

# Data Exchange Standard – Character Set

Office of the Corporate Chief Information Officer, Strategy and Governance Branch

Version: 1.0

<b>Approved by:</b> Corporate Chief Information Officer	<b>Owner:</b> IMT Policy Program
<b>Approval date:</b> 2014-12-31	<b>Review date:</b> 2021-01-01
<b>Contact:</b> <a href="mailto:IMT.Policy@gov.ab.ca">IMT.Policy@gov.ab.ca</a>	<b>Policy Instrument type:</b> Technical Standard

## Contents

Contents.....	2
Standard Statement.....	3
Authority .....	3
Scope .....	3
Standard Specification.....	3
References and Supporting Resources .....	4
Appendix A.....	5

## Standard Statement

Data exchange standards are required to ensure the clarity and the accuracy of data exchanged between Government of Alberta (GoA) applications.

The standard below describes the GoA data standard for character sets shared between GoA applications and applies to exchange of information gathered by GoA applications.

## Authority

Internal use only.

## Scope

This standard applies to all Ministries within the GoA.

## Standard Specification

Character Set			
<b>Description</b>	Universal character set that represents every character in the Unicode Character Set (UCS).		
<b>Format</b>			
	UTF-8		<p>Unicode Transformation Format 8-bit is a variable-width encoding that can represent every character in the Unicode character set.</p> <p>UTF-8 has the characteristic of preserving the full US-ASCII range, providing compatibility with file systems, parsers, and software that rely on US-ASCII value but are transparent to other values.</p> <p>UTF-8 characters from the U+0000..U+10FFFF range (the UTF-16 accessible range) are encoded using sequences of 1 to 4 octets.</p> <p>UTF-8 has a one-octet encoding unit and encodes characters as a varying number of octets, where the number of octets and the value of each depend on the integer value assigned to the character. The encoding form characteristics are described in RFC 3629.</p>

## References and Supporting Resources

- UTF-8: <http://www.utf-8.com/>
- RFC 3629, UTF-8 A transformation format of ISO 10646: <https://tools.ietf.org/html/rfc3629>
- ISO/IEC 10646, Information Technology – Universal Coded Character Set (UCS): <http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html>

## Appendix A

Types of Standards	Description
<b>Technical Standard</b>	These are detailed, unique standards that have developed in response to government IMT policies. Technical standards are intended to be replicable, transferable, and adaptable across ministries and other government agencies. Examples of these could include address data standards or specifications for a single identifier for transacting with government electronically.
<b>Product Standard</b>	An IMT product or specific technology oriented standard that facilitates the task of planning for enhancements and acquisitions within the government's broad information systems environment. As a definitive list of the numerous technologies either employed or under evaluation by Workplace Technology Services, product standards are critical in establishing conformity, interoperability and interchange-ability. Examples of these could include a government-wide standard for document, record management and database, and the list of core products for government workstations.
<b>Process Standard</b>	An established, mandatory business practice that supports IMT projects and existing systems to improve the outcome, diminish risks, and increase reliability. Examples could include business continuity planning processes, threat risk assessment processes, etc.
<b>Reference Standard</b>	An IMT industry standard (either a national or international formal or de facto standard) that has been adopted for use by the Province of Alberta. A Reference Standard may be adopted either as stand-alone or as a precursor to a customized standard or policy document. Examples could include the 1024 bit RSA standard for public key encryption.