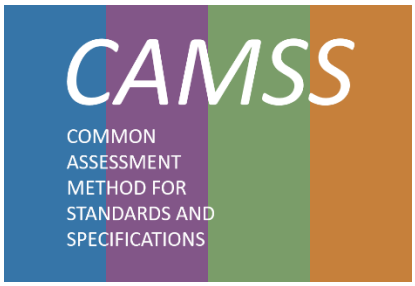


ASSESSMENT SUMMARY

Unicode Transformation Format - 8

Internet Engineering Task Force (IETF)

1. INTRODUCTION

The present document is a summary of the assessment of UTF-8 carried out by the CAMSS Team using the CAMSS EIF assessment scenario. The purpose of this scenario is assessing the compliance of a standard or technical specification with the European Interoperability Framework (EIF)[1].

2. ASSESSMENT SUMMARY

Unicode Transformation Format - 8 (UTF-8) is a variable width encoding character defined by the Unicode Standard. The UTF-8 has become one of the most used for the World Wide Web, being present in more than 90% of all web pages.

2.1. Interoperability Principles

Interoperability principles are fundamental behavioural aspects that drive interoperability actions. They are relevant to the process of establishing interoperable European public services. They describe the context in which European public services are designed and implemented.

The specification fully supports the principles setting context for EU actions on interoperability:

- **Subsidiarity and proportionality**
  UFT-8 is included in 2 national catalogues of recommended specifications. They belong to the Netherlands and Sweden. The National Interoperability Framework (NIF) of these Member States is fully aligned with at least 4 out of 5 sections of the European Interoperability Framework (EIF)[2] according to the National Interoperability Framework Observatory (NIFO)[3] factsheets.

The specification partially supports the principles setting context for EU actions on interoperability:

- **Openness**
  RDF N-Triples is a lined-based plain text format that performs encoding RDF graphs, this encoding is carried out by using UTF-8. As RDF is the main technology used to publish Linked Open Data (LOD). However, N-Triples presents a tedious process if the purpose is encode big amounts of data.

  Moreover, during the development of a specification, draft versions of the document are made available for informal review and comment making an evolving working document readily available to a wide audience, facilitating the process of review and revision. In addition, UTF-8 is publicly available for implementation and use, and therefore for everyone to study, for free on IETF's webpage. Like all the IETF standards, UTF-8 is a free and open technical specification, built

[1] https://ec.europa.eu/isa2/eif_en

[2] https://ec.europa.eu/isa2/sites/isa/files/eif_brochure_final.pdf

[3] https://joinup.ec.europa.eu/collection/national-interoperability-framework-observatory-nifo/nifo-factsheets

on IETF standards and licenses from the Open Web Foundation. Finally, UTF-8 is widely adopted in environments developing and implementing ICT solutions including innovative solutions and maintained by IETF.

- **Transparency**
  By allowing communications over the internet, the UTF-8 fosters the visibility and comprehensibility of administrative rules, processes, data, services and decision-making of public administrations. In addition this specification ensures the availability of interfaces with internal information systems of a public administration.

- **Reusability**
  UTF-8 is an open specification that is available for free and published in collaborative platforms for the reuse of solutions. Additionally, it is a sector agnostic specification.

- **Technological neutrality and data portability**
  The UTF-8 is independent from any specific technology and/or platform and fosters data portability between systems and applications.

Technical Specification partially supports the principles related to generic user needs and expectations:

- **User-centricity**
  The purpose of UTF-8 is not related to the implementation of the once-only principle. Therefore this criterion is not considered.

- **Inclusion and accessibility**
  The purpose of UTF-8 is not related to e-accessibility. Therefore this criterion will not be considered.

- **Security and privacy**
  UTF-8 is an element with high relevance in the data and information exchange. Moreover, it can be read by both parts implied in the communication ensuring the reliability of transferred data. Consequently, UTF-8 eases the trustworthy data exchange between administration and stakeholders.

- **Multilingualism**
  UTF-8 supports many languages and can accommodate pages and forms in different cases including a mixture of these languages.

The Technical Specification partially supports the foundation principles for cooperation among public administrations:

- **Administrative Simplification**

UTF-8 eases the digital data exchange between administration and stakeholders, therefore, it is a mean for the reduction of administrative burden.

- **Preservation of information**
  The purpose of UTF-8 is not related to long term preservation of electronic records. Therefore this criterion is considered.

- **Assessment of effectiveness and efficiency**
  There are no evidences of existing studies or documentation assessing the standard or specification in terms of effectiveness and efficiency to justify the criterion.

2.2. Interoperability Layers

The interoperability model which is applicable to all digital public services includes:

- Four layers of interoperability: legal, organisational, semantic and technical;
- A cross-cutting component of the four layers, 'integrated public service governance';
- A background layer, 'interoperability governance'.

The Technical Specification supports the implementation of digital public services complying with the EIF interoperability model:

- **Interoperability governance**
  UTF-8 is already associated to an EIRA ABB in the European Library of Specifications (ELIS). More specifically, UTF-8 can define the interoperability aspects of the "Data" ABB of the EIRA Semantic View. UTF-8 have existing validation tools, and is validated by 10 Members States highly used by public administrations.

  Finally, UTF-8 is compliant with the regulation on standardisation 1025/2012, UTF-8 has been identified by Commission Implementing Decision and included in the European list of ICT Standards for e-procurement.

- **Integrated public service governance & Legal Interoperability**
  After being evaluated compliant with the regulation on standardisation 1025/2012, UTF-8 has been identified by Commission Implementing Decision. During the evaluation process, all the Member States are invited to share their doubts. The positive evaluation of UTF-8 and its identification is considered an interoperability agreement.

- **Organizational interoperability**
  UTF-8 is not a business process modelling standard or specification and does not define organisational interoperability aspect. The purpose of the specification is not related to organisational Interoperability.

- **Semantic interoperability**

UTF-8 is not defining a cross-sector reusable data model. The purpose of the specification is to define a schema for encoding data.

- **Technical interoperability**
  UTF-8 is an open specification available for everyone for study or use.

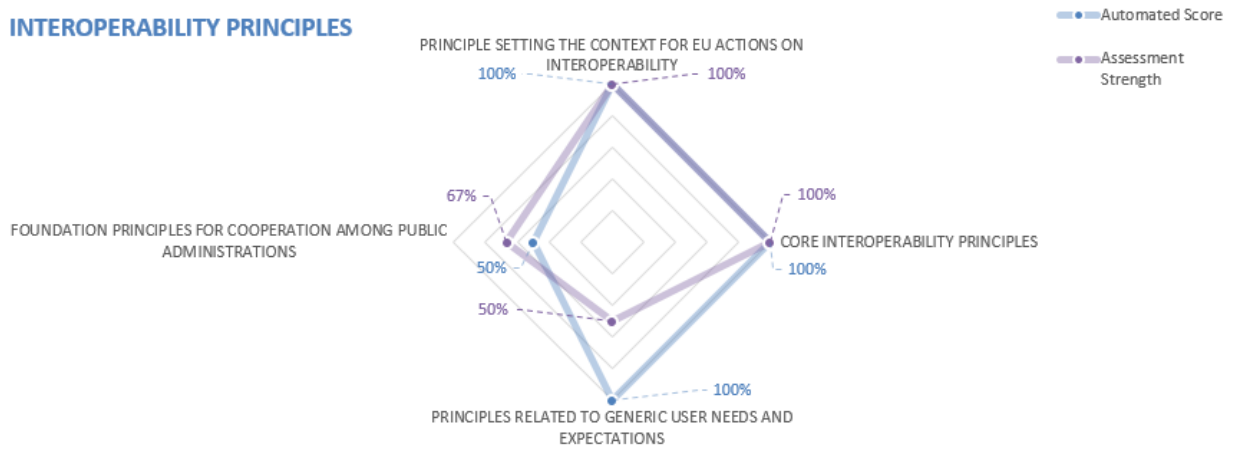
3. *ASSESSMENT RESULTS*

This section presents an overview of the results of the CAMSS assessments of UTF-8. The Assessment "Strength" indicator measures the reliability of the assessment by calculating the number of applicable criteria. On the other hand, the number of favourable answers and the number of unfavourable ones are used to calculate the "Automated Score" per categories.

Category	Automated Score	Assessment Strength	Favourable	Unfavourable	Not Applicable
Principle setting the context for EU actions on interoperability	100%	100%	1	0	0
Core interoperability principles	100%	100%	15	0	1
Principles related to generic user needs and expectations	100%	50%	2	0	2
Foundation principles for cooperation among public administrations	50%	67%	2	0	1
Interoperability layers	95%	91%	18	0	4
Overall Score	**94%**	**86%**	**29**	**0**	**8**

**The technical interoperability layer is covered by the criteria corresponding to the core interoperability principle ''Openness''.*

With an 86% of assessment strength, this assessment can be considered representative of the specification compliance with the EIF principles and recommendations.



The Overall Automated Score of 94% demonstrates that UTF-8 highly supports the European Interoperability Framework in the domains where it applies.

Figure 1 Assessment Results – Interoperability Principles

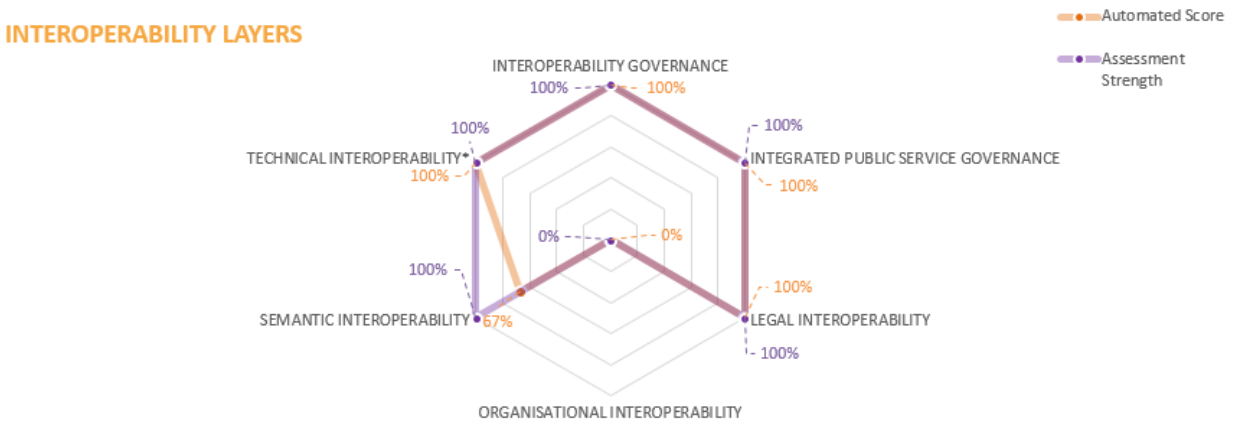


Figure 2 Assessment Results - Interoperability Layers