



HAL
open science

An Electronic Document for Distributed Electronic Services

Gerard Wawrzyniak, Imed El Fray

► **To cite this version:**

Gerard Wawrzyniak, Imed El Fray. An Electronic Document for Distributed Electronic Services. 15th IFIP International Conference on Computer Information Systems and Industrial Management (CISIM), Sep 2016, Vilnius, Lithuania. pp.617-630, 10.1007/978-3-319-45378-1_54 . hal-01637502

HAL Id: hal-01637502

<https://inria.hal.science/hal-01637502>

Submitted on 17 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

An Electronic Document for Distributed Electronic Services

Gerard Wawrzyniak¹, Imed El Fray¹

¹ West Pomeranian University of Technology, Szczecin
Faculty of Computer Science and Information Technology, Poland
{gwawrzyniak@ebstream.com, ielfray@zut.edu.pl}

Abstract. The paper presents the role of documents in the implementation of various types of transactions. The main features of the document determining its usefulness in the effective exchange of legal information, ensuring the authenticity, integrity and non-repudiation of origin are presented. Considering the general background of the document, the concept of an electronic document having significant (in terms of legal effectiveness) features of traditional document as well as those features that allow its operation and processing in the virtual space has been presented. An example of the use of an electronic document in the implementation of typical transactions, which are reflecting traditional electronic transaction will be discussed as well.

Keywords: traditional document, electronic document, electronic transaction, XML, electronic services.

1 Introduction

Every business, regardless of the area in which it is conducted (business, administration, etc.) apart from the flow of material goods or performance of services, is associated with transfer of capital and information.

From the nature of things, the production and delivery of products and service delivery are

a tangible – real. The flow of capital used to be implemented in the real world as the physical transfer of money in the past, now is almost completely dominated by the flow of information in virtual reality (electronic). The exchange and processing of information accompanying the economic processes, administrative processes etc., at all stages of their execution is carried out in many ways: talking, letters, faxes, e-mail - can be enumerated, but the importance of information is relevant to from the point of view of legal effectiveness [1-3]. At some stage of the business process execution there is a need to ensure the legal effectiveness of information being exchanged at least for purposes of legal proceedings, in case of differences in the interpretation of

Corresponding author: Imed El Fray, e-mail address: ielfray@zut.edu.pl, full postal address: West Pomeranian University of Technology in Szczecin, Faculty of Computer Science and Information Technology, ul. Zolnierska 52, 71-210 Szczecin, Poland.

events and commitments undertaken by parties. Such possibility is provided by document that regardless of its form (technical implementation) must ensure authenticity, non-repudiation of origin, integrity and privacy [4-6].

Furthermore, the document as an electronic document must ensure the possibility of electronic transport, storage, and first of all the processing using electronic means.

The existing formats of electronic documents actually files, do not meet requirements discussed above.

For example, the format of Microsoft Office "*.doc, .docx" or Libre Office (Open Office) does not provide legal effectiveness (no electronic signatures applied), it also do not give the possibility of electronic data processing.

Adobe format *.pdf in its basic form provides the ability to sign a document (whole content) using number of signatures - but it can not generate the signatures of various parts of a document by different people. These forms do not reflect the real (paper) forms in which it is possible to perform such signatures. It is also possible to use XFDF format, based on XML and representing data and annotations, which are included in PDF form file. Automatic processing of data from a PDF document requires a conversion (data mining) document into an XML data file (XFDF).

Electronic document, especially the one that is used by human is composed of presentation layer and data layer. It should reflect the traditional documents and forms in which these two layers are contained in a single object - a piece of paper (or multiple pages) - this is the immanent feature of a document. Mentioned above formats of documents meet this requirement. However, as already mentioned, do not meet critical requirements (legal effectiveness, ease of electronic processing). Such requirements would meet the electronic document based on XML (ease and interoperability of processing, eith legal efficiency provided by electronic signatures (many signatures covering different parts of the document, done by many people) with the use of PKI, including in its structure both description of visualization and introduced by human data.

A similar solution is presented in the XFDL form (Extensible Forms Description Language) proposed by Canadian company Pure Edge [7]. In XFDL forms both data and presentation of data is stored in one XML file which syntax described using the DTD. This solution does not provide the possibility to define and then implement such features: as default values of each form field, validation rules of the data, help messages and hints. In the XFDL forms it is not possible to define a method to build XML messages based on data stored in the form, and support protocols to exchange information with network services (eg. SOAP), which would allow to build messages, sending them to the server, receiving a response in real time as synchronous mode and present (update) data in a form for presentation.

The aim of the article is to present, with regard to described above drawbacks, the format of an electronic form with syntax formally described usig XML Schema XSD. The form contains description of the form visualisation, data types of form fields, and electronic signatures with indication of elements sign by each signature, calculation rules and also the following new features (relating to the XFDL), such as definitions of default values, validation rules, mechanisms for building XML messages, descriptons how to send XML (or the entire form) using SOAP or SMTP (the list of recipients, encryption) protocols, a mechanism for interpreting XML messages received in SOAP response. These features enable to implement electronic services of

transactions in real time.. Other aim of the paper is to show possibility an benefits of such document in distributed services definition and execution.

2 Document in electronic service delivery

On behalf of the European Commission, the company Cap Gemini Ernst & Young, has developed a 4-point scale assessment of the progress of the e-Europe initiative for evaluating the availability of online services for citizens and businesses [8]:

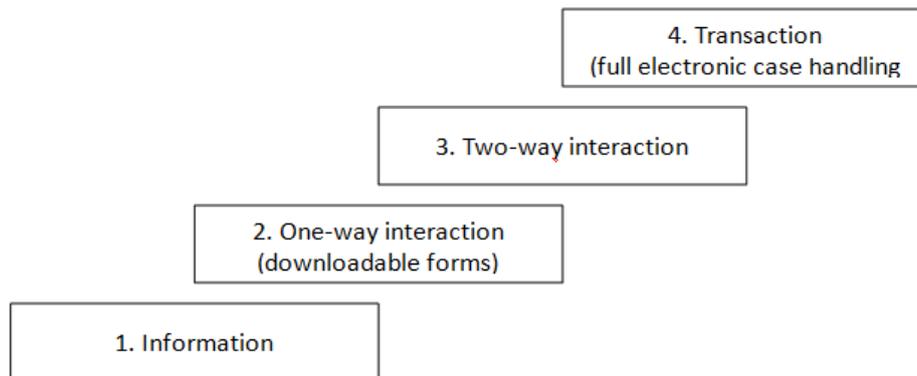


Fig 1. Maturity levels of electronic services

0 – The service is not (in any form) provided electronically:

- no publicly accessible website/s, or those websites does not meet the criteria for levels 1-4;

1 – Information:

- publication of electronic information on the service provided;

2 – One-way interaction:

- the possibility to download forms;

3 – Two-way interaction

- the possibility to fill out forms or electronic submission of the application or requests;

4 – Transaction

- full electronic execution of the case, taking into account the decision, informing the recipient, delivery and payment if applicable;

Levels of electronic services maturity are shown in the figure (Fig.1)

Achieving the highest level does not always mean executing the service at the transaction level, for example, some services are one-sided information reported to the authority and do not require a decision. Therefore, also among the criteria's for assessing the progress of e-Europe, used by the Commission of the European Union a key place takes accessibility for citizens and businesses to public services and the quality of these services. However, you can see that the achievement of 3rd and especially 4th maturity level requires:

- ensure the authenticity, integrity and non-repudiation of origin of exchanged information it means the legal effectiveness of that information, and
- the need for automation of the service and hence the need to automate the processing of information associated with the service execution.

Public services should therefore be considered in these two dimensions: in terms of legal regulations and technical implementation.

3 Document and its features

The document is any tangible or symbolic indication, fixed or saved in order to prove the phenomenon of physical or mental [9], it means that document is a resource used as evidence. However, the concept of the document in the language of computer makes some problems, it is commonly equated - the concept of electronic document and the concept of the file.

The file is a collection of data stored in permanent memory (hard disk or other media), which can be seen by the user of the operating system as a unitary, indivisible, closed as a whole, having a distinct name and a specific structure. It is the result of working with a computer program, writable and then readable.

It seems that the most preferred would be to combine the two approaches, and without changing the concept of the primary document concept, state that:

- An electronic document is a document in the form of a file, which has all the features of the document and all the features of the file.
- Data are stored on a medium. In the case of electronic information the data can be processed, communicated and interpreted by computer. A set of data may sometimes constitute a record [10].
- The international Council on Archives [10] Committee on electronic records defines a record as ‘a specific piece of recorded information generated, collected or received in the initiation, conduct or completion of an activity and that comprises sufficient content, context and structure to provide proof or evidence of that activity’.

Legal effectiveness of the document is ensured by document features:

- integrity - the immutability of content (document content),
- authenticity - by knowledge of the origin of the document,
- non-repudiation - no possibility to denial of the document authorship.

These definitions and referenced to the concept of the document. These features are implemented using a handwritten signature or electronic signature. Moreover, as already stated above, documents have features such as:

- Durability - the possibility to store on a durable medium (in relation to traditional documents, such as paper, parchment, papyrus, stone tables, images etc., in relation to an electronic document can be a disk, tape, memory, etc.)
- Autonomy - it constitutes a single, closed as whole, standalone entity (independent, autonomous - data and visualization is one object).

- Interoperability - readable / interpretation by any of the entities involved in the transaction

If the document is to be used in the execution of electronic services for the exchange of legally binding information, it must be possible to.

- its transmission in electronic form,,
- automatic reading and writing of data in file, automatic processing across different systems,
- automatic verification of features for its legal effectiveness,
- recording and storing in electronic warehouses (databases, file systems).

General features of document and specific features of electronic document are depicted on the following figure (Fig 2):

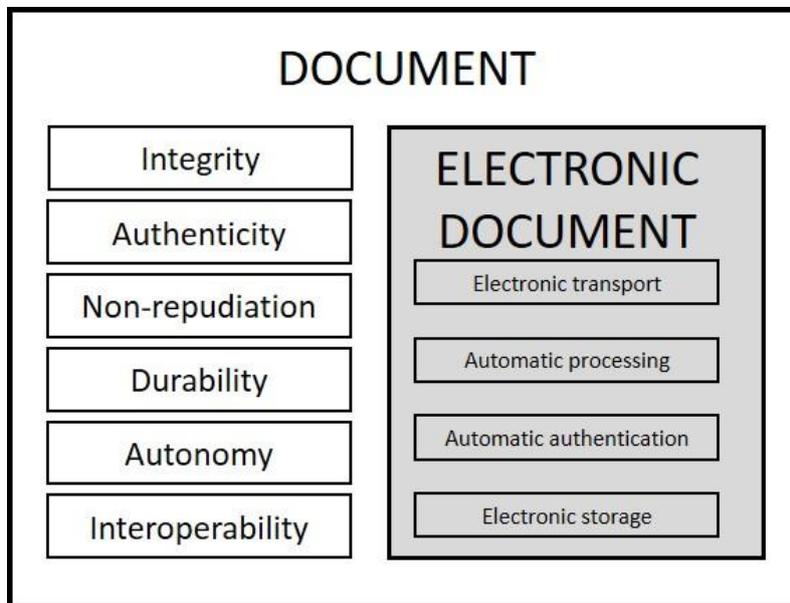


Fig.2 Each electronic document is a document and has all of its features

The European Union Directive [11] and its national implementations in the form of national acts [12], equates in terms of legal effectiveness, the handwritten signature and the electronic signature. Thus electronic signature plays a key role in ensuring the legal effectiveness of an electronic document. In addition, it is possible and even desirable to implement other features of the handwritten signature in electronic signatures, such as:

- the possibility of signing one or more parts of the document by one signature,
- ability to place in the document number of signatures (signing - including the indicated part of the document)
- possibility of countersigning.

Such possibilities are given by XML Signature format – XMLdSig [13] and its development – XAdES [14]. Below, see figures Fig.3 and Fig.4, the example of model-

ling the contents of signature in electronic ebForm document is presented (ebForm is xml format of electronic document in the shape of electronic form defined by author:

```
<ds:XPath>(ancestor::*//parent::value or ancestor-or-self::ds:Signature/ancestor::signature//parent::ebForm or ancestor::content/ancestor::attachment or ancestor::*//ancestor::*//ancestor::mailStatus or ancestor::*//ancestor::value/ancestor::soap-RPC or ancestor::body/ancestor::request or ancestor::header/ancestor::request or ancestor::body/ancestor::response) or (ancestor::*//parent::value/ancestor-or-self::*//signedBy/signer[text()='clientsignature'] or ancestor-or-self::ds:Signature/ancestor::signature/countersignedBy/signer[text()='clientsignature'])</ds:XPath>
```

Fig.3. Example of definition of signature content (XML element XPath) signing the client signature (one of the many users) in the ebForm form)

```
<ds:XPath>(ancestor::*//parent::value or ancestor-or-self::ds:Signature/ancestor::formSignature or ancestor-or-self::ds:Signature/ancestor::signature//parent::ebForm or ancestor::content/ancestor::attachment or ancestor::*//ancestor::*//ancestor::mailStatus or ancestor::*//ancestor::value/ancestor::soap-RPC or ancestor::body/ancestor::request or ancestor::header/ancestor::request or ancestor::body/ancestor::response)</ds:XPath>
```

Fig.4. Example of definition of signature content (XML element XPath) signing the form template in the ebForm form

4 Syntax and semantics of data in electronic document

Each file, holds in its content some information is built according to specific rules. These rules define the file format and are composed of the syntax, which is the definition of relationships between elements (data) stored in the file, and the semantics (meaning), a description of the interpretation of the data placed in certain syntactic constructions.

To ensure unambiguous and clear interpretation (by peoples and systems) of information contained in the file, the file format must be used in accordance with principles set out in the specification of the file format.

The problem of determining the format of the data containing information necessary for the implementation of electronic services, is crucial from the point of view of interoperability and universality. It is not a major problem, through the widespread use for this purpose XML. XML has become the informal but extremely popular and the applicable standard [15-17].

In addition to many features of XML, such as:

- platform independence and software being applied using,
- support for a universal set of national characters UNICODE
- clarity and ease of processing.

Special attention should be paid to:

- the ability to define the syntax and automatic verification of information structure with the syntax,
- the ability to import syntactic structures defined in another place that is possibility to use data structures already defined,
- the ability to identify both the document and components imported using namespace (namespace) [18].

At present the centralized model of electronic services implementation and execution is used in Poland.

This is such approach to the construction of electronic public services, in which the document formats, data dictionaries and ways of delivering services are defined, implemented and executed by a designated central organization and other organizations have to subordinate to this central organization. This approach causes the separation of the responsibility for the implementation of services in its real terms from implementing the same services in the electronic (virtual) dimension. The model is complicated by the possibility of existing a few centers (tax authorities, local authorities and government) responsible for different areas of activity This is shown in the Fig.5.

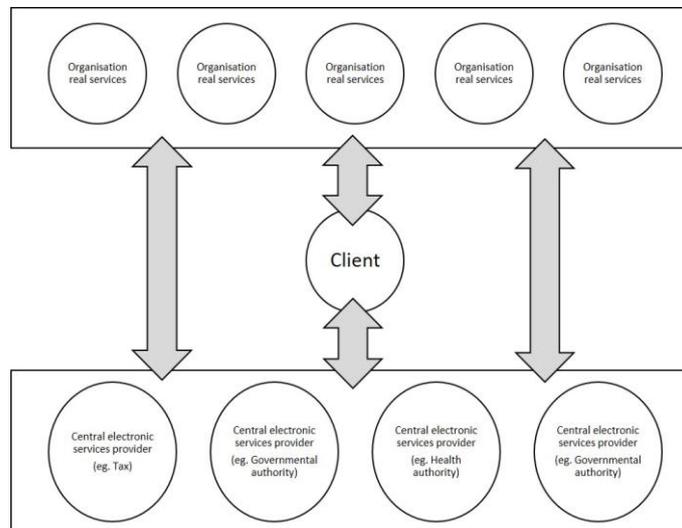


Fig.5. The centralized model of electronic services - the existence of several different centers is possible

Considerations in the area of syntax and semantics of the data in an electronic document and the features and capabilities of XML are of particular importance in light of the concept of decentralization of public services and posting obligation to implement them to the places (institutions, organizations), where those services are actually carried out – see figure Fig.6.

The decentralized model of defining the specifications and publishing electronic documents, is presented in this article assumes that:

- document formats (syntax and meaning) are defined and published by these organizations, which are responsible for the implementation of electronic service using defined document,
- data structures and data are defined in organizations using namespace that has the form of a URL, as in the domain name owned by the organization responsible for implementing the service. This provides a unique and unambiguous terms used in the exchange of documents.

- Appearance (visualization of the document), taking into account specific regulations as well as freedom of visualization within a certain range (tax form, invoice)

Using the possibility of identifying formats (syntax and semantics) of specific documents and data using namespace makes it possible to avoid the need for a use the central repository of model (specifications) of documents (or electronic forms). Namespace has the form of a URL, which part is the domain name owned entity responsible for the specification of certain documents. It is therefore possible "exit" the responsibility for defining these specifications and definitions of the central office to the organization responsible for implementing the service, ensuring the uniqueness of identifiers and compliance concepts contained in an electronic document with the terms defined in the act regulating the implementation of the service.

Practical use of the principles presented above, would be the universal application of the rule of expansion of legal acts which describe the public service, by the entity implementing this service by the syntax definition of electronic documents (in the form of a DTD or XML Schema) with a description of its meaning. A description of the syntax and semantics would be a complement to this part of the act, where documents and forms in its traditional form are described

Such an approach would at the same time the specification of requirements for software vendors in regards not only to data formats and visualization, but also the acceptable range of visualization freedom (eg. strictly defined tax form, or the appearance of freeform invoice).

The decentralization of the electronic services execution is:

- reflection of the actual geographical and functional decentralization of services delivery;
- linking responsibility for the electronic dimension of the service with the real service (one entity responsible for the execution of services in real and electronic dimension);
- The implementation of electronic services, complementary/extension services in a manner consistent with local customs and requirements.

Easier and unambiguous definition of the concepts related to the particular service - these definitions, as a part of the act regulating the implementation of services are created in the same place and in the same act.

A special, very important role in the exchange of information associated with the delivery of services regardless of method of realization: whether it is e-service, or provided traditionally, regardless of the model - the central or decentralized, regardless of the area (industry), plays the form. Natural need to enter data in the appropriate space limits the possibilities of interpreting the meaning of these data so that the form is a convenient means to collect the data necessary for the service, as well as the processing of these data, which is important especially for automatic execution of these processes. Taking into account the fact that the currently performed service may be fully or partially carried out automatically and partially, by human it would be relevant to make use of such an electronic document (form), with the use of suitable

software, which could be used by a man and simultaneously processed in computer systems.



Fig.6. Decentralized (distributed) model of electronic services – organization delivers services in real and electronic form

5 Electronic form

Electronic document containing the information necessary to execute the service is processed automatically. Identification of the document type and the particular data using namespace allows unambiguous interpretation of these data (interpretation is published in the type specification document and published in the act of describing the realization of the service as described in the discussion on the implementation of services in the decentralized model) by all entities taking part in the service. The use of a document in XML format syntax as XML schema, makes possible to automatic verification of document syntax and automatic document processing.

If documents are exchanged between systems that implement the services (or services that form part of other service), then visualization of the data is not necessary, because the machine executing the service has been implemented in accordance with publicly available specifications described above.

However, if the party taking part in the execution of the service is the human, the document must be presented to the user. Human must understand the document, type the information to it and consciously sign this document.

Similar to the traditional paper document, the most convenient form of the document for this purpose is the form. The electronic form it is a document containing free places for data representing desired information, and to ensure the characteristics of the document - signature. The data are entered in the form in which the appearance of the form and arrangement of fields for entered data determine the semantics of these data. In terms of the form it is a document in which visualization and data are one inseparable entity, in which the semantics of data is determined by the graphical environment of the data. Thus, if an electronic document will contain both the data and its presentation it will be electronic implementation - a metaphor of the traditional paper

form readable / understandable by man and at the same time possible to automatic processing using commonly available tools.

It should be noted that the concept of an electronic form presented above, in accordance with its traditional understanding deviates from the currently existing electronic forms, in which an separation of presentation layer and the data layer is implemented (eg. An HTML in which "forms" are used to collect data to be sent to the server, or separating data from presentation of XML documents with reference to the transformation file XSLT). Visualization layer is responsible for presenting the form view, the user enters the required data, which are sent to the service provider (after separating them from the presentation layer - the graphical environment determining the semantics of these data). This approach prevents a uniform semantic interpretation of the transmitted (stored) data.

In the proposed solution the form is an electronic document based on XML, with syntax defined using XSD [XML Schema Definition], applying electronic signatures based on XMLDSig/XAdES [XML Advanced Electronic Signatures], containing both the data and the description of the presentation (visualization) of the form in one single file – see Fig.7.

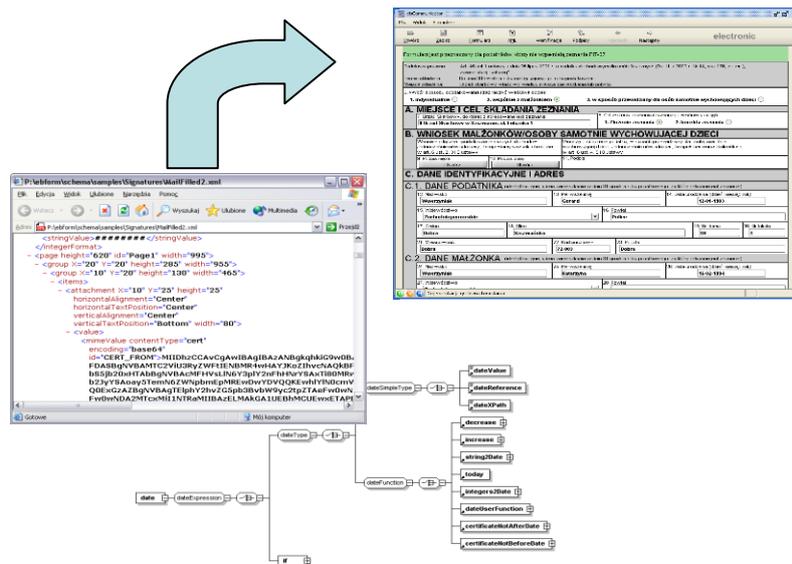


Fig.7. Electronic form (syntax - XSD, data XML, presentation XML) – example of Polish tax declaration form

6 Features of electronic form

In addition to its "traditional functionality" electronic form can and should be enhanced with the features offered by the Information Technique such as:

- possibility to verify the formal validity of data,
- possibility to define default values,

- possibility to define of help and tips,
- providing a dynamic (or lock) input fields.
- calculating the value of each field based on the data entered into other fields.
- ability to define the mechanisms of communication and ways of exchanging data between the form and the environment through a well known protocols as email or web services protocols.

All these features are part of the form and are part of the document together with the data and presentation. This is due to the need to ensure that the description of the method of processing and presentation of data is accessible and known to all users in order to avoid doubts in interpretation.

In the following figure (Fig,8) the example of definition of text field ("textField" element) with its properties is presented (ebForm format).

```

<textField X="0" Y="2" height="20" horizontalAlignment="Right"
id="Pismo.MiastoPowstania.Komponent" width="175" xsi:type="stringField">
  <value>
    <stringValue id="Pismo.MiastoPowstania">Łęczna</stringValue>
  </value>
  <defaultValue>
    <substring>
      <stringReference>Urzad.KodMiasto</stringReference>
      <beginIndex>
        <integerValue>7</integerValue>
      </beginIndex>
      <endIndex>
        <stringLength>
          <stringReference>Urzad.KodMiasto</stringReference>
          </stringLength>
        </endIndex>
      </substring>
    </defaultValue>
    <background>
      <colorValue id="Kolor.Pol">218 231 248</colorValue>
    </background>
    <helpText>
      <stringValue>Należy wpisać nazwę miejscowości.</stringValue>
    </helpText>
    <tipText>
      <stringValue>Należy wpisać nazwę miejscowości.</stringValue>
    </tipText>
    <validity id="Pismo.MiastoPowstania.Poprawnosc">
      <rule>
        <condition>
          <not>
            <isEmpty>
              <dateReference>Pismo.MiastoPowstania</dateReference>
              </isEmpty>
            </not>
          </condition>
          <description title="Brak miejscowości przy dacie">Należy wpisać
nazwę miejscowości przy dacie powstania dokumentu.</description>
        </rule>
      </validity>
    </textField>

```

Fig.8. Example of defining the features in the form (value, default value, background, help text, text, hints, validation rules)

Form features defined in this way, make it possible to use the form as an information medium that is synonymous with the traditional document (form) and to use the form as a intelligent interface to services, where form is a generator of messages sent out, or processor of message received from service, and involving human.

Very wide range applications of electronic forms could be enumerated. Such forms is a generic solution with can be used in different areas regardless of business area – form small companies for ordering and invoicing, thru activities associated with internal company management up to support for Client – Business, Client – Administration and Business – Administration operations. It can be applied wherever legally effective information is required or forms are being used (exchanged) both by human and machines, or by human only.

7 The use of electronic forms – applications/implementations

Electronic forms based on the concept presented in the article can be used wherever legal effectiveness is required, regardless of the area of activity and the size of the business.

Electronic form looks exactly like its traditional counterpart. The logic of usage is also identical.

It can be transferred/delivered to recipient using widely known mechanisms like: simple mail as attachment to mail message or SOAP as Soap envelope content. It also can be transferred as ordinary file.

It can be simply processed by information systems using standard XML parsers and then integrated with internal organization's systems including data retrieval and transferring data to internal databases. For example:

- Exchange documents (agreements, orders, others) using SMTP, as attachment,
- Implementation of simple document flow even composed of few steps of document exchange within company or between companies (holiday arranging, internal reports etc...) and integration with internal systems and databases using SMTP or services implemented using SOAP protocol.
- Legally effective delivery of a document to the public authority with the official synchronous (immediately) confirmation of delivery / reception. The confirmation is digitally signed part of the document, and document is automatically integrated with internal authority system,
- Legally effective delivery of a document to resipient (human) using SMTP with confirmation signature of recipient (client) returned to sender.
- Implementation of fiscal charges undeniably linked with the document.

This shows, that presented concept of the form, in contrast to formats presented in the introduction (.doc, odt, pdf/xfdf, xfdl), allows:

- exact implementation of traditional documents.
- automatic document processing.
- automatic signatures verification (authorization, authentication).
- document integration with existing systems.
- ensuring the legal effectiveness of the document.

8 Summary

Electronic services are typically implemented using dedicated, closed - demanding and accounts of users. Legal effectiveness of documents used for the transaction is

provided only in the space of the system. The autonomy and interoperability of document is violated - it can not exist out of dedicated system.

Another way to implement such transactions is the use of signed PDF documents and PDF forms. The use of such documents is a source of serious problems and costs associated with processing its.

In turn, the use of e-mail (SMTP) does not provide legal effectiveness and also creates serious problems with messages processing.

The proposed concept of electronic documents in the shape of forms using the XML format, containing both a description of visualization and data input, signed with electronic signatures and containing other elements (calculations, verification rules, defaults, etc.).

The concept of electronic form presented in this paper is a subject of work in progress at the stage of wide testing. Results of work will be presented in future papers.

References

1. Lucie Veseláa, Miroslav Radiměškýa, The development of electronic document exchange. *Procedia Economics and Finance* 12 (2014) pp. 743 – 751.
2. Guidelines on best practices for using electronic information. How to deal with machine-readable data and electronic documents, Brussels 1997
3. Radoslaw Petermann, Electronic document, ArchNet, Lodz 1998
4. Yu-Chi Chen, Gwoboa Horng ↑ , Chang-Chin Huang, Privacy protection in on-line shopping for electronic documents. *Information Sciences* 277 (2014) 321–326
5. R. Sion, B. Carbunar, On the computational practicality of private information retrieval, in: *Proc. of Network and Distributed System Security Symposium, 2007*
6. K. Sakurai, Y. Yamane, Blind decoding, blind undeniable signatures, and their applications to privacy protection, in: R. Anderson (Ed.), *Proc. 1st International Workshop on Information Hiding, Lecture Notes in Computer Science*, vol. 1174, Springer-Verlag, 1996, pp. 257–264.
7. Extensible Form Description Language (XFDL) 4.0, W3c Note, September 1998, <https://www.w3.org/TR/1998/NOTE-XFDL-19980902>
8. Digitizing Public Services in Europe: Putting ambition into action 9th Benchmark Measurement, December 2010, Capgemini, IDC, Rand Europe, Sogeti and DTi for: European Commission, Directorate General for Information Society and Media, December 2010
9. Briet, S. (1951). *Qu'est-ce que la documentation?* Paris: Documentaires Industrielles et Techniques
10. International council on archives ICA.ORG, www.ica.org
11. Directive of the European Parliament and Council 1999/93 / EC of 13 December 1999. On a Community framework for electronic signatures
12. Act of 18 September 2001. On electronic signature. OJ 2001 No. 130, item. 1450
13. XML Signature Syntax and Processing, <http://www.w3.org/TR/xmlsig-core/>
14. XML Advanced Electronic Signatures (XAdES), <http://www.w3.org/TR/XAdES>
15. Extensible Markup Language (XML), <http://www.w3.org/XML/>
16. Blake Dournaee XML Security, McGraw-Hill Osborne Media, 2002
17. V. Hassler, *Security Fundamentals for E-Commerce*, Artech House, 2001
18. Namespaces in XML, <http://www.w3.org/TR/REC-xml-names/>