: 3]. Thus the presentation of structured data like annual reports on the WWW becomes very simple.

Furthermore, XML enables the structured, automated data exchange using the World Wide Web. Data can be transferred directly from the desktop of investors or analysts into their applications. This is the most significant advantage of XML. Today investors and analysts have to input financial data from annual reports manually, which is very costly. Another example are news agencies like Bloomberg or Reuters. Bloomberg for example imports (manually) the data from annual reports of several thousand corporations worldwide. The potential savings through automated data processing are obvious. In addition, XML may even lead to changed working conditions for financial analysts. It enables the integration of the sender’s and the receiver’s applications and additionally provides all other possible benefits of digitized data like advanced graphical user interfaces, hypertext and information retrieval.

5. Concluding Remarks

Capital markets are based on complex communication networks. Within these networks the annual report represents one of the most important communication tools regarding financial information about corporations. In addition to the traditional printed versions, a growing number of corporations provide their annual reports on their homepages. The acceptance of these digital documents is limited due to deficiencies concerning content, security and presentation format. We show that accounting standard setters that want to extend current standards by specifying rules for digital annual reports must consider content-related standards as well as the technology of digital signatures. Based on the specification of such rules the use of the new emerging Internet standard XML could add significant benefits to digital reports and therefore trigger the emergence of a new genre. Using XML to structure information on the World Wide Web can enable automated processing of financial data and reduce information costs of capital market participants. Since XML is a new standard, which is not yet sufficiently adopted, the future will show whether it will be used in corporate reporting and whether its potential benefits will be exploited.

Acknowledgements

This work is part of the research projects B3 “Economics of Standards in Information Networks” and B4 “Standardization of Information Intermediation in International Capital Markets” of the interdisciplinary research program “Competitive Advantage by Networking - the Development of the Frankfurt and Rhine-Main Region”. We gratefully acknowledge the financial support from the German National Science Foundation (Deutsche Forschungsgemeinschaft). We are also very grateful to Dominic Deller, Mark Ginsburg, Frank Ladner, Christoph Weber, Tim Weitzel, and the anonymous referees for their comments.

References


