

**EFFICIENT EXTENSIBLE
MARKUP LANGUAGE (XML)
INTERCHANGE (EXI)**



IJIS Institute

*IJIS INSTITUTE TECHNICAL ADVISORY
COMMITTEE WHITE PAPER*

ACKNOWLEDGEMENTS

The IJIS Institute would like to thank the following members of the IJIS Institute Technical Advisory Committee for their work in authoring this whitepaper:

Dave Rader

Metatomix

Bob Slaski

Open Networks

Adrian Withy

Tiburon

The IJIS Institute would like to thank the following individuals and their sponsoring companies for their dedication and input to this document:

Joe Mierwa, *Committee Chair*

URL Integration

Prem Neelakanta, *Committee Vice-Chair*

Analysts International

Tom Carlson

Tom Carlson Consulting

Dominique Caspers

URL Integration

Marc Clifton

Interacx, LLC

Mike Hulme

Unisys

Fred A. Maymir-Ducharme

IBM

Catherine Plummer

JISP Liaison

Ian Warner

Computer Projects of Illinois (CPI)

Ben Acker, *Contributing Member*

CIBER

Bob Lorenz, *Contributing Member*

BIO-Key International

Vijay Mehra, *Contributing Member*

Deloitte Consulting

Vipul Mishra, *Contributing Member*

Syscon Justice Systems, Ltd.

The IJIS Institute would also like to thank the U.S. Department of Justice (DOJ) Office of Justice Programs (OJP) Bureau of Justice Assistance (BJA) for their comments and feedback.

This document draws heavily on the materials and content developed and published by the W3C EXI Committee and the authors hereby acknowledge their contributions.

This project was supported by Grant No. 2008-DD-BX-K009 awarded by the Bureau of Justice Assistance. The Bureau of Justice Assistance is a component of the Office of Justice Programs, which also includes the Bureau of Justice Statistics, the National Institute of Justice, the Office of Juvenile Justice and Delinquency Prevention, and the Office for Victims of Crime. Points of view or opinions in this document are those of the author and do not represent the official position or policies of the United States Department of Justice.

TABLE OF CONTENTS

Acknowledgements	i
Table of Contents	ii
List of Figures	ii
List of Tables	iii
Section 1. Purpose and Scope	1
Section 2. Concept	2
Section 3. Benefits of EXI	4
EXI Compression	4
Benefits of EXI Compression	4
Section 4. Relationship to NIEM and JRA	8
EXI and JRA	8
Section 5. Policy Implications	10
Section 6. Design Guidelines	11
Section 7. Interoperability	13
Section 8. Example	14
Section 9. W3C Efforts	17
Section 10. Alternative Solutions	18
Section 11. Related Technologies	20
Section 12. Hardware Accelerators	21
Section 13. Availability	22
Section 14. Summary	23
Section 15. References	24
Appendix A: Real World Example	25
About the IJIS Institute	31

LIST OF FIGURES

Figure 1 - Simple representation of EXI data transmission	5
Figure 2 - EXI compactness compared to Gzipped XML	6
Figure 3 - EXI encode speed (without compression)	7
Figure 4 - Example 1 notebook (XML document)	14
Figure 5 - Example 2 EXI Body Stream	15
Figure 6 - EXI Compression	16
Figure 7 - Example 4 BRDMLRECCEORD.xml XML Instance	28
Figure 8 - Example 5 BRDMLRECCEORD Uncompressed EXI Instance	29
Figure 9 - Example 6 BRDMLRECCEORD Compressed EXI Instance	30

LIST OF TABLES

Table 1 - Summary of design guidelines used to create EXI..... 12

SECTION 1. PURPOSE AND SCOPE

This emerging technology paper considers Efficient XML Interchange (EXI) as an emerging technology applicable to integrated justice and public safety data sharing. The need for efficient or binary XML has risen dramatically with the increased use of XML. EXI was identified as an efficient XML solution through of the extensive and open process that was used by the [World Wide Web Consortium \(W3C\)](#). This W3C effort was performed by the XML Binary Characterization Working Group and later by the EXI Working Group. In the absence of EXI, simple compression solutions such as Gzip have been used to achieve some of the goals of EXI.

The paper provides a brief overview and summary of the implications of EXI (<http://www.w3.org/XML/EXI/>) in the context of the [Justice Reference Architecture \(JRA\)](#) and the [National Information Exchange Model \(NIEM\)](#). XML has gained significant use and adoption within many information technology-related fields; integrated justice and public safety are no exception.

The [Global Justice XML Data Model \(GJXDM\)](#) and NIEM provide a common data model within the integrated justice and public safety industry; however, they are based on XML which increased the size and complexity of exchanges; this increase in size impacts network load, processing (parsing) time, and storage (memory and physical storage) of systems which process the XML data. EXI is a specification for encoding XML messages into a binary representation which provides two primary benefits:

1. Reduction in the size of the XML documents which use less network bandwidth and require less storage space.
2. Efficient encoding of the XML documents which improves processing speed, is less susceptible to processing issues (e.g., requiring escaping of special characters such as "<" or ">"), and provides a modest level of security by encoding.

SECTION 2. CONCEPT

Efficient XML Interchange (EXI) is "a very compact, high performance XML representation that was designed to work well for a broad range of applications. It simultaneously improves performance and significantly reduces bandwidth requirements without compromising efficient use of other resources such as battery life, code size, processing power, and memory."
- W3C¹

In essence, EXI is a compact XML representation that is:

- ◆ Efficient – since EXI is compact, it takes less space for memory, storage, and bandwidth;
- ◆ Straightforward – EXI uses a grammar-driven approach to encoding and provides for "processors" which are application programs that encode and decode an EXI stream; and
- ◆ Encoded – encoding of the EXI stream reduces the content size and also avoids data reserved character issues (e.g., <, >, &, etc.)

EXI was developed around five key principles – the format had to be general, minimal, efficient, flexible, and interoperable. The first two principles – general and minimal – account for the non-invasiveness of EXI. The specification is meant to leverage existing XML

technologies such as schemas rather than replace them with another format or methodology. Efficiency is provided by several components such as the compact nature of EXI streams and the fact that EXI uses information from the schema to improve compactness and processing efficiency. EXI provides flexibility in that schemas are not required to achieve accurate, complete, or compact EXI streams. EXI is compatible with the XML Information Set, or XML Infoset, which permits EXI to have the potential for minimizing the impact on XML application interoperability while still providing a more efficient alternative to XML syntax and grammar.

EXI Streams are the basic structure of EXI documents and consist of an EXI Header followed by an EXI Body. The EXI Header contains document metadata information such as the format version and options that were used during encoding. The EXI Body contains a sequence of "events" which comprise the contents of the body. This can best be explained by a comparison to XML syntax notation as follows:

"XML items are encoded into one or more EXI events; for example, an attribute named foo can be encoded as AT("foo") and an element named bar as the pair of events SE("bar") and EE."
- W3C²

¹ <http://www.w3.org/TR/2009/CR-exi-20091208/>

² <http://www.w3.org/TR/2009/WD-exi-primer-20091208>

For more information on the EXI Body refer to Section 2.1.2 EXI Body of the W3C EXI Primer, <http://www.w3.org/TR/2009/WD-exi-primer-20091208/#exiBody>. Please refer to [Section 8: Example](#) in this document for

more information and examples of EXI documents.

The following sections of this concept paper will delve further into EXI, expanding upon the benefits of EXI and the possible relationships between EXI, NIEM, and JRA.

SECTION 3. BENEFITS OF EXI

This section will provide an overview of the benefits of EXI in the context of information sharing architectures. As EXI's primary goal is efficiency of XML data representation, the benefits of EXI adoption are seen in areas where XML is used on-the-wire, such as XML web services, as well as where XML is persisted to physical storage mechanisms, such as a relational database. In both on-the-wire communication as well as persistence, the size and efficiency of data representation can be critical.

EXI Compression

At its highest level, EXI processing is done by an *EXI Processor* that can assume the role of either an *EXI Encoder* or an *EXI Decoder*. Figure 1 shows an XML document moving through a typical EXI data processing workflow.

EXI uses the XML Infoset data model in its application of a binary representation and compression algorithm. By making assumptions as to the structure of the data and not relying on the literal text representation, EXI is able to compress XML into a structured sequence of bytes without the verbose tagged structure.

Quoting the W3C EXI Working Draft:

“EXI achieves broad generality, flexibility, and performance, by unifying concepts from formal language theory and information theory into a single, relatively simple algorithm. The algorithm uses a grammar to determine what is likely to occur at any given point in an XML document and encodes the most likely alternatives in fewer bits. The fully generalized algorithm works for any language that can be described by a grammar (e.g., XML, Java, HTTP, etc.); however, EXI is optimized specifically for XML languages.”

- W3C EXI Candidate

Recommendation (Section 3, Basic Concepts)

Benefits of EXI Compression

The two primary benefits of EXI are data compression and processing efficiency. A decreased data footprint will have an impact on bandwidth as well as database hardware requirements. Increased

processing efficiency has the potential to alleviate processing strain on server hardware, as well as enable the spread of XML capabilities to less-capable mobile devices. This section contains charts from the *Efficient XML Interchange Evaluation, Second Public Working Draft, 7 April 2009*, which are themselves derived from the detailed measurements in the *Working Draft of the Efficient XML Interchange Measurements Note*, updated (25 July 2007). Figure 2 is a diagram that compares EXI compression to Gzip, a commonly used compression technology.

As shown in Figure 2, EXI can compress an XML document anywhere from 1.4 times to 100 times the document's original size and over 10 times that of Gzip.

A secondary benefit to EXI is the encoding and decoding speed of EXI data compared to that of XML. Figure 3 shows the relationship between the speed of encoding and decoding raw XML to that of EXI encoded data.

As shown in both Figure 2 and Figure 3, EXI can be orders of magnitude more efficient to encode and decode versus raw XML. The median speed of encoding performance gain was 2.4 times faster, and 6.7 times faster decoding.

An additional benefit of EXI adoption is increased cross-platform interoperability as implementers adopt EXI as an alternative to platform-specific binary encoding mechanisms.

It is additionally worth noting that although EXI encoding can be a more efficient mechanism for encoding XML, there are trade-offs in the ability for EXI to be queried, audited, and verified for integrity.

These performance tests show that EXI can dramatically decrease bandwidth, storage, and performance requirements of XML intensive applications.

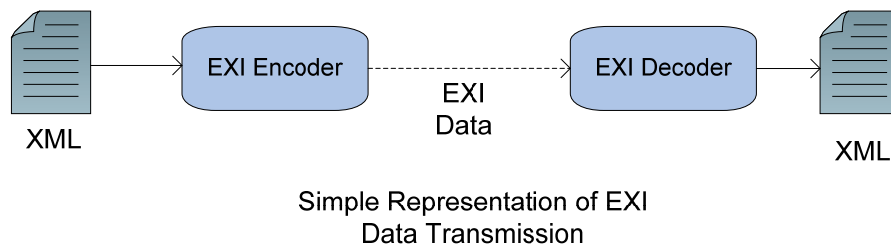


Figure 1 – Simple representation of EXI data transmission.

EXI Compactness Compared to Gzipped XML

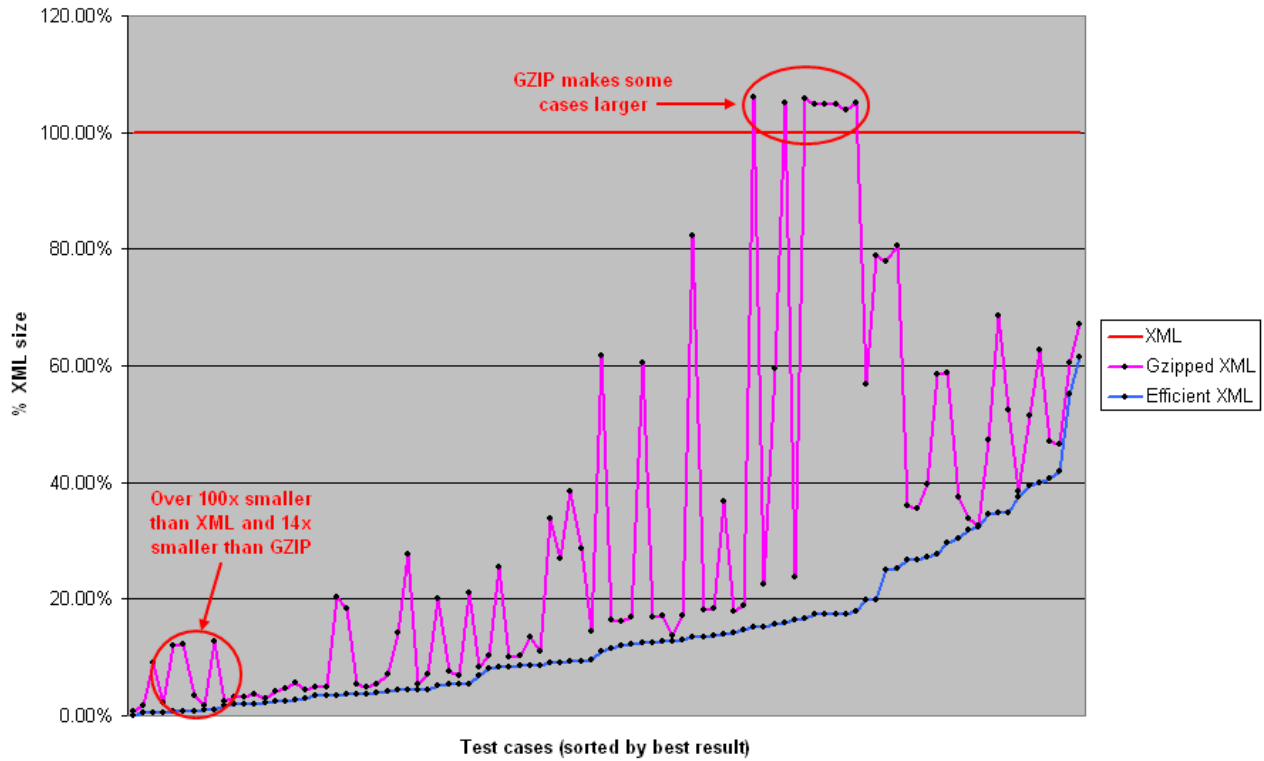


Figure 2 – EXI compactness compared to Gzipped XML.

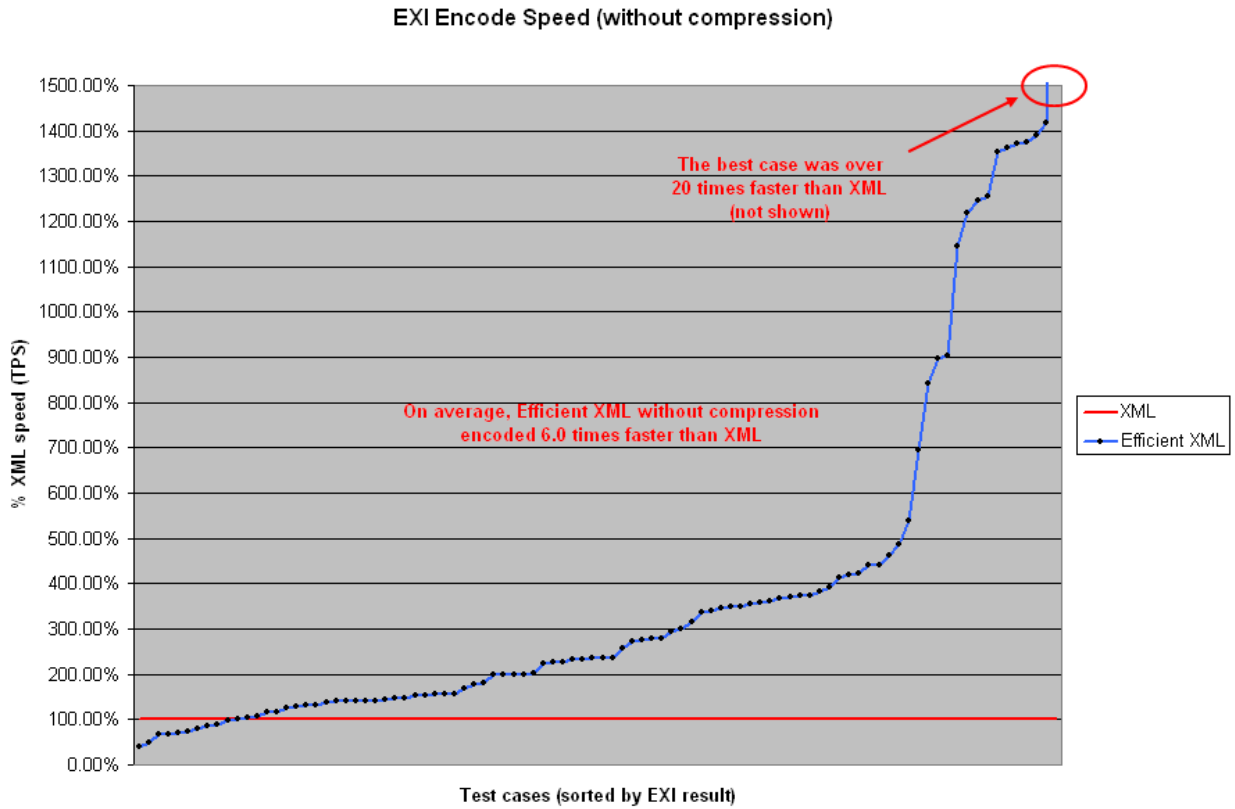


Figure 3 – EXI encode speed (without compression).

SECTION 4. RELATIONSHIP TO NIEM AND JRA

EXI has a strong relationship to NIEM in that EXI is an XML technology. The adoption of EXI, however, does not create dependencies with NIEM, as EXI is designed to be transparent to the XML interchange process and easily integrated into current XML implementations. As NIEM is built on fundamental XML technology, EXI is a compression process and algorithm that is independent of NIEM implementation. This will allow NIEM to develop in parallel without forming negative couplings on EXI standards.

EXI provides numerous benefits for NIEM exchange packages; the foremost address data volume (by compression of data) and data security (by encoding of data). To illustrate this process, envision an exchange of data from a state agency, such as a Department of Public Safety, to the [National Crime Information Center \(NCIC\)](#). This exchange of information includes common information such as a subject's name, address, and date of birth, and it also includes sensitive information such as the subject's Social Security information and charges related to an ongoing case. In most integrated justice applications, this information is sent in XML format (NIEM XML) over a secure (SSL) communication channel. EXI can aid the sending of this information by encoding the XML information. Encoding the information entails transforming the data from a standard, human-readable, XML format into a binary representation of the data. The encoded data yields information which is not human readable (e.g., the data is obfuscated) and the data is significantly compressed as compression algorithms are applied during encoding which replace

redundant data sets with more compact pointers. (For more information on how EXI compresses data see [Section 3. Benefits of EXI: EXI Compression](#)).

The [U.S. Department of Justice's](#) Logical Entity eXchange Specification (LEXS) is a NIEM-conformant specification that provides for law enforcement information exchanges. Because LEXS is based entirely on NIEM XML, LEXS exchange packages can be exchanged using EXI in the same manner as any other NIEM exchange package.

EXI and JRA

The JRA has the potential for interdependencies with EXI. As EXI becomes incorporated into web service implementations and web service contract standards, the JRA will need to take into account the adoption of such standards and develop interchange best practices.

One important consideration is the use of encryption. XML data that is already encrypted by an application cannot be processed by EXI. Consequently, JRA best practices must carefully consider when and where encryption occurs.

Other areas where EXI will have an impact include WS-Policy statements and the Web Service Description Language (WSDL). At present, *EXI Best Practices* recommend a very permissive policy with respect to accepting EXI transactions. In the future more complex policies will be defined and codified via the WSDL. There may also be tradeoffs with other binary data technologies, e.g., XML Optimized Packaging (XOP) and Message Optimization Transmission Method.

It is also important to point out that EXI can be used in information sharing systems in varying ways. Take the above Department of Public Safety system for example – this system may already have the logic to send XML data to NCIC but would like to leverage the encoding and compression capabilities from EXI. This could be accomplished in one of two ways: 1) alter the existing DPS system to include encoding and compressing the data with EXI; or, 2) intercept the message as it leaves the DPS system and perform the encoding and compression before it is sent to NCIC. JRA offers a better approach as system functionality is generally encompassed as reusable ‘services’. JRA-compliant systems could implement an EXI service which is used to encode and compress data at the appropriate time. The benefit to this

approach is that core business logic will remain abstracted from EXI; often developers of the core business logic will not need to know about EXI as it is a service which is called and applied as needed in the service process chain.

It is also possible to use EXI for exchanges that do not implement JRA Service Interaction Profiles. For example, EXI-encoded NIEM can be used with proprietary exchanges such as IBM MQ® or in conjunction with simple REpresentational State Transfer (REST) exchanges.

Overall, EXI can promote NIEM and JRA adoption by addressing the biggest user concerns: transaction sizes, network bandwidth, and processing load.

SECTION 5. POLICY IMPLICATIONS

This section explains the policy implications of EXI and how EXI compression affects current XML technology policy.

EXI has the potential to be entirely transparent to the current XML exchange process. In the simplest case, existing XML exchanges can be transformed into EXI as they are exiting one system and transformed back to XML 1.0 format on the way in to the receiving system. However, there are some concerns.

One important consideration is that the transformation to and from EXI produces a functionally equivalent XML document but not an identical one. EXI is based on the formal XML Infoset grammar rather than the less rigorous XML 1.0 specification. A reconstructed EXI may be and look different than the original document in superficial respects such as the elimination of “white space” and redundant namespaces. There may be “non-repudiation” concerns if XML documents are not delivered exactly as sent. These same problems exist and are being solved on a lesser scale today with the canonicalization requirements associated with digital signature. Verification of a digital signature requires an exact copy of the original document not a functional XML Infoset equivalent.

Transparent EXI would not be possible where data is encrypted since EXI is based

on an understanding of the XML document and underlying schema. Consequently, provisions would need to be made for EXI to be incorporated into the security model and associated policies of all organizations that are participating in the EXI exchanges.

The smaller size of EXI documents may indirectly influence implementation decisions and policies as well. For example, the dramatically smaller size may allow the direct transmission of XML documents to mobile data computers that currently are highly-restricted by limited narrowband communications. Routine transmission of additional data, such as full driver history, may be practical and, consequently, mandated. The smaller size may also influence data retention policies. Storage has ever increasing capacity but is not unlimited. EXI offers the potential for more than an order of magnitude decrease in storage requirements which conversely could mean a retention period that is an order of magnitude longer.

The Justice Reference Architecture provides the framework to address policies and implementation issues associated with the use of EXI for information exchange.

SECTION 6. DESIGN GUIDELINES

This section provides a summary of the design guidelines that were used to create EXI. The Efficient XML Interchange Evaluation, W3C Candidate Recommendation, provided a

comprehensive assessment of EXI when considering the desired XML Binary Characterization properties. The W3C assessment is incorporated below.

Desired Property	Efficient XML Interchange Assessment
Directly Readable and Writable	Implementations can read and write EXI streams directly via standard XML APIs, such as DOM, SAX and StAX.
Transport Independence	EXI can be used over TCP, UDP, HTTP and various wireless and satellite transports.
Compactness	In general, the W3C assessment measured an order of magnitude reduction in size (greater compactness) for EXI in comparison to XML.
Human Language Neutral	EXI supports all standard character set encodings.
Platform Neutrality	The EXI format specification does not make particular assumptions about the platform architecture. An implementation already exists for several popular server, desktop and mobile platforms, including Java EE/SE, Microsoft .NET, Java Mobile Edition and .NET Compact Framework.
Integratable into XML Stack	EXI was designed to integrate well into the XML stack, neither duplicating nor requiring changes to functionality at other layers in the XML stack. It builds on the XML Infoset data model. It implements the same character encodings as text XML and supports the common interfaces as existing XML parsers and serializers. As such, it can be inserted into existing XML applications with minimal time and cost.
Royalty Free	EXI will be made available per the W3C patent policy.
Fragmentable	EXI can represent any collection of XML fragments extracted from any collection of XML documents. All schema optimization, bit-packing and XML compression algorithms apply equally to fragments.
Roundtrip Support	EXI supports lossless equivalence for PSVI, Infoset and lexical applications, such as XML Digital Signatures. The EXI "preserve" option can be used when this property is needed.
Generality	EXI is competitive with alternatives across a diverse range of XML documents, applications and use cases.
Schema Extensions and Deviations	EXI includes schema optimizations that support arbitrary schema extensions and deviations. Applications may specify strict or extensible schema handling and may provide a full schema, partial schema or no schema at all.

Desired Property	Efficient XML Interchange Assessment
Format Version Identifier	EXI header includes version.
Content Type Management	EXI can be used in various contexts, some which use a media type and some which use content encoding, or both.
Self-Contained	When schema optimizations are not used, EXI documents are always self-contained.
Processing Efficiency	Current implementations achieve performance several times faster than XML using both in-memory tests and more realistic scenarios that involve file and network IO. These implementations do not depend on compile-time schema-binding techniques that make dynamically acquiring, loading or updating schemas impractical or impossible.
Additional properties that were not assessed with the formal testing processing are listed below.	
Streamable	EXI can be used in streaming applications.
Small Footprint, Widespread Adoption, Space Efficiency, Implementation Cost, Forward Compatibility	These properties are generally associated with implementations. Since there are currently only a limited number of EXI implementations, the only assessment that can be made is that EXI does not prevent or preclude a positive outcome for each of these properties.

Table 1 – Summary of design guidelines used to create EXI.

SECTION 7. INTEROPERABILITY

This section summarizes the issues and capabilities associated with EXI and XML interoperability. The W3C has published a draft *EXI Best Practices* document that largely deals with interoperability. The key interoperability issues are distinguishing EXI from XML in a mixed environment, fully preserving the XML Infoset when needed and insuring the integrity of XML security features, particularly XML signature and XML encryption.

Because of the widespread adoption of XML and the limited use of EXI, EXI documents should only be sent when specifically requested or by prior agreement. Differentiating XML from EXI is accomplished by using the distinguishing bits which is a leading bit sequence of "10". This allows EXI to be processed separately and allows the application to avoid sending EXI to XML parsers. HTTP content negotiation is the recommended approach for dynamically determining the EXI exchange capabilities. EXI also defines the use of an optional EXI "cookie" as a means of differentiating EXI and XML.

EXI provides for varying degrees of precise retention of XML 1.0 by offering "fidelity" options. The areas of fidelity are the same as those associated with the XML Infoset and the broad issue of canonicalization: superfluous white space, preserving comments, redundant or less than optimal namespace references, and lexical consistency when arranging elements and attributes. There are also some restrictions associated with file names and MIME-type specifications that need to be considered when exchanging EXI. Fidelity issues need to be considered when precise rendering is important. Provisions have been made to retain necessary fidelity for message-level security. In particular, EXI canonicalization is consistent with the use of XML signature and XML encryption.

The ability to "view source" is also recommended as a best practice. XML developers have become accustomed to this capability because of the text-based nature of XML.

SECTION 8. EXAMPLE

This section provides an example of XML to EXI encoding.

EXI is a binary representation of XML data and, as such, is difficult to provide a meaningful example as the XML data is encoded. Encoding makes the document difficult to show a direct mapping from an XML instance to the corresponding EXI representation. While this is a drawback in explaining EXI, it can be a benefit in exchanging data in an EXI format as it adds a layer of obfuscation to the payload.

An EXI document consists of Header and Body section. The Header section defines encoding properties needed to decode the Body section along with other metadata related to the EXI document (e.g., encoding options, compression, or user defined options for encoding/decoding). The Body section contains the encoded version of the XML document per the encoding options specified in the Header section.

The following section is a summary from the [EXI Primer, Section 2.1 EXI Streams](#).

The Body section of an EXI document is composed of a sequence of EXI events. XML items are encoded into one or more events; for example, an attribute named *type* can be encoded as AT("type") and an element named *Person* as the pair of events SE("Person") and EE. AT, SE and EE are acronyms for attribute type, start element, and end element, respectively. The EXI Specification defines all events needed to encode an XML document into EXI. This allows decoding of the EXI document using the Header options and the decode events provided in EXI.

To better illustrate an XML document mapping to an EXI (uncompressed) encoded document, consider the following example XML depicting two notes within a notebook (Figure 4). The sequence of EXI events corresponding to the body of this XML document is shown in Figure 5.

```
<?xml version="1.0" encoding="UTF-8"?>
<notebook date="2007-09-12">
  <note date="2007-07-23" category="EXI">
    <subject>EXI</subject>
    <body>Do not forget it!</body>
  </note>
  <note date="2007-09-12">
    <subject>Shopping List</subject>
    <body>milk, honey</body>
  </note>
</notebook>
```

Figure 4 - Example 1 notebook (XML document).

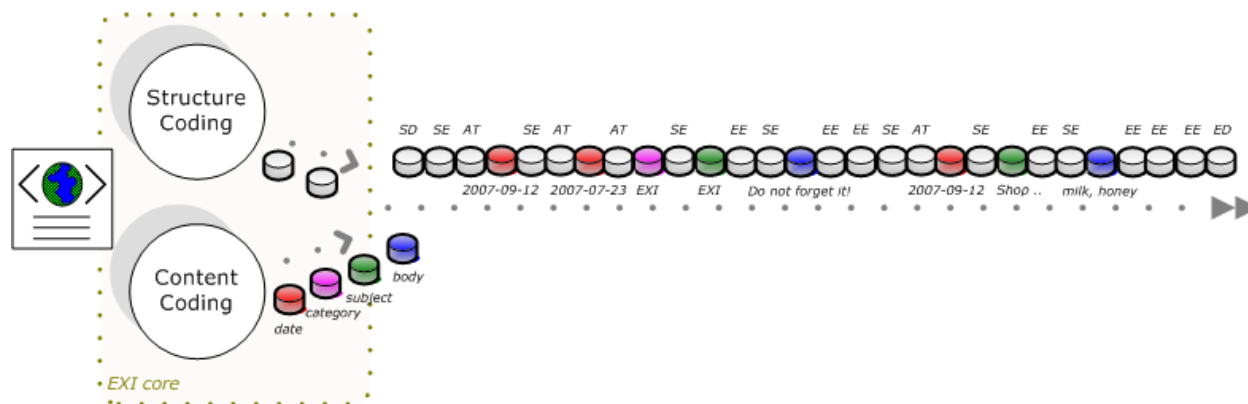


Figure 5 – Example 2 EXI Body Stream.

This sequence of EXI events can be easily mapped to the structure of the XML document shown above. Every document begins with a SD and ends with an ED. The SE event indicates the “Start Element”, the first SE correlates to the <notebook> element, followed by the AT event which correlates to the @date attribute to the <notebook> element. Subsequent events can be traced correspondingly to their counterparts in the XML document listed in Figure 4 - Example 1.

It is important to note that the above example, Figure 5 - Example 2 EXI Body Stream represents an EXI encoded instance of an XML document which has not been compressed. As mentioned before, the EXI specification provides functionality for compression within the EXI document. Typically, compression of exchange information would be applied on the entire item being exchanged as a whole (e.g., the entire XML document would be compressed). Because of EXI's inherent knowledge of the XML structure, EXI provides the ability to compress within a document.

The mechanism used to combine homogeneous data is simple and flexible enough so that it can be used in all forms of

EXI encoding. Element and attribute values are grouped according to their qualified names while structure information, like Event Codes, is combined. To keep compression overhead at a minimum, smaller QName channels are combined while larger channels are compressed separately.

To illustrate this process, refer to Figure 5 - Example 2 above. The figure uses grey buckets for structure information and colored buckets for content information. The color is determined by the associated QName (e.g., date, category, subject, body). XML instances can be seen as a combination of structure and content information. The content information can be further divided in different sections according to the context (surrounding structure as indicated by a QName). EXI treats XML instances this way and uses these implied partitions, so called channels, to provide blocked input to a standard compression algorithm. This blocking of similar data increases compression efficiency.

Note: An alignment phase creates a byte-aligned representation of event codes and content items that is more amenable to compression algorithms compared to unaligned representations. Most compression algorithms operate on a series of bytes to identify redundancies in the octets.

By combining smaller channels into the same compressed stream while others are compressed separately EXI keeps the compression overhead at a minimum. The mechanism to determine whether channels are combined or compressed separately is guided by the number of value content items present in the EXI stream. For small

documents (≤ 100 value content items) EXI uses a single compressed stream while larger documents (> 100 value content items) result in several independent compressed streams. The reader is referred to the [EXI specification](#) for further details.

The notebook example, Figure 4 - Example 1 above, falls in the first category and is encoded as a single compressed deflate stream containing first the structure channel, followed by the QName channels in the order they appear in the document (date, category, subject, body).

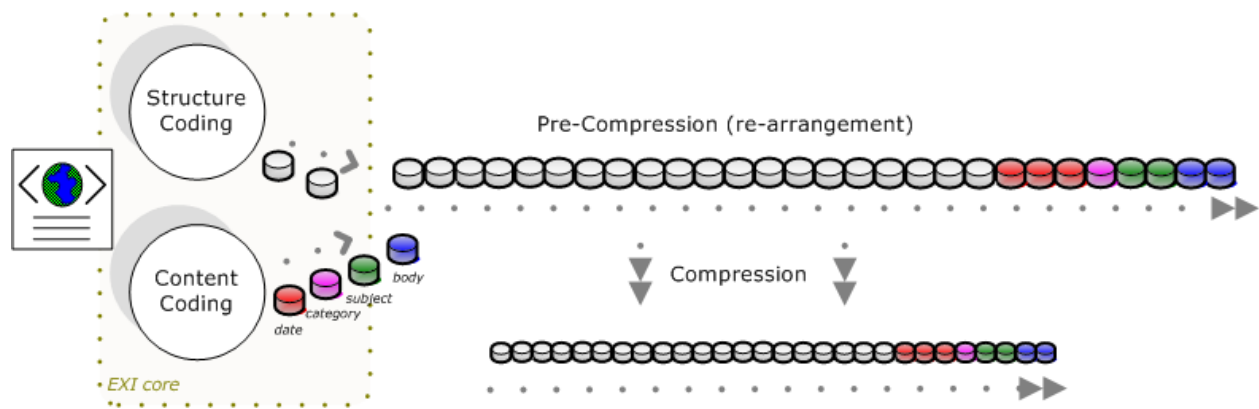


Figure 6 – EXI Compression.

SECTION 9. W3C EFFORTS

There are five levels of W3C recommendations:

1. Working Draft
2. Last Call Working Draft
3. Candidate Recommendation
4. Proposed Recommendation
5. W3C Recommendation

The first EXI Candidate Recommendation was published December 2009. The timetable for approval as a Recommendation is uncertain.

SECTION 10. ALTERNATIVE SOLUTIONS

This section discusses alternative solutions to EXI. There are several alternatives to the features EXI provides; however, there are few alternatives which bring equivalent benefits to integrated justice and public safety systems. The following list shows alternatives to EXI that offer similar functional benefits as EXI, namely compactness of the XML document and encoding of the XML content:

- ◆ Fast Infoset - designed to optimize compression, serialization, and processing, uses tables and indexes to accomplish benefits, generally results in a more compact representation of the XML document; however it does not provide encoding, is extremely extensible - for more information see: <http://www.w3.org/TR/exi-measurements/#contributions-FI>
- ◆ X.694 ASN.1 with BER - provides similar functionality to EXI using ASN.1 to encode document contents, requires a schema for the document, some components of a document cannot be represented in the encoded document (randomly occurring comments and processing instructions) - for more information see: <http://www.w3.org/TR/exi-measurements/#contributions-ASN1-BER>
- ◆ X.694 ASN.1 with PER - similar to X.694 ASN.1 with BER but uses X.694 to map from XML schema document to ASN.1, provides a more compact document than BER - for more information see: <http://www.w3.org/TR/exi-measurements/#contributions-PER>
- ◆ Xebu - derived from research into XML processing on mobile devices, intended to have small foot print, models an XML document as a sequence of tokenized events similar to StAX/SAX - for more information see: <http://www.w3.org/TR/exi-measurements/#contributions-xebu>
- ◆ Extensible Schema-Based Compression (XSBC) - uses a document schema to build a 'lookup table' for an XML document and uses this lookup table to transcode the XML document into a binary format, XML Schema is required, straightforward mapping between XML document and the binary representation - for more information see: <http://www.w3.org/TR/exi-measurements/#contributions-xsbc>
- ◆ Fujitsu XML Data Interchange Format (FXDI) - designed to provide greater efficiencies in encoding and compactness, more complex solution requiring the Fujitsu Schema Compiler to compile a schema into a "schema corpus", works well with conventional documents - for more information see: <http://www.w3.org/TR/exi-measurements/#contributions-fxdi>
- ◆ X.694 ASN.1 with PER + Fast Infoset - uses both X.694 with PER and Fast Infoset, accounts for lack of schemaless support in X.694 ASN.1 with PER, more performant with schema - for more information see: <http://www.w3.org/TR/exi-measurements/#contributions-perfi>

- ◆ Efficiency Structured XML (esXML) - encodes XML data model in a flexible, compact, and efficient to process manner, provides a variety of encoding methods, and uses a meta-structure for describing the data mapping - for more information see:
<http://www.w3.org/TR/exi-measurements/#contributions-esxml>

Of these alternatives, Fast Infoset has gained considerable support including an open source implementation as part of the Sun Glassfish initiative. Fast Infoset is based upon the proven Abstract Syntax Notation 1 (ASN.1).

SECTION 11. RELATED TECHNOLOGIES

EXI is an efficient binary representation of the XML Infoset. The XML Infoset was defined as the more semantically rigorous definition of XML 1.0. Current and future XML specifications use or will use the XML Infoset as the base underlying specification. All of the issues of fidelity are inherent to the XML Infoset since does not provide a byte-for-byte representation of XML 1.0 even though it is fully functionally equivalent.

XML-binary Optimized Packaging (XOP) has a goal similar to EXI in that it provides an efficient representation of binary “objects” within XML documents without

the transformation to base64 encoded text. EXI provides this capability for the entire XML document. The Message Transmission Optimization Method (MTOM) defines how to transfer XOP “packages” using SOAP and how to specify the packages with WSDL.

In order to avoid intermediate translation of binary data between the application and the underlying transport, the Web Services Description Language must make provisions for binary data.

SECTION 12. HARDWARE ACCELERATORS

Hardware XML accelerators can be used to process XML more efficiently. Depending on the application, increases in efficiency can be one or two orders of magnitude. Accelerators typically use proprietary internal binary XML formats.

Because of the proprietary formats used by accelerators, the same proprietary accelerator would need to be used on both ends of a compressed data exchange. As a result, accelerators cannot be used for solving the bandwidth problems in a heterogeneous computing environment.

SECTION 13. AVAILABILITY

The availability of EXI processors has been limited but is increasing. The following is a list of software vendors who supply Efficient XML processing libraries for both .NET and the open source community.

AgileDelta (www.agiledelta.com) provides a Java EXI SDK as well as a .NET EXI WCF binding.

EXificient (exificient.sourceforge.net) is an open source implementation.

SECTION 14. SUMMARY

A fully open process has been used to select and validate EXI through the W3C. The W3C EXI effort was performed in an exhaustive two step process - the XML Binary Characterization Working Group and later by the EXI Working Group. EXI is a technically superior solution. However, EXI is currently stalled in terms of building full consensus. The W3C EXI standard is currently a Candidate Recommendation which must formally proceed to Proposed Recommendation and, finally, Recommendation.

Competing solutions, notably Gzip and Fast Infoset, are already available and deployed.

Fast Infoset is based on the widely used ASN.1 data standard and available in the open source Sun Glassfish. Gzip is very simple, open source and very effective way to address compression but does not address more efficient compression.

EXI is very much still emerging and there is not a clear consensus on the EXI standard. Commitment to EXI should be limited until at least the W3C has approved EXI as a W3C Recommendation.

SECTION 15. REFERENCES

- ◆ EXI Format specification [Candidate Recommendation](#), 08 December 2009
- ◆ EXI Primer, [Second Public Working Draft](#), 08 December 2009
- ◆ Efficient XML Interchange Evaluation, [Second Public Working Draft, 7 April 2009](#)
- ◆ EXI Impacts, [First Public Working Draft, 3 September 2008](#)
- ◆ EXI Primer, [Second Public Working Draft](#), 08 December 2009
- ◆ EXI Best Practices , [First Public Working Draft, 19 December 2007](#)
- ◆ Working Draft of the [Efficient XML Interchange Measurements Note](#), updated (25 July 2007)
- ◆ Measurement Test Framework - see [below](#) (18 May 2007)
- ◆ [Raw results](#) of candidates testing

APPENDIX A: REAL WORLD EXAMPLE

To better illustrate an XML document as it would appear in EXI, uncompressed and compressed, consider the following example.

The following data has been provided by the [MOVES Institute](#) as an example of an XML instance and its representation in EXI, both compressed and uncompressed. All data provided below can be obtained in their original form from the “Encoded Files Directory” link on the [MOVES Institutes EXI Examples](#) site.

Example 4 (Figure 7) provides the XML instance the ASMTF/brdmlreceord/BRDMLRECCEORD.xml XML document which will be encoded into an EXI document.

```
<?xml version="1.0"?><?iris lemc=0511231028?><mtf:national_bridge_demolition_reconnaissance_order
mtfid="BRDMLRECCEORD" xmlns:mtf="asmtf2006:message" xmlns:xsi="http://www.w3.org/2001/XMLSchema-
instance"><operation_identification_data><operation_codeword>BIG
BUSTER</operation_codeword></operation_identification_data><message_identification><message_identifier>
BRDMLRECCEORD</message_identifier><originator>DJFHQ</originator><message_serial_number>001</messag
e_serial_number><month_name>MAR</month_name></message_identification><reference><serial_identifier>
A</serial_identifier><type_of_reference><reference_description>VTC</reference_description></type_of_refere
nce><originator>LHQ</originator><date_and_or_time_of_reference><day_time_and_month_of_reference><day
>11</day><hour_time>12</hour_time><minute_time>30</minute_time><time_zone>Z</time_zone><month_n
ame>MAR</month_name></day_time_and_month_of_reference></date_and_or_time_of_reference></referen
ce><geodetic_datum><geodetic_datum>AFG</geodetic_datum></geodetic_datum><organization_designator_of
_drafter_releaser><unit_designation_name>1-
503IN</unit_designation_name><unit_size_indicator>BN</unit_size_indicator><geographical_entity>AS</geogr
aphical_entity><unit_role_indicator_code_a>INF</unit_role_indicator_code_a><unit_role_indicator_code_b>RE
CCE</unit_role_indicator_code_b><unit_role_indicator_code_c>ABN</unit_role_indicator_code_c><unit_role_i
ndicator_code_d
xsi:nil="true"/><higher_formation_name>2BDE</higher_formation_name><organization_code><armed_service>
A</armed_service></organization_code><unit_identification_code_uic><nationality>AS</nationality><armed_se
rvice>F</armed_service><file_sequential_location_number>1393</file_sequential_location_number></unit_ide
ntification_code_uic></organization_designator_of_drafter_releaser><map_data_without_geodetic_datum><m
ap_or_chart_series_designator>1501</map_or_chart_series_designator><map_or_chart_suffix_number>16</m
ap_or_chart_suffix_number><group_of_fields><map_or_chart_sheet_number>NM32-
1ESSEN</map_or_chart_sheet_number><map_or_chart_edition_number>3-
DMG</map_or_chart_edition_number></group_of_fields></map_data_without_geodetic_datum><task_serial_
number><task_serial_number>12-
TH4</task_serial_number></task_serial_number><obstacle_identifier><obstacle_identification><obstacle_numb
er><regional_identifier>A</regional_identifier><corps_troops_or_unit_designator>A</corps_troops_or_unit_de
signator><division_troops_or_brigade_designator>2</division_troops_or_brigade_designator><tactical_obstacle
_identifier>A</tactical_obstacle_identifier><tactical_obstacle_number>12</tactical_obstacle_number></obstacl
e_number></obstacle_identification></obstacle_identifier><obstacle_zone_belt_name_or_number><obstacle_z
one_belt_number_or_name><obstacle_zone_belt_name>BUFFALO1</obstacle_zone_belt_name></obstacle_zo
ne_belt_number_or_name></obstacle_zone_belt_name_or_number><location_and_nearest_feature><geograp
hic_place_name>ADELAIDE</geographic_place_name><group_of_fields><utm_grid_reference><here>10ADA123
45678</here></utm_grid_reference></group_of_fields><group_of_fields><utm_grid_reference><here>10ADA1
23456790</here></utm_grid_reference></group_of_fields></location_and_nearest_feature><trafficable_when
_prepared_for_demolition><trafficable_indicator>YES</trafficable_indicator></trafficable_when_prepared_for_
demolition><bridge_demolition_status><demolition_status>PRO</demolition_status></bridge_demolition_statu
s><reconnaissance_timing><dtg_to_submit_report_by><day>12</day><hour_time>13</hour_time><minute_ti
me>0</minute_time><time_zone>Z</time_zone><month_name>MAR</month_name><year>2003</year></dtg
_to_submit_report_by><no_earlier_than_dtg_for_recon_to_begin><day>12</day><hour_time>11</hour_time>
<minute_time>0</minute_time><time_zone>Z</time_zone><month_name>MAR</month_name><year>2003</
year></no_earlier_than_dtg_for_recon_to_begin><dtg_for_recon_to_be_completed><day>12</day><hour_tim
e>10</hour_time><minute_time>0</minute_time><time_zone>Z</time_zone><month_name>MAR</month_na
```

```
me><year>2002</year></dtg_for_recon_to_be_completed></reconnaissance_timing><to_be_mined><mining_i
ndicator>YES</mining_indicator></to_be_mined><is_mined><is_mined>NO</is_mined></is_mined><to_be_ma
rked><to_be_marked>YES</to_be_marked></to_be_marked><is_marked><is_marked>NO</is_marked></is_ma
rked><bridge_identification><bridge_identification_military>TELSRA
1</bridge_identification_military><bridge_identification_civilian>SYDNEY
HARBOUR</bridge_identification_civilian></bridge_identification><name_of_waterway_or_feature><waterway_
or_feature_name>DERWENT
RIVER</waterway_or_feature_name></name_of_waterway_or_feature><bridge_function><road_type>A</road
_type></bridge_function><name_of_route_to_task_site_road_or_river_segment><name_of_route_to_task_site
_road_or_river><name_of_route>BONNIE
CAUSEWAY</name_of_route></name_of_route_to_task_site_road_or_river><route_start_point><grid_referenc
e_utm><here>16ADA1234</here></grid_reference_utm></route_start_point><route_intermediate_point><inte
rmediate_point_number>1</intermediate_point_number><grid_reference_utm><location_utm_1000_meter><u
tm_grid_zone_column>10</utm_grid_zone_column><utm_grid_zone_row>W</utm_grid_zone_row><_100_000
_meter_square_column>D</_100_000_meter_square_column><_100_000_meter_square_row>A</_100_000_m
eter_square_row><utm_1000_meter_easting>12</utm_1000_meter_easting><utm_1000_meter_northing>34</
utm_1000_meter_northing></location_utm_1000_meter></grid_reference_utm></route_intermediate_point><
route_release_point><grid_reference_utm><location_utm_1000_meter><utm_grid_zone_column>12</utm_grid
_zone_column><utm_grid_zone_row>W</utm_grid_zone_row><_100_000_meter_square_column>D</_100_00
0_meter_square_column><_100_000_meter_square_row>A</_100_000_meter_square_row><utm_1000_meter
_easting>12</utm_1000_meter_easting><utm_1000_meter_northing>34</utm_1000_meter_northing></locati
n_utm_1000_meter></grid_reference_utm></route_release_point></name_of_route_to_task_site_road_or_riv
er_segment><equipments_data_segment><equipments_data><group_of_fields><equipment_type>BOLT
CUTTERS</equipment_type><equipment_allocation>25</equipment_allocation><quantity_required>28</quanti
ty_required><duration_required_in_hours>345</duration_required_in_hours></group_of_fields><group_of_fiel
ds><equipment_type>GOGGLES</equipment_type><equipment_allocation>20</equipment_allocation><quantit
y_required>30</quantity_required><duration_required_in_hours>456</duration_required_in_hours></group_o
f_fields></equipments_data></equipments_data_segment><equipments_data_segment><material_data><grou
p_of_fields><type_of_material>CANVAS</type_of_material><quantity_authorized_with_unit_of_measure><qua
ntity_authorized>123</quantity_authorized><unit_of_measurement_combat_engineers>FT</unit_of_measur
ement_combat_engineers></quantity_authorized_with_unit_of_measure><quantity_required_and_unit_of_meas
urement><quantity_required>250</quantity_required><unit_of_measurement_combat_engineers>FT</unit_of_
measurement_combat_engineers></quantity_required_and_unit_of_measurement></group_of_fields></materi
al_data><organization_designator_1><unit_designation_name>1-
503IN</unit_designation_name><unit_size_indicator>BN</unit_size_indicator><country_of_the_world>AS</cou
ntry_of_the_world><unit_role_indicator_code_a>INF</unit_role_indicator_code_a><unit_role_indicator_code_
b>RECCE</unit_role_indicator_code_b><unit_role_indicator_code_c>ABN</unit_role_indicator_code_c><unit_r
ole_indicator_code_d
xsi:nil="true"/><higher_formation_name>2BDE</higher_formation_name><organization_code><armed_service>
A</armed_service></organization_code><unit_identification_code_uic><nationality>AS</nationality><armed_se
rvice>F</armed_service><file_sequential_location_number>1393</file_sequential_location_number></unit_ide
ntification_code_uic></organization_designator_1><rendezvous_details_1><rendezvous_location><rendezvous_
position_utm_10_meter><utm_grid_zone_column>32</utm_grid_zone_column><utm_grid_zone_row>S</utm_
```

```

grid_zone_row><_100_000_meter_square_column>T</_100_000_meter_square_column><_100_000_meter_sq
uare_row>H</_100_000_meter_square_row><utm_10_meter_easting>1150</utm_10_meter_easting><utm_10
_meter_northing>1410</utm_10_meter_northing></rendezvous_position_utm_10_meter></rendezvous_locati
on><rendezvous_dtg><day>4</day><hour_time>14</hour_time><minute_time>5</minute_time><time_zone>Z
</time_zone><month_name>APR</month_name><year>2003</year></rendezvous_dtg></rendezvous_details_
1></equipments_data_segment><organisation_function_segment><organisation_function><organization_functi
on>PROTECT</organization_function></organisation_function><organization_designator_2><unit_designation_n
ame>1-
503IN</unit_designation_name><unit_size_indicator>BN</unit_size_indicator><country_of_the_world>US</cou
ntry_of_the_world><unit_role_indicator_code_a>INF</unit_role_indicator_code_a><unit_role_indicator_code_
b>RECCE</unit_role_indicator_code_b><unit_role_indicator_code_c>ABN</unit_role_indicator_code_c><unit_r
ole_indicator_code_d
xsi:nil="true"/><higher_formation_name>2BDE</higher_formation_name><organization_code><armed_service>
A</armed_service></organization_code><unit_identification_code_uic><nationality>US</nationality><armed_se
rvice>A</armed_service><file_sequential_location_number>1393</file_sequential_location_number></unit_ide
ntification_code_uic></organization_designator_2><communication_radio_frequency_and_call_sign><call_sign_
and_frequency><specific_call_sign>BADBIRD</specific_call_sign><mandatory_space>
</mandatory_space><context_quantity_one_decimal_pt_permitted_8_char_max>987.4</context_quantity_one
_decimal_pt_permitted_8_char_max><radio_frequency_unit_of_measurement>MHZ</radio_frequency_unit_of
_measurement></call_sign_and_frequency><own_call_sign_on_all_arms_net>31C</own_call_sign_on_all_arms
_net></communication_radio_frequency_and_call_sign><rendezvous_details_2><rendezvous_location><rendez
vous_position_utm_10_meter><utm_grid_zone_column>32</utm_grid_zone_column><utm_grid_zone_row>S</
utm_grid_zone_row><_100_000_meter_square_column>T</_100_000_meter_square_column><_100_000_met
er_square_row>H</_100_000_meter_square_row><utm_10_meter_easting>1150</utm_10_meter_easting><utm
_10_meter_northing>1410</utm_10_meter_northing></rendezvous_position_utm_10_meter></rendezvous_l
ocation><rendezvous_dtg><day>4</day><hour_time>14</hour_time><minute_time>5</minute_time><time_zo
ne>Z</time_zone><month_name>APR</month_name><year>2003</year></rendezvous_dtg></rendezvous_det
ails_2></organisation_function_segment><acknowledgement_requirement><acknowledgement_requirement_in
dicator>YES</acknowledgement_requirement_indicator><group_of_fields><instructions_or_unit><instructions_f
or_acknowledging>AOC EXT
2345</instructions_for_acknowledging></instructions_or_unit></group_of_fields></acknowledgement_require
ment></mtf:national_bridge_demolition_reconnaissance_order>

```

Figure 7 – Example 4 BRDMLRECCEORD.xml XML Instance.

Example 5 (Figure 8) provides an uncompressed EXI encoding instance of the XML instance of the ASMTF/brdmlrecceord/BRDMLRECCEORD.xml XML document. Please note that the following data has been inserted below as an image to maintain encoding information represented in the EXI encoding.

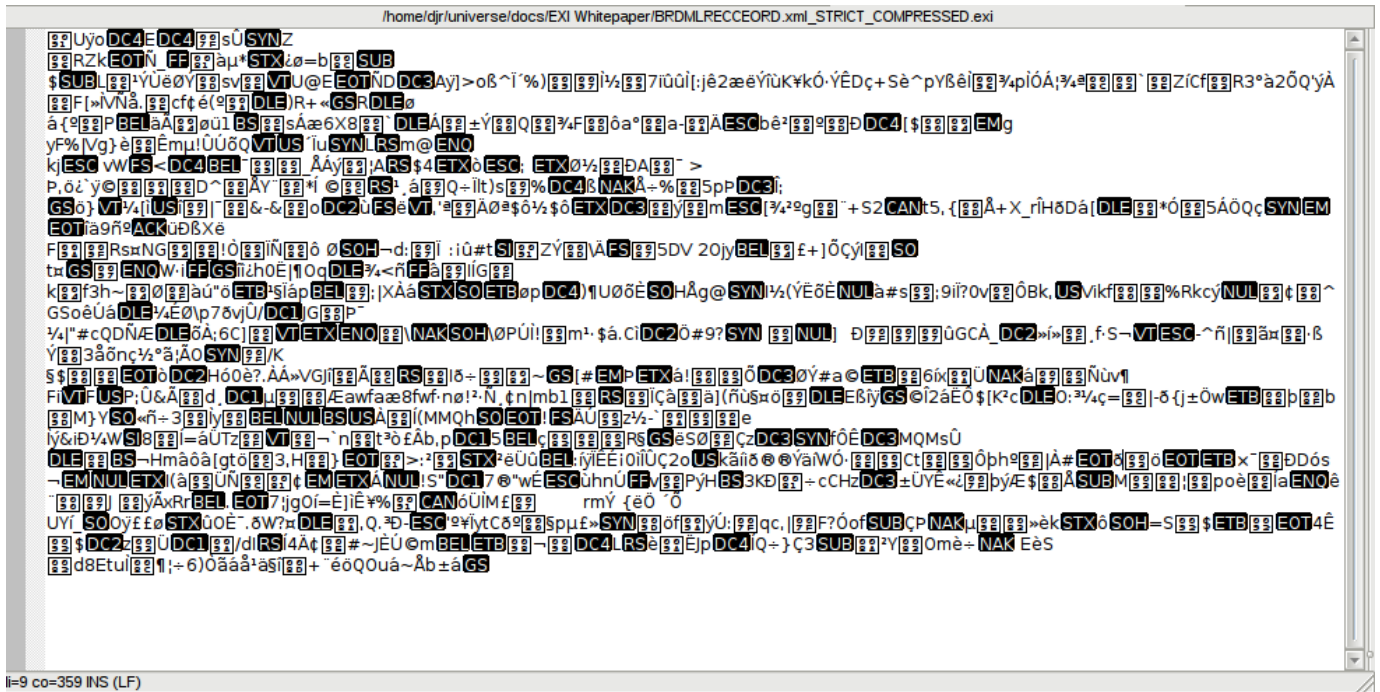


Figure 9 – Example 6 BRDMLRECCEORD Compressed EXI Instance.

Example 5 and 6 (shown in Figures 8 and 9) both illustrate encoding and compression. Even the uncompressed EXI document provides compression (elimination of redundant or unneeded data within the document) over the XML instance; however, it is clear to see that the document listed in Example 6 (Figure 9) is much smaller in size due to EXI's internal compression.

For more information and examples on EXI see the [MOVES Institute's](#) website or the [W3C's EXI Primer](#) and [EXI 1.0 Encoding Examples](#).

ABOUT THE IJIS INSTITUTE

The [IJIS Institute](#), a 501(c)(3) nonprofit corporation, represents industry's leading companies who collaborate with local, state, tribal, and federal agencies to provide technical assistance, training, and support services for information exchange and technology initiatives. Serving as the voice of industry, the IJIS Institute unites the private and public sectors to improve mission-critical information sharing for those who protect and serve our communities.

The IJIS Institute was founded in 2001 as a result of the [U.S. Department of Justice's](#) interest in raising private sector participation in the advancement of national initiatives affecting justice and public safety, and more recently homeland security. Today, the IJIS Institute represents the [leading companies](#) serving these and other related sectors. The IJIS Institute provides assistance to government agencies by bringing industry to the table in a constructive role, and continuing to drive toward achieving high regard for the companies that are dedicated to helping the public sector find high value solutions. The IJIS Institute is funded through a combination of federal grants, industry contributions, and partnership agreements.

The IJIS Institute does its valuable work through the contributions of its member companies. The IJIS Institute thanks Technical Advisory Committee for their work on this document.

The IJIS Institute also thanks the many companies who have joined as members that contribute to the work of the Institute and share in the commitment to improving justice, public safety, and homeland security information sharing.*

[Abstractian Group, The](#)
[Accelerated Information Management](#)
[Accenture](#)
[Advanced Justice Systems](#)
[Advanced Technology Systems \(ATS\)](#)
[AmCad Court Case Management, LLC](#)
[Amicus Group, The](#)
[Analysts International](#)
[Appriss, Inc.](#)
[ARC Consulting](#)
[Archer Group, The](#)
[Avaya Government Solutions](#)
[BAE Systems](#)
[Bask Enterprises](#)
[BIO-key International, Inc.](#)
[BlueStreak Connect](#)
[Booz Allen Hamilton](#)
[BruckEdwards, Inc.](#)
[Cardinal Point Partners, LLC](#)
[Cisco Systems, Inc.](#)
[Citizant](#)
[CivicUS](#)
[CommSys Incorporated](#)
[Computer Projects of Illinois, Inc.](#)
[Core Technology Corporation](#)
[CourtView Justice Solutions](#)
[DaPro Systems, Inc.](#)
[Datamaxx Applied Technologies, Inc.](#)
[Deloitte Consulting](#)
[Development Services Group](#)
[DigitalBridge](#)
[Eadie Consulting](#)
[Eastport Analytics, Inc.](#)
[Emergitech](#)

[ERDG](#)
[ESRI](#)
[Harbitter Consulting Group](#)
[HCW Consulting](#)
[Health Information Designs](#)
[Hewlett Packard Company](#)
[Hitech Systems, Inc.](#)
[HRInterop, LLC](#)
[Hunter Research](#)
[i2 Inc.](#)
[IBM](#)
[ICUC, Inc.](#)
[Infinite Solutions Inc.](#)
[Intellitech Corporation](#)
[Interacx](#)
[Intergraph Public Safety](#)
[Interop-Solutions, Inc.](#)
[iNovate Solutions](#)
[InTime Solutions, Inc.](#)
[IxReveal, Inc.](#)
[Justice Served](#)
[Justice Systems, Inc.](#)
[Justice Technology Consulting](#)
[Keane Federal Systems](#)
[Marquis Software Development, Inc.](#)
[Memex, Inc.](#)
[Metastorm, Inc.](#)
[Metatomix](#)
[Microsoft Corporation](#)
[Motorola](#)
[MTG Management Consultants, LLC](#)
[New Dawn Technologies](#)
[Niche Technology](#)
[Nixle](#)
[NOWHERETOHide.ORG](#)
[Nuance Communications](#)
[Online Business Systems](#)
[Open Networks](#)
[Optimum Technology](#)
[Patriot Data Solutions Group \(PDSG\)](#)

[PCSS, Inc.](#)
[Police Central, Inc.](#)
[PNL Associates, LLC](#)
[Presynct Technologies, Inc.](#)
[PST Technologies](#)
[PSTG Consulting](#)
[Public Engines](#)
[Public Safety Consulting, Inc.](#)
[RADCA Co., LLC](#)
[Raytheon Company](#)
[RCC Consultants, Inc.](#)
[Readiness Resource Group Inc.](#)
[SAL Consulting LLC](#)
[SituCon Systems](#)
[Soos Creek Consulting, LLC](#)
[Soteria Systems, LLC](#)
[Spillman Technologies](#)
[SRA International, Inc.](#)
[Sypherlink, Inc.](#)
[Syscon Justice Systems, Ltd.](#)
[Tetrus Consulting Group](#)
[Tiburon](#)
[Tom Carlson Consulting](#)
[Total Computer Group](#)
[TriTech Software Systems](#)
[Trusted Federal Systems, Inc.](#)
[Unisys](#)
[URLIntegration](#)
[Verint Technology, Inc.](#)
[Versaterm Corporation](#)
[VisionAIR](#)
[Visual Analytics, Inc.](#)
[Vortex](#)
[Watch Systems, LLC](#)
[Waterhole Software](#)
[Whys Solutions, LLC](#)
[Yellow House Associates, LLC](#)

**as of January 2010*