XML-Signatures and the Presentation Problem

Fraunhofer Institut
Sichere Telekooperation

XML-Signatures and the Presentation Problem

Talk at the University of Natal, Durban, ZA

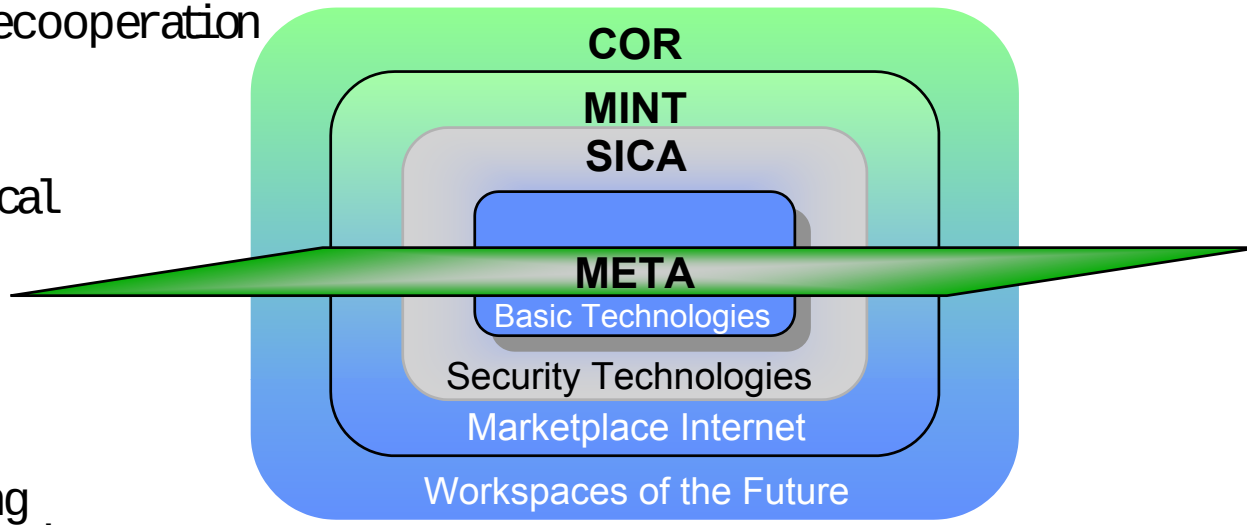
16 May 2003

Thomas Kunz, Ulrich Pordesch, **Andreas U. Schmidt**

Fraunhofer Institut
Sichere Telekooperation

Fraunhofer Institute Secure Telecooperation (SIT)

- Development of security technologies
- Incorporation of security technologies in pre-existing applications
- Realisation of innovative forms of telecooperation
- Activities based on XML, e.g.
  - TMF - Telematikplatform for medical research networks: Medical data in XML
  - Media@com (e-government): X.509 authentication protocol
  - Diploma- and PhD theses about XML-Encryption, secure auctioning systems using XML-Dsig, presentation problem, *etc.*



Fraunhofer Institut Sichere Telekooperation

Our collaboration:

Thomas Kunz

- Studies in informatics Uni Frankfurt/Main
- Since 2001 research assistant at SIT.MINT
- Themes: secuity in e-business processes, security-solicies

Ulrich Pordesch

- Studies in informatics TU-Darmstadt
- Since 1997 PhD student/researcher at SIT
- Schwerpunkt: judicial requirement analysis
- PhD Dissertation (TU-Ilmenau): Die elektronische Form und das Präsentationsproblem (the electronic form and the presentation problem)

Dr. Andreas U. Schmidt

- Studies in mathematics and physics, University Frankfurt/Main, 1999 PhD in Mathematics
- 1999/2000 researcher at the GMD Institute SIT, AG MINT. Subject: digital signatures in XML
- 2000/01/02 research stays in ZA (UDW) and Italy (Univ. Pisa)
- since october 2002 researcher at SIT.MINT. subject: security-policies

Overview

1. Common Content- and Signature Formats
2. The Presentation Problem
3. Pros and Cons of XML as Content- and Signature Format
4. Remaining Problems and Possible Solutions

Usual Content Formats

Word processors: Word

Spreadsheets: Excel

E-Mail: Mime-ASCII

Internet: HTML

Financial data exchange: EDIFACT

Archiving: PDF

Fax/Scan: TIFF

Pictures: JPEG, GIF

„Security properties“

- No secured authenticity/integrity
- Content and authorship repudiable
- No judicial proofs

Therefore even today in many use cases

- Print
- Sign
- Scan, and/or archive as paper …

Common Signature Formats (I): CMS

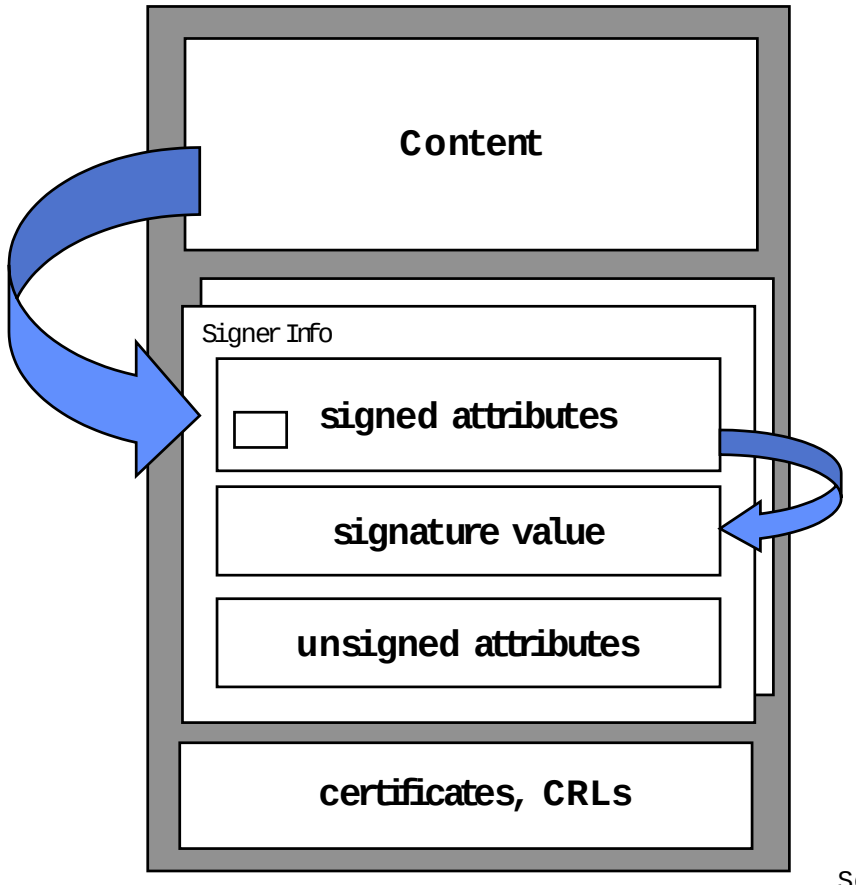
ASN.1-Syntax, binary coding

Signature generation

- Hash content (e.g. file)
- Hash the hash value and further attributes (time, used algorithms, content formta)
- Generate signature value
- Add certificates and further attributes (endorsing signature, time stamp,..)....

Form CMS-Container: Content integrated or external

Numerous ASN.1-based standards for attributes, certificates, CRLs, ...



Fraunhofer Institut Sichere Telekooperation

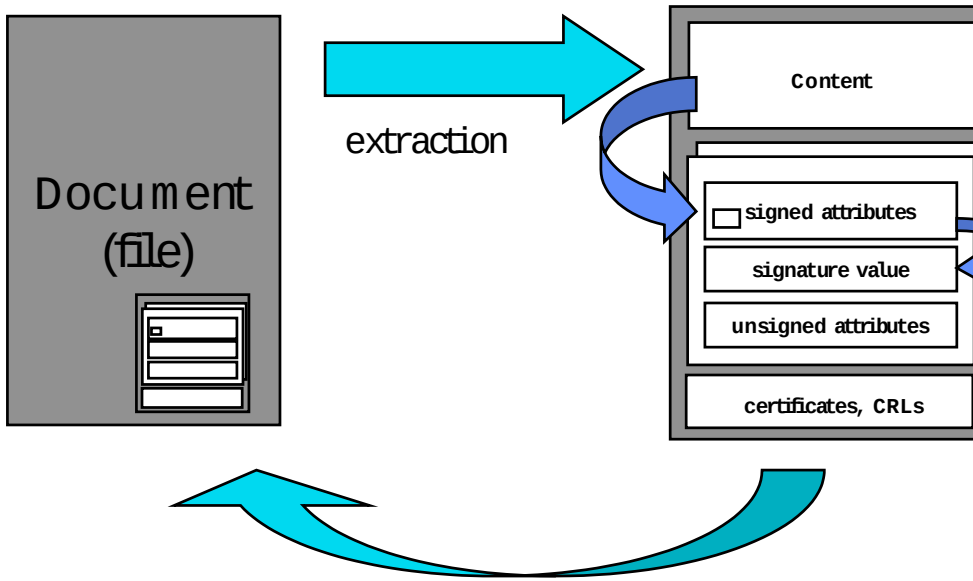
Common Signature Formates (II): Usage

a) Signed message within a file in CMS format

b) Document remains unchanged, signature placed in additional file in CMS format

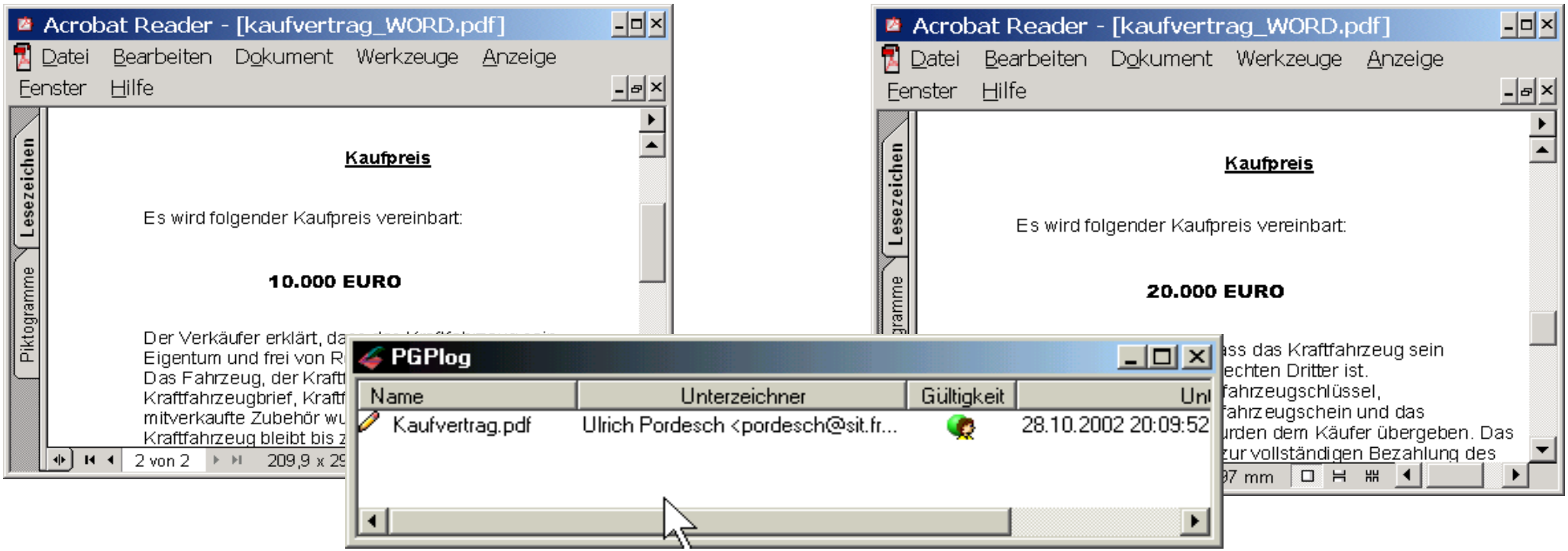
c) Integration in usual document formats

- Selection, transformation coding of document data by application yields signable content
- Signature is placed into document file, document format (syntax) is extended by a CMS-block for that
- Example: PDF



Fraunhofer SIT Institut Sichere Telekooperation

The Presentation Problem (I)



Two presentations of the very same signed document contents differ significantly and lead to (legally) incommensurable interpretations.

Fraunhofer SIT Institut Sichere Telekooperation

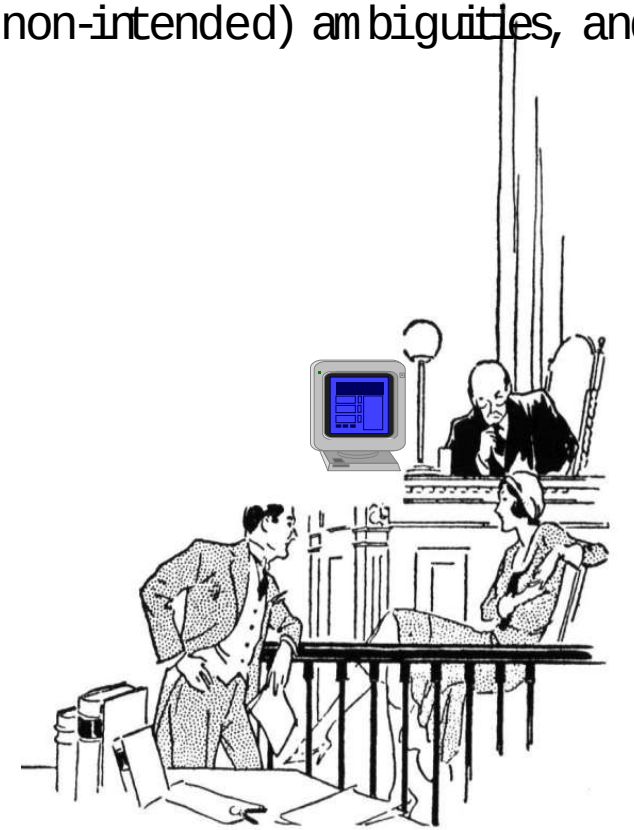
The Presentation Problem (II)

A variety of possible presentations results from (intended and non-intended) ambiguities, and misrepresentation due to error and/or abuse.

Unique presciptions are lacking for

- Data format and syntax
- Presentation of content
- User interface
- User-system-interaction

In conclusion: Judicial value of signed data might be limited, although it contains a technically correct digital signature



Fraunhofer SIT Institut Sichere Telekooperation

Presentation Problem with Common Formats

Content format

- Content itself is mostly non-unique concerning representation; but worse:
- **What** is signed is often not recognizable:
  - Origin, selection, transformation, coding is opaque
  - Content is not human-readable, syntax and pertinent semantics are not disclosed
- With CMS-signatures: Only one file can be signed, but not, for instance, stylesheets, txtual descriptions, presentational variants, etc.

Signature format

- Content format *can* be defined (as a MIME-type), but is often not used properly („data“)
- Differing content formats can be used in parallel signatures
- Signature data itself is implementation dependent

Fraunhofer SIT Institut Sichere Telekooperation

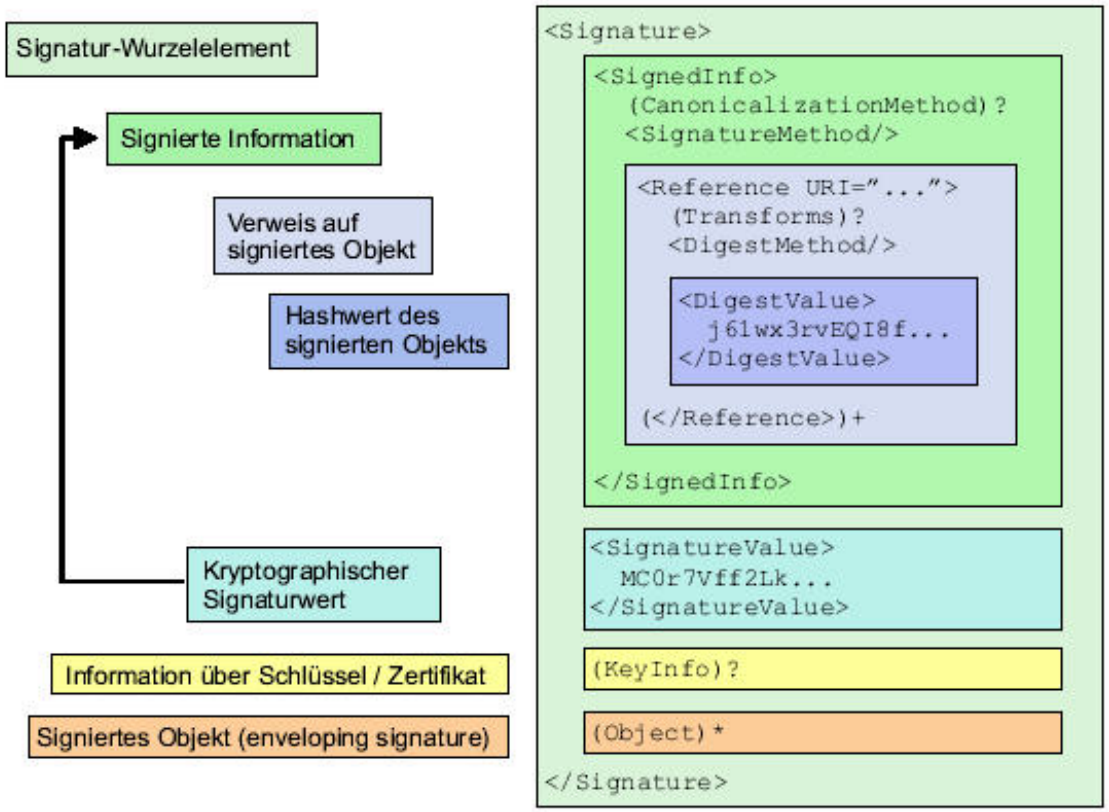
XML as Signature- and Content-Format

The common standard of IETF and W3C for XML-Signatures
(XMLDsig, W3C Recommendation 12.4.2002) provides a format for signatures of:

- Structured (XML) and unstructured data in single or multiple files
- from different application contexts (interoperability),
- in a multitude of variants (enveloped, enveloping, detached, selected content, ...)
- Web-wide distributed or local,
- mobile or with fixed location,
- with low demands on implementations,
- using established as well as advanced cryptographc standards,
- with open-source implementations

Fraunhofer Institut
Sichere Telekooperation

Basic Structure of XML-Signatures (XMLDSig)



Fraunhofer SIT Institut Sichere Telekooperation

XML Components Associated to the Presentation Problem

Syntax: schema

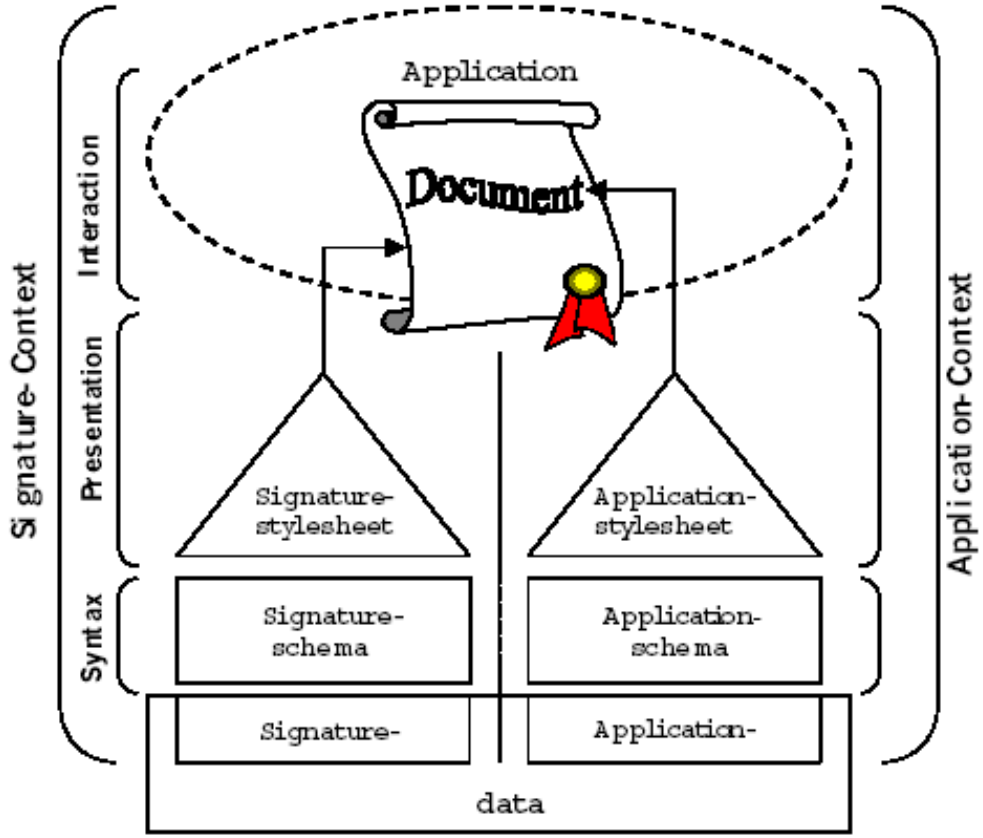
- Old: DTD – SGML's `heritage´
- XML-Schema
  Syntax descriptions for XML in XML

Presentation: Stylesheets

- CSS - for presentation of (X)HTML in the Web browser
- XSL Extensible Stylesheet Language:

  XSL Transformations
  (XML to XML, text, or (X)HTML)

  XSL Formatting Language: detailed presentation prescriptions for XML-Documents



Fraunhofer SIT Institut Sichere Telekooperation

Advantages of XML Pertaining to the Presentation Problem

- **Separation of signature- and application context:**

  Syntax and presentation can be determined sepearately for Application data and signature data

  These components can be made attributable – by (XML)-signatures – to responsible parties
- Transparent codierung (human readable [?] XML)
- Adapted presentation variants possible by association to stylesheets
- Uniqueness of syntax through use of namespaces (fixing semantic context of the XML-elements)

Fraunhofer SIT Institut Sichere Telekooperation

Further Advantages of XML-DSig

- Signing of multiple data objects enables authentication of associated stylesheets and schema definitions
- Transformations (XSLT): data objects can be changed before signing them: inessential and potentially damaging parts can be eliminated (e.g. script code)
- Transparent selection of content parts before signing via Xpath (powerful)
- Transparent normalisation

Fraunhofer Institut
Sichere Telekooperation

XML-Signatures: Flexibility *vs.* Security

Flexibility

- Explicit re-coding, execution of externally defined Operations before signing
- Content-selection with XPath/XPath Filters: useful for forms and workflow applications where parts of documents are to be signed by various parties
- Canonicalization: re-codings that keep the signature invariant – should also leave the semantics of the content invariant

(In)security

- Does the application of c14n, transformations, selction operations really leave the semantics of the signed content unaffected? The signed doc might have a meaning which differs from the original one.
- C14n and XPath even create problems on the syntactic level, i.e. for interoperability: Even for simple use-cases, XPath expressions become unintelligible and error-prone.
  (Standard was developed with emphasis on lightweight implementation and high functional power)
- On higher semantic levels: What is it, really that is signed – the XMLDSig standard knows not of content types. This meaning has to be provided within the application context.

Fraunhofer Institut Sichere Telekooperation

Requirements w.r.t. Presentation Problem

- **Signed Stylesheets:** The authenticity of the stylesheet used for the presentation must be ensured by digitally signing it
- **Binding of Context:** Syntax (Schema) and presentation must be **bound** to the signed XML data, i.e. their usage must be explicitly prescribed *and* signed.
  XMLDSig has no high-level semantics for that.
- Otherwise: Integrity of context components becomes a problem
  (is the singed stylesheet the same that has been used for presentation?)
- **Ambiguity:** To which level of detail shall context components be presented, if they are signed? E.g. Should also signed stylesheets be presented?
- **Presentation of Signature:**
  - Should there be a unique, application independent (standardized, „officially regulated“) presentation of signature, certificates, timestamps, etc. ?
- **Interaction level is underspecified:** Even advanced XML form description languages do not specify user interaction; it is usually implemented by scripting languages.

Fraunhofer Institut Sichere Telekooperation

XAdES (ETSI – Specification)

Extension of W3C-specification (success doubtful) by signature attributes

- ASN.1-time stamps, -certificates, - CRLs
- Counter signatures

Assignment of data formats (signed)

- Text descriptions of data
- Format specification via Object Identifier (OID) or MIME-Type with encoding

```
<ds:Object>
    <QualfyingProperties>
      <SignedProperties>
        <SignedSignatureProperties>
          (SigningTime)
          (SigningCertificate)
          (SignaturePolicyIdentifier)
          (SignatureProductionPlace)?
          (SignerRole)?
        </SignedSignatureProperties>
        <SignedDataObjectProperties>
          (DataObjectFormat)*
          (CommitmentTypeIndication)*
          (AllDataObjectsTimeStamp)*
          (IndividualDataObjectsTimeStamp)*
        </SignedDataObjectProperties>
      </SignedProperties>

    </QualifyingProperties>
</ds:Object>
```


Fraunhofer SIT Institut Sichere Telekooperation

Further Directions

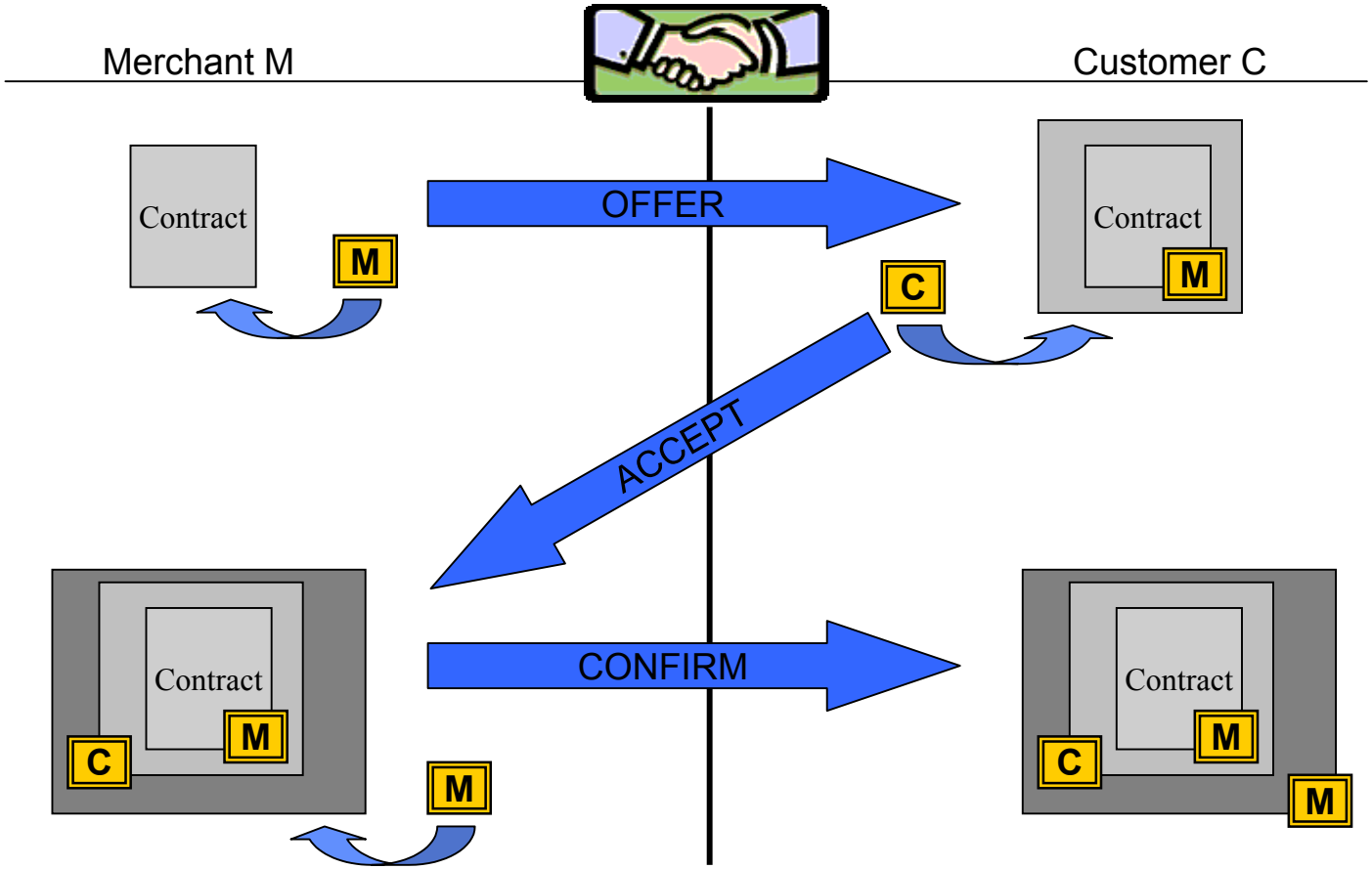
Syntax and presentation of signed data

- Depends too much on applications for central prescriptions.
  Should there be guidelines?
- Evaluation, registration, and certification (by digital signatures) of stylesheets and schemas for applications
- Presentation of signature (nontrivial due to the large variety of signature types):

  „Officially prescribed" stylesheets for signature presentation could form a base infrastructure for XML signature applications.
- Binding of context elements: A meta standard based on XMLDSig could provide pertinent semantics and functionality.
- Expressive elements for user interaction...

Fraunhofer Institut
Sichere Telekooperation

An Example: BaKo – A Protocol for Mutual Non-Repudiation



Fraunhofer SIT Institut Sichere Telekooperation

Presentation of Signatures: Doubly Signed ACCEPT Document

Fraunhofer Institut Sichere Telekooperation

Presentation of Signature: Detailed

XML-Signature (detaillierte Darstellung)	
Signature Information	
Canonicalization Algorithm	http://www.w3.org/TR/2000/CR-xml-c14n-20001026
Signature Algorithm	http://www.w3.org/2000/09/xmldsig#rsa-sha1
Reference URI: #ts245908	
Transformation:	http://www.w3.org/TR/2000/CR-xml-c14n-20001026
Digest Algorithm:	http://www.w3.org/2000/09/xmldsig#sha1
Digest Value:	[illegible]
Reference URI: #245908	
Transformation:	http://www.w3.org/TR/2000/CR-xml-c14n-20001026
Digest Algorithm:	http://www.w3.org/2000/09/xmldsig#sha1
Digest Value:	[illegible]
Signature Value:	
[illegible]	
X509 Certificate Information	
Issuer Name:	CN=Certifier, O=www.certifier.com, L=Frankfurt, C=DE
Serial Number:	1
Signer (Subject Name):	CN=Merchant, O=www.merchant.com, L=Darmstadt, C=DE
Object ID: ts245908	

Angebot ID: [illegible]	
Datum:	03.12.2000
Uhrzeit:	17:32:33

Object ID: 245908

Fraunhofer Institut Sichere Telekooperation