Business requirements for authentication and integrity of XBRL documents

Prepared by Eric E. Cohen (eric.e.cohen@us.pwc.com) on behalf of the

American Institute of Certified Public Accountants' (AICPA) Assurance Services Executive Committee (ASEC)

Editor's Draft[1]/Discussion Draft, dated November 7, 2007

Disclaimer: This document is a discussion draft only, has not been reviewed by members of the AICPA ASEC, and does not necessarily represent the views of that committee or my employer. This version was prepared for discussions with the XML Security Specifications Maintenance Working Group[2] of the W3C.

Table of Contents

Summary

Electronic documents are increasingly replacing paper reports and forms in compliance processes as the basis for the commercial use and government filing of information. Why? Because the Internet helps filers and users realize cost, timing and accuracy efficiencies, and permits information to be reused by the market in a way never before feasible.

These benefits may also amplify risks, as traditional controls relating to the authentication and integrity of paper are abandoned for the sake of efficiency.

[1] As such, it may and does contain unfinished text in the process of being prepared.
[2] http://www.w3.org/2007/xmlsec/

The audit profession has recognized that this creates concerns and opportunities as it electronic presentation-based documents (information represented by HTML or PDF) are replaced with XML-based XBRL[3] (Extensible Business Reporting Language) documents. XBRL is an XML-based language for communicating financial statements[4], tax returns[5] and other business reports, forms and data.

The purpose of this document is to discuss business requirements for authentication and integrity of XBRL documents. While the concentration is on information exchanged using XBRL specifically, solutions that meet these business requirements are likely to be useful for other types of digital information in fields such as medicine (e.g., medical records), law (e.g., contracts) and banking.

The increasingly sophisticated electronic information supply chain provides new ways for organizations to mistakenly expose information to the public, as well as opportunities for parties to purposefully deceive. The audit profession is not *responsible* to mitigate all of these dangers, but hopes to take part in a market-driven, collaborative effort to reduce its/our risks, improve the market, and help all who might benefit from improved business reporting.

How can we innovate for the Web? How can we make the business reporting supply chain on the Web *as* trustworthy - if not more trustworthy - than the paper-paradigm environment of today? How can we design trust on business reporting so it crosses borders, languages and cultures?

[3] http://www.xbrl.org

[4] See http://www.sec.gov/spotlight/xbrl.htm for more on the U.S. Security and Exchange Comission's move toward XBRL filings; there are parallel programs going on in Canada (http://www.csa-acvm.ca/html_CSA/xbrl.html) and around the world

[5] In the UK, HMRC has mandated XBRL for filings by 2011, for example.

Examples/Use Cases of Authentic XBRL

To help the reader understand the issues presented in this document, we provide a series of use cases. The phrase "authentic XBRL" is not meant to differentiate something that is "valid XBRL" from something that is part XBRL and part something else; it means the XBRL documents are of verifiable origin, are genuine, are the assertions of management with no loss of data integrity in the representation; they are genuine, if not perfect. Other attributes are discussed later in this document.

A situation is described, followed by how the situation is dealt with today, and then how the situation might be improved in the "*authentic XBRL environment*" - the environment where solutions for authenticity, integrity and other technical, business and regulatory issues have been developed and implemented.

Protecting the Protectors

Whether in the digital XBRL world or even in today's "analog" HTML and PDF environment, expressed responsibility and assurance related to business reports is important to communicate, but difficult to verify. When, for example, there is an auditor's report on a business reporting document, a textual representation of the auditor's report including a textual representation ("/s/PricewaterhouseCoopers") or a graphical presentation is provided today. An investor wishes to verify that the assurance was indeed provided by that auditor. Likewise, the audit firm wishes to efficiently monitor the Internet for the misuse of its name.

In the past, the *investor* would have few techniques to gain comfort that the auditor's opinion and signature were authentic and unchanged. They might trust information submitted to the SEC, so rely on the document on EDGAR as a baseline. They might see other verification that the named auditor was indeed the auditor of record for the time of the report. The *auditor* would regularly search the Internet for the use of their name, researching obvious anomalies. Graphical images or information deep within databases not accessible by search engines pose a greater problem.

In the new authentic XBRL environment, investors will be able to verify the validity of the auditor's report, easily determining whether it was signed by an appropriate agent of the appropriate audit firm and quickly receiving feedback that the opinion has remained unchanged since it was signed.

Auditors will also benefit from this environment; users can more easily provide feedback of an auditor's name being associated with assurance they did not, in fact provide. Auditors will also have greater control of the appropriate use of their signature, with the ability to revoke or draw other attention to situations with which their name and opinion letter is associated.

Protecting the Protectors - applied

Filings with the SEC Voluntary Filing Program may include:

- Contextual tags that indicate whether specific facts are audited or unaudited. These tags are not standard to XBRL today, but there are currently numerous ways to express this information.
- An XBRL-tagged representation of the original auditor's report (or other service provided) on the original paper-based report which is being represented with XBRL (which I will here call the "*audit report*".)
- A separate HTML file with the attestation (or other service provided) report on the XBRL-related documents (which I will here call the "*attestation report.*"
- And, possibly in the future, an XBRL representation of the auditor's report on the XBRL file.

Today, the only reason that the market knows that the audit report is actually associated with the named auditor, or that the attestation report was actually provided by the named attester, is that someone who forged an actual auditor's identity could be punished by the SEC or others. In the new authentic XBRL environment, such documents can be digitally signed by the auditors and attesters with a means of evaluating those signatures for validity and appropriateness .

Signing Signals Special Services

An investor is looking at an annual report online. There is a lot to read!

What happens today? Balance Sheets, Income Statements, supplementary schedules, Notes, MD&A, links off-site; where is the auditor's report? I think this paragraph discusses what is covered by the assurance and what isn't, but as I click around on hyperlinks sometimes I go to sections I know are part of the annual report, sometimes I go to sections I am not sure are part of the annual report. How do I know what is covered and what isn't?

In the new authentic XBRL environment, there is not only a trustworthy auditor's report that can be verified for integrity and for feedback on the auditor, but the links from the assurance to the report itself are explicit. The user can have their application quickly display what content is covered by the assurance through colored highlights, graying out non-assured data or other means suitable to those relying on the assurance. Should there be multiple types of assurance provided, multiple parties indicating that they have authorized or provided assurance, or other complexities, the applications can help differentiate between those, and enable the reader to easily switch between reported facts and the relevant certification or assurance on those facts.

The Case of the Sticky Assurance

An investor sees a press release with the quarterly sales results from a company in which they have considered investing. The numbers look promising; are they too good to be true? Are they what the company reported?

In the past, the investor would choose some more "authoritative" sources to verify the number. They would go to the EDGAR system and search for a filing from the company. Alternatively, they would try to find the investor relations section of the corporate web site to also look for the filing.

In the new authentic XBRL environment, the investor clicks on the number to see if there are any assurance attributes that have been provided. Indeed, the fact does have assurance attributes (as well as links to some analyst reports). The investor clicks to see that the number was drawn from the 10Q filed with the SEC (with the ability to drill to the text on the EDGAR site or to load the XBRL version of the 10Q into their I-Metrix or Xinba spreadsheet). They can also see instantly from the signatures on the facts that the facts were signed by management and the external auditor, the level of assurance the auditor provided on the XBRL representation of the facts (either individually or pointing to the original document from which the individual fact was drawn) and also see the type of assurance that was provided on the source document represented by the XBRL document.

I Feel the Need - The Need for Speed

Today, the closest the US market comes to near real-time data is postings to the EDGAR site. While information from EDGAR is available from direct access, through RSS feeds[6], and through access to commercial resellers of the data, the possibilities of more data, delivered more often, and closer to the time of the triggers for the information are growing. When data "jumps" into models without human intervention, the need to have trust of the source and the integrity of the information increases. For the move to real-time reporting[7], automated trust is foundational.

Under current circumstances, having XBRL information reduces the mechanical tasks of entering information into models, but doesn't reduce the need to trust the source or check what has been imported against some original source on a trusted site, such as the SEC's EDGAR system.

In the new authentic XBRL environment, an analyst will be feeding a wide variety of XBRL feeds into their modeling systems on "as real-time a basis as possible". For information from certain known organizations, or with assurance provided by certain auditing firms, the data will flow into their systems immediately and be fully considered in their analysis. For those with the companies or auditors in another "tier", the information will be included, but discounted or not given the same weight. Information from non-trusted sources will be ignored, highly discounted, or subject to manual follow up. Organizations like the AICPA, CFA and FEI may provide subscriptions to databases of appropriate organizations, roles and persons to keep such listings internally up-to-date.

[6] http://www.sec.gov/Archives/edgar/xbrlrss.xml

[7] Section 409 of the Sarbanes-Oxley Act of 2002 is entitled *Real-Time Issuer Disclosures* and Congress and the SEC have both expressed a desire to move to real time reporting.

About this document

In light of the consideration to develop a schedule to make the filing of XBRL documents with the SEC mandatory (as well as recognizing that other compliance documents, such as tax returns, have already begun that evolution), a set of business requirements is being developed to help bring together the needs of the profession and communicate them to those who can provide solutions.

What is this document, and who is it for?

This document is a draft of a collection of business requirements, primarily concerning the authentication and integrity requirements related to XBRL related documents (taxonomies and instances) and secondarily to other compliance oriented documents. It is accompanied by explanatory information helpful to those who will be engaged in moving the satisfaction of these requirements forward, including:

- Participants in the compliance reporting supply chain
- Experts in XML and security standards and technology
- Vendors interested in providing solutions in this space

The organization is roughly organized around the following issues:

1. We have already provided a few stories that describe some of the challenges and opportunities. From there, we move to:
2. Why are these business requirements being developed?
3. How are the provided requirements organized?
4. What are the requirements?
5. What are the next steps?
6. What underlying philosophies or believes undergird the next steps?
7. What technologies are obviously relevant?
8. Examples/Use cases of reliable XBRL
9. A brief technical overview of XBRL technology issues
10. Examples of risks and problems

Why are these business requirements being developed?

The growth of XBRL has highlighted the need for secure documents

XBRL (Extensible Business Reporting Language) is an international agreement on the use of W3C Recommendations related to the XML[8] (Extensible Markup Language) family to simplify and standardize business reporting information. XBRL is best known for its use to digitize financial statements, so the information is more easily collected, published, discovered and consumed. Unlike other electronic, but analog, formats like HTML, XBRL uses XML tags to associate contextual information with reported facts, so the individual facts can be identified consistently and reused[9].

[8] http://www.w3.org/XML/

[9] Where XML describes, XBRL formally defines the facts, their attributes and interrelationships.

XBRL is also used to provide a standardized, generic and holistic format for moving operational, business and accounting information to and from ERP systems in detail[10]. XBRL is being adopted by governments around the world[11] to improve their compliance processes and reduce compliance burden, replacing tax returns filed on paper forms or their presentation-centric electronic counterparts, HTML and traditional PDF. XBRL is also being adopted by corporations and other market participants to communicate their own messages more precisely and exchange business reporting information of all kinds - statistical, sustainability, statutory and other information financial and non-financial - more efficiently.

When these reports, forms and other filings were prepared “on paper,” parties who wished to communicate their authority or who needed to otherwise assume responsibility for those documents would pick up a pen and sign the document. That signature, sometimes accompanied by a notary’s stamp, would provide comfort for a later user that the document was legitimate, its publication was authorized, and its content trustworthy.

For the average person in the United States, taking pen to paper and writing a signature is still a regular experience in a few areas. We sign our signature to credit card receipts, for example. In other areas, that act, once common, is going away. An example is the move from paper mail to e-mail; our “John Hancock”, written with a flourish at the end of a missive has decreased markedly. Just as signatures on credit card receipts are starting to go away (with the “chip and pin” credit card required after February 14, 2006, UK consumers began to experience “No PIN, No Plastic”), the usefulness of signatures on electronic filings for anything but basic legal requirements is dissipating.

In an electronic business reporting world, there is no physical paper on which to sign, and no literal pen with which to sign the document, form, or auditor’s opinion; signatures are electronically noted (whether a graphical representation of a paper-based signing, or some other indication). This was true before XBRL, as HTML and PDF “analog” versions of paper reports began supplanting paper. Business reporting supply chains that rely on this information realized that some solution was necessary.

Many of the attributes associated with current paper processes (e.g. authentication, authorization, integrity) must be translated to electronic processes. Relying on the system to provide trust - as opposed to providing trust directly at the document level - has seen numerous failures, as authorized web sites have been exploited and even secure browsing has led users to the wrong place and the wrong information.

Today's tools are unsuitable

Falsifying auditor's reports and associating auditor's improperly with reporting is not new to the world of the Internet, and especially not to XBRL. In 2006, for example, the Inter-American Development Bank noted that a company in a competitive bidding situation attempted to bounce back from being

[10] This is the XBRL Global Ledger Framework, an agreed-upon semantic of the data fields in typical ERP systems, expressed using XBRL. Learn more at http://www.xbrl.org/GLTaxonomy

[11] Such as the Netherlands Taxonomy Project' learn more at http://www.xbrl-ntp.nl/english

disqualified by their financial situation in a fraudulent way - changing their financials and also falsifying the auditor's report[12].

One “solution” that is in place right now for financials on the Internet, which tries to solve today’s problems with yesterday’s tools, is to consider a set of traditional signatures (made by pen on paper) that are “on file”, accessible only on an exceptional basis, while letting the reader see an electronic image indicating that those signatures exist (e.g., “/s/PricewaterhouseCoopers”). This does not solve the authentication, authorization, and integrity issues that readers face, some of which were pervasive even in the paper-paradigm environment. Anyone can reproduce a physical signature while the consumer’s cost to verify it is so great that it is virtually impossible to do so (authenticity).

Likewise, there is no way to verify that the document has been unaltered (integrity), or that the supposed signer is actually authorized to sign on behalf of the auditor or management (authorization). The use of physical signatures provides these attributes for paper documents, to a degree; however, electronic documents can be more easily copied, altered and falsified and a physical signature does not provide these same characteristics for electronic documents. Some means of being able to digitally check the reliability of documents is necessary.

An example of today's solutions falling short

Companies House in the UK[13] has been the victim of fraudsters, finding ways around their controls (which do not include digital signatures), with an estimated loss of over $100 million (£ 50 million) annually[14]. Today's solution is to recommend that companies monitor the Companies House system to see if someone is entering unauthorized filings on their behalf. Despite Companies' House controls and advice, my firm, PricewaterhouseCoopers and others (including KPMG) have found our names misused on filings. Court orders have to be obtained before the filings can be removed; Exchanges may need to suspend dealings in shares when things go wrong. Digital signatures have been raised as an obvious tool to help battle the problems.[15]

The (intuitively obvious) solution shouldn't precede the problem statement

Accordingly, the use of ‘digital signatures’ to ensure the authentication, authorization and integrity of electronic documents as they are processed appears to be the most efficient way to provide these attributes across the wide range of preparers, auditors, investors, analysts, regulators and other business reporting supply chain participants. A digital signature solution can help ensure that these documents can have the authenticity and integrity of their paper parallel - and in many cases, *more* trustworthy and confirmable

[12] http://www.iadb.org/integrity/oii_ar06/inv_external.cfm?language-english&detail=box6
[13] This statement is not meant to be a criticism of Companies House or their personnel.
[14] According to Accountancy Age.
[15] http://www.accountancyage.com/accountancyage/analysis/2183448/digital-signatures-protect

authenticity and integrity - the widespread use of digital signatures (and I would propose, in particular, XML Signature[16]) has yet to take hold.

XML Signature comes into play as "document level" signing is not the ultimate solution. More fine-grained signing is necessary, and XML Signature provides the granularity necessary for the new electronic world.

However, as this document is meant to communicate business requirements, my goal is not to declare *the* solution, as different solutions may be necessary to meet current priorities, based on the available technical infrastructure and in light of the business and regulatory environments, and to solve the problem more holistically in the future, as well as meeting the needs not just of XBRL for financial reporting in the United States, but of many information exchange requirements, for many purposes, on a global and interoperable basis.

For this process to be ultimately efficient, the market needs to have a limited number of digital signature solutions. We need such a solution for compliance reporting processes that meets specific business requirements across the full range of participants.

How are the business requirements organized?

The business requirements are organized by process areas.

A draft of the process areas would include:
1000) Setting up/prearrangements
2000) Signing
3000) Storing
4000) Communicating
5000) Retrieving
6000) Discovering, checking and viewing
7000) Following up
8000) Acting after-the-fact (including revoking signatures, obsolete documents, etc.)

The requirements are also characterized with additional information, including:
1) Priority/time line (urgency, importance short/medium/"ultimate")
2) Importance
3) How they relate to the primary *technical* issues:

Authorization	Was the publication of this information authorized by appropriate people?
Authentication	Is the information here the actual assertions made by the publishers?
Accessibility	If the information is available, can it reliably be retrieved?
Integrity	Has the document been changed in a way that alters the information in it since it was initially published?
Obsolescence	Has the information been updated, changed or otherwise become obsolete?
Time stamping	Was any signature appropriate at the

[16] http://www.w3.org/Signature/

	time, even if the signer has since changed roles or left the organization?
Non-repudiation	Non-repudiation is especially important in the government filing arena.
(Others?)	

The scope of this document does not include confidentiality. Confidentiality, which is usually associated with ***encryption****, is vitally important to securing and exchanging non-public information, such as that normally associated with XBRL's Global Ledger Framework (XBRL GL). Encryption requirements will be described and defined in a separate document.*

4) How they related to the primary business issues:

Services	Greater visibility/transparency/communication of services provided or *not provided*
Users	Insulating non-technical participants (auditors, investors) from the technical aspects of XBRL and security
Management	Being able to manage electronic documents as or more easily than their analog counterparts
Differentiation	Clearly delineating the differences between the owners of assertions (i.e., management) and assurors of the assertions (i.e., the auditor), as well as other parties
(Others?)	

What are the requirements?

This table represents a broad range of requirements for working with XBRL information. The use of the word "sign" is not meant to be limited to "digital signature", but to express the manual task of "signing" something today; the electronic equivalent could potentially be any number of technologies and metaphors (e.g., publishing a hash on one's own web site as a way of expressing the ownership and/or responsibility).

Also not represented in the table are some general principles:

- Recommendations we make will be for interoperable, global solutions. While implementations may have local requirements, the ultimate goal is a single, harmonized solution.
- Although our concentration is primarily on compliance documents exchanged in the business reporting supply chain, the solution we foresee is one that will be equally relevant to exchanging medical information (HIPAA), banking and mortgage documents, or legal documents.

[Note: this is a first draft; the priority/timing and assignment to areas are not "random" but have not yet been the subject of much thought or any collaboration. They are a starting point for discussion.]

Description	Priority/timing	Areas
1000) Setting up/prearrangements		

Description	Priority/timing	Areas
1010. Establish workflow and roles • Need to differentiate between types of participants and roles of the participants • Types of participants: Company, preparer, auditor, notary, analyst, regulator, others • Roles or participants: Those authorized to publish information on behalf of the organization, management, employees with accountability to management • Need to be able to register and communicate authorized users	High/mid-term	Authorization Differentiation

Description	Priority/timing	Areas
1020. Establish standardized "reasons" for signing • Authorization of publication • Responsibility for publication • Types of assurance/services/opinions • We saw this; it has not been altered since we saw it (but we make no other claims) • We provide assurance for this; it is trustworthy • Please read inside; you will see our assertion, which we here sign with the materials with which it is related • We are “associated” with this document	High/mid-term	Authorization Differentiation
1030. Setup isn't just for the signer; consumers will potentially have databases of trusted sources/signers for automated acceptance and trust.		
2000) Signing		
2010. Permit multiple signatures with serial/parallel workflow	High/short-term	Authorization Authentication

Description	Priority/timing	Areas
2020. Capture/Questions about the reason for the signing	High/mid-term	Authorization Authentication
2030. Allow signing from different organizations, different certifiers; what if different signers use different digital signature syntaxes/algorithms, different certificate authorities? (Vendor and platform independence)	High/short-term	Authorization Authentication Integrity
2040. Provide standardized user interface allowing visual accessibility for signers who are not technical experts, including considerations for sensory disabled/impaired persons	High/mid-term	Users
2050. Sign multiple documents simultaneously	High/short-term	Integrity
2060. Sign specific areas of documents	High/short-term	Authentication Services
2070. Sign different areas of documents with different reasons for signing	High/mid-term	Authentication Services
2080. XBRL files, when signed, still (optionally) remain valid XBRL. (XBRL practice is that instances are payload and that any use would be after "unwinding" any transport, handling security.)	Medium/mid-term	Integrity

Description	Priority/timing	Areas
2090. Ability to sign and bind together instances, original and extension taxonomies, optional additional files (e.g., assurance HTML, stylesheet)	High/mid-term	Integrity Users Management
3000) Storing		
3010. Allow identification of "open" signed documents.	Medium/mid-term	Management
3020. Signature bound with documents, even if signature is stored in a separate file.	High/short-term	Management
4000) Communicating		
4010. Allow "subscription" by topic to signed documents, including updates to previously published documents	Medium/long-term	Users Management
4020. Signatures usable in Web Services environment.	High/short-term	Integrity
5000) Retrieving		
5010. Allow querying of authorized signers by role within organizations for non-anonymous consumers with limitations by role	Medium/long-term	Management
5020. Allow role-based, rule-based retrieval of information	Medium/mid-term	Management

Description	Priority/timing	Areas
6000) Discovering, checking and viewing		
6010. Questions about the Signature: • Whose signature is this? • Is the person who signed it still with the firm? Does that matter? • What does it apply to?	High/short-term	Authentication Authorization Obsolescence
6020. Questions about the Document: • Is this document – The real document • The original • The one I asked for • Without change – The most recent version of the document • If superseded, we would know about it – Actually from whom we think it is	High/short-term (except for obsolescence, which is mid to long term)	Authentication Integrity Obsolescence

Description	Priority/timing	Areas
6030. Questions about the Signer: • Authorization – Was this published by the appropriate representative of management/officer – Was it signed by the appropriate firm and an appropriate representative of the auditor	High/mid-term	Authorization
6040. Public accessibility - technology should, in most cases, not be limited to a proprietary technology or be limited to a specific "network" (only available within a corporate/through a VPN)	High/mid-term	Authentication Accessibility
6050. Provide visual feedback for non-technical consumers, including considerations for sensory disabled/impaired persons	High/mid-term	Users
6060. Establish a protocol for problems and errors How might you signal that a signature exists to be checked? How do you make clear what is **not** covered by the signature?	Medium/mid-term	Management Services

Description	Priority/timing	Areas
6070. Portability - pointing back to original document and its taken as a wholeness	High/long-term	Accessibility
6080. Online and offline checking options	Medium/short-term	Authentication
6090. Checking On and offline How do you check a signature? With what information returned? Can we have probabilistic/stochastic/not “black and white” but “green/yellow/red” options? Dealing with shades of error (e.g., right company, wrong representative)	Medium/short-term (checking) to mid to long term (stochastic)	Authentication Users
6100. Receiving Did I get what I actually asked for? Or just a valid document? Did I get the most recent version of the document, if earlier versions have been updated/corrected	High/short to medium-term	Accessibility Obsolescence
6110. Provide different feedback based on consumer (anonymous, by role)	Medium/long-term	Management
6120: Identify all parts of a document with a related signature	High/medium-term	Users

Description	Priority/timing	Areas
6130. Check information related to the signature	High/medium-term	Management
6140. Need a standardized set of return messages that are consistent and can be machine-interpreted for appropriate follow up/response in an automated environment.	High/medium-term	Management
6150. Need an approach that ultimately is not bound by language, so consumer can trust information reported and assured in any language.	High/medium-term	Users
7000) Following up		
7010. Provide feedback mechanism for issues from consumers based on signing	Medium/medium-term	Management
8000) Acting after-the-fact (including revoking signatures, obsolete documents, etc.)		
8010. Providing feedback to "new" checkers that a document has been superseded in some way	Medium/long-term	Obsolescence
8020. Providing a way to communicate to those who previously relied on a document that something has changed	Medium/long-term	Obsolescence

Description	Priority/timing	Areas
8030. Provide a mechanism for updating security on documents in the future (annual update/refresh?) to compensate for weakening signatures.	Medium/long-term	Integrity
8040. Ability to withdraw auditor's report.	High/mid-term	Authentication

What are the next steps?

Trade-offs relevant to short and long term issues:

Possible short term approach	"Ultimate" approach	Comments
Signed at the document level	Signed at the "data" level	Not dealing with data level assurance here per se, but signing pieces as opposed to the document as a whole.
Basic capabilities	Ease of use	We can start to get protection, but it may not be a transparent process at the beginning until developers create them once a foundation is laid.
Centralization	Pervasive / integrated	We can have solutions tied to people coming to the auditor's web site, the SEC, the AICPA in the short term instead of one that is widespread and integrated
Published hashes	Digital on the way to XML Signature	
"Analog" digital signature	"Digital" digital signature	

What underlying philosophies or believes undergird the next steps?

(This section to be filled out).

- Vendor and platform independence
- Standards-based
- Global, interoperable, holistic
- Collaborative, market developed
- Win-win-win for all participants

What technologies are obviously relevant?

- SHA1 (RFC 3174), SHA2, MD5
    - http://www.faqs.org/rfcs/rfc3174.html
    - http://tools.ietf.org/html/rfc4634
    - http://userpages.umbc.edu/~mabzug1/cs/md5/md5.html
- XML Security
    - http://www.w3.org/Security/
- IETF / W3C XML Signature
    - http://www.w3.org/Signature/
- IETF Cryptographic Message Syntax/ PKCS #7 (RFC 2630)
    - http://www.ietf.org/rfc/rfc2630.txt
- ETSI XML Advanced Electronic Signatures (XAdES - TS 101 733)
- IETF Time-stamp protocol (RFC 3161)
    - http://www.ietf.org/rfc/rfc3161.txt
- W3C XML Security Maintenance
    - http://www.w3.org/TR/xmldsig-core/
- W3C XML Canonicalization and Exclusive XML Canonicalization
    - http://www.w3.org/TR/xml-c14n/
    - http://www.w3.org/TR/xml-exc-c14n/
- OASIS DSS
    - http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=dss
- OASIS DSS-X
    - http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=dss-x
- Digital Object Identifier (DOI) Systems
    - http://www.doi.org/

What are the open questions?

1. Anyone, anywhere. How does that impact x.509 certificates or other internal-only certificates/signatures?
2. Dealing with issues Technical, Physical and Procedural
3. XBRL says to unwind security, transport, and handling and then validate. XBRL inside or outside? If inside, can it be added without modifying specification, taxonomies?
4. XBRL Canonicalization is different than XML Canonicalization; changes in XBRL that have no impact on the information therein (e.g., the physical order of facts in an XBRL instance normally don't matter) would potentially need to be taken into account.
5. XBRL "elements" can be signed, but if the namespace information changes, the safety at the element level isn't enough. The namespace can be unaltered, but if the taxonomy files change, the safety at the instance level isn't enough.
6. Differentiating between the XBRL documents and VIEWS of the XBRL documents.
7. Differentiating between digital signatures for evidentiary purposes and digital signatures for assurance/comfort purposes.
8. "The SEC's EDGAR[17] web site is a "trusted" site for US filings, practically speaking. Why don't people just confirm business reporting information that they find elsewhere with that company's postings at EDGAR? In Canada, they likewise have SEDAR[18]. There are equivalent sites globally. *Answer*: 1) The Companies House story, told earlier, shows that even trusted sites aren't perfect. 2) Not all companies that we are interested in have to post their filings to these sites. 3) It takes a lot of effort to find information at these sites; being able to have trustworthy access to data from any source (company web site, press release, or analyst report) with comfort of the authenticity and integrity of the data without having to manually confirm and compare is important to the efficient functioning of the market. 4) It is important to introduce trust at points before and after the "trusted third party" posting.

[17] http://www.sec.gov/edgar.shtml

[18] http://www.sedar.com/homepage_en.htm

XBRL's Unique Security Issues: An Introductory Overview

XBRL is a sophisticated Specification, and it is not appropriate to attempt to describe the technical nuances of the XBRL Specification in this document. However, we felt it might be helpful to illustrate some basic facts about how XBRL represents information and why simple file-level or even XML Signature functionality may not be enough for secure XBRL.

XBRL is XML. So it seems intuitive that XML tools should work well with XBRL. The XBRL Specification sets out the syntax and semantics of XBRL taxonomies and instances, based on W3C Recommendations, including XML, XML Namespaces, XML Schema and XLink. Some of the design deserves some extra attention when considering how to design a technical solution.

Facts in instances refer to taxonomies, which are (almost always) customized

An instance document (the primary holder of the company data) is a set of facts. It refers to a series of XML Schema (.xsd) files, each of which is related to a number of XLink (.xml) files which store additional information - descriptions, definitions, interrelationships and other attributes.

There are collaboratively developed taxonomies (schema/linkbase groupings) that are used in many cases. In addition, it is almost always a certainty that company data files include extensions to the collaboratively developed taxonomies - the "Extensibility" of XBRL. These extension, or customized, taxonomies can completely change the meaning of the individual and collective facts in the instance document.

Understanding any fact in an XBRL instance document, when there is customization, requires understanding what the taxonomies look like "at run time". Therefore, at the file level, signing an instance document as a whole ("file level") without also securing the referred to taxonomies is not secure. The process of bringing together the related taxonomies "at run time" is called "DTS discovery", where DTS stands for "discoverable taxonomy set".

Where the needs of security become more granular, securing XBRL data is not as simple as signing the content of an element, the element and its content, or even an XPath expression.

Facts in instances have many internal cross-references

For example, a simple fact in an instance may look like this:

```
<usfr-pte:SalesRevenueNetGoods contextRef="myc_0001193125-07-
  225883_STD_p9m_20060930_4" unitRef="USD" decimals="-
    6">22539000000</usfr-pte:SalesRevenueNetGoods>
```

Unless the namespace associated with "usfr-pte", which should look like this:

```
usfr-pte="http://www.xbrl.org/us/fr/common/pte/2005-02-28"
```

and which, in many cases, should be drawn only from the official location; the contextual information provided in the context referred to by the contextRef, which might look like this:

```
<context id="myc_0001193125-07-225883_STD_Inst_20070930_2">
<entity><identifier
scheme="http://www.sec.gov/CIK">0000009999</identifier></entity>
<period><startDate>2007-01-01</startDate><endDate>2007-12-
31</endDate></period></context>
```

and the *unit of measure* information associated with the unitRef, which might look like this:

```
<unit id="USD"><measure>iso4217:USD</measure></unit>
```

are secured, this fact is not secure.

The topic of *XBRL Canonicalization*, as a specialization of XML Canonicalization[19], is also one that may be of interest to the reader. Canonicalization, in simplistic terms, is the process of taking an XML document and running it through some steps that get rid of trivial differences that might otherwise say two logically equivalent documents are actually different.

XBRL Canonicalization as an extension to Canonical XML has not been formally defined, but one example of the special issues is provided here.

Because interrelationships between XBRL items are generally defined in linkbases, their order in an instance document is not an issue in processing. Two facts, such as:

```
<usfr-pte:InterestExpense contextRef="myc_0001193125-07-
225883_STD_p9m_20060930_4" unitRef="USD" decimals="-
6">291000000</usfr-pte:InterestExpense>
<usfr-pte:SellingGeneralAdministrativeExpenses
contextRef="myc_0001193125-07-225883_STD_p9m_20060930_4"
unitRef="USD" decimals="-6">1035000000</usfr-
pte:SellingGeneralAdministrativeExpenses>
```

do not change in meaning if they are instead in the instance document in this order:

```
<usfr-pte:SellingGeneralAdministrativeExpenses
contextRef="myc_0001193125-07-225883_STD_p9m_20060930_4"
unitRef="USD" decimals="-6">1035000000</usfr-
pte:SellingGeneralAdministrativeExpenses>
<usfr-pte:InterestExpense contextRef="myc_0001193125-07-
225883_STD_p9m_20060930_4" unitRef="USD" decimals="-
6">291000000</usfr-pte:InterestExpense>
```

[19] http://www.w3.org/TR/xml-c14n

Of course, where order is defined by a *sequence* in the complexType definition (used extensively in XBRL GL, used much less in financial reporting), order is important.

Risks and Problems

A short list of some of the risks and problems we are trying to deal with a solution for authentication and integrity.

One of the best known examples of corporate information being counterfeited by another is that of Emulex Corp., where a falsified press release resulted in the loss of $110 million to investors.

In another case, a Company wishes to mislead a company in a bidding situation. The Company falsified its financial statements, and provided an auditor's report without permission.[20] Forged auditor's reports also figured in a number of other situations, including SEC vs. Ocumed Group Inc.21, United States Attorney for the Northern District of Ohio charges against Jesse Bonner[22], New Zealand fraudster Wi Nepia, who made off with $160,000NZ using a falsified audit report[23]

The SEC reported numerous cases it investigated in Internet sweeps where third party "investment advisors" made suggestions based on falsified information, where both management's results and auditor's report are falsified.

Former employees can cause problems. In the Streamedia fraud, a former Chairman and Vice President issued a damaging press release reportedly on behalf of his former company. In the same way, ex-employees of audit firm may certify their favorite client's financial statements after leaving firm.

Auditors who determine that their audit report is wrong must inform those who relied on the audit report in the past, as well as those who might rely on it in the future.
- http://ca10.washburnlaw.edu/cases/2003/04/01-4147.htm

[20] http://www.iadb.org/integrity/oii_ar06/inv_external.cfm?language-english&detail=box6
[21] http://www.sec.gov/litigation/litreleases/lr18723.htm
[22] http://cleveland.fbi.gov/dojpressrel/2006/securities_fraud022306.htm
[23] http://www.salient.co.nz/index.php?a=2142&c=55