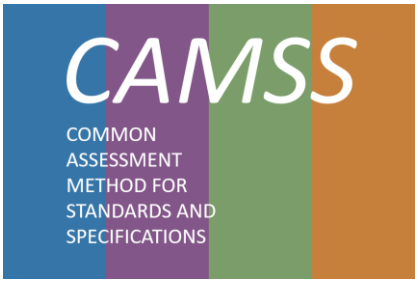

ASSESSMENT SUMMARY v1.0.0

Shape Expressions (ShEx)[1]

World Wide Web Consortium (W3C)[2]

[1] Shape Expressions specification: Shape Expressions (ShEx) 2.1 Primer

[2]W3C website: https://www.w3.org/

Change Control

Modification	Details
Version 1.0.0	
Initial version	

TABLE OF CONTENT

1. INTRODUCTION

The present document is a summary of the assessment of the **Shape Expressions (ShEx)** carried out by CAMSS using the CAMSS Assessment EIF scenario[3]. The purpose of this scenario is to assess the compliance of a standard or specification with the European Interoperability Framework (EIF)[4].

2. ASSESSMENT SUMMARY

Shape Expressions (ShEx) is a language for describing RDF graph structures. A ShEx schema prescribes conditions that RDF data graphs must meet to be considered "conformant": which subjects, predicates and objects may appear in a given graph, in what combinations and with what cardinalities and datatypes. In the ShEx model, an RDF graph is tested against a ShEx schema to yield a validation result that flags any parts of the data which do not conform.
ShEx schemas are intended for use in validating RDF data, communicating interface parameters and data structures, generating user interfaces and transforming RDF graphs into other data formats and structures.

The specification has been developed by World Wide Web Consortium (W3C), which is an international community concerned with evolving the World Wide Web by developing protocols and guidelines to ensure and enhance its growth.

2.1. EIF 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 does not support the principles setting context for EU actions on interoperability*:**

- **Subsidiarity and proportionality**
  No Member State includes the ShEx in their national catalogue with Their National Interoperability Framework (NIF) in alignment with the three categories 1. Conceptual model for integrated public services provision, 2. interoperability layers, and 3. interoperability principles.

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

- **Openness**
  Given that the main uses of ShEx are validating instance data and transforming RDF graphs into other data formats and structures, it can be a good enabler for the publishing of open data, since it is a tool through which to assure is well structured to be published as open data.

[3] CAMSS Assessment EIF Scenario 5.1.0: EUSurvey - Survey (europa.eu)

[4] ISA[2] programme: https://ec.europa.eu/isa2/eif_en

The development process has been carried out by W3C[5] to make it accessible to the different stakeholders and it also includes a public review. Moreover, the Shape Expressions Community Group[6] is the developer community that maintains this specification. It is important to mention that the language expressed by ShEx is stable, and there can be found documentation of its supporting processes on GitHub[7] and the ShEx Community Group webpage. Although there are not many market uptake indicators, ShEx has been proven efficient and has support from interest groups that are involved in the development of cross-border initiatives. It is also publicly available for free on W3C's webpage, and It is licensed on a royalty-free basis for its implementation or study.

- **Transparency**
  The Shape Expressions (ShEx) language provides a structural schema for RDF data. This can be used to document APIs or datasets, aid in the development of API-conformant messages, minimize defensive programming, guide user interfaces, or anything else that involves a machine-readable description of data organization and typing requirements. This can help to the exposure of interfaces to access public administrations' services and to foster the comprehensibility of administrative procedures.

- **Reusability**
  ShEx is built around RDF and is highly dependent on it since its purpose is to validate RDF graphs. Nonetheless, it has many known implementations that range through a variety of business domains allowing for RDF graphs validation across business domains.

- **Technological neutrality and data portability**
  ShEx is independent of any specification and does not rely on any technology or platform. Moreover, extension mechanisms are provided by external shapes, often used to describe very large value sets. Regarding data portability, ShEx can be used for serialization in a variety of human and machine-readable formats, as well as it allows for the communication of data structures associated with some process or interface, thus, facilitating the transfer of information.

***The specification partially supports the principles related to generic user needs and expectations*:**

- **User-centricity**
  ShEx has been designed to enhance the semantic and technical interoperability layers of ontologies expressed as RDF graphs and, in turn, RDF is a European Standard used by ùblic administrations in projects for sharing linked data. In this sense, ShEx allows for the reuse of

[5] W3C Process Document: https://www.w3.org/2018/Process-20180201/#Policie

[6] ShEx Community Group: Shape Expressions Community Group (w3.org)

[7] ShEx in GitHub: GitHub - shexSpec/spec: ShEx specification

information, although not in a complete manner, insofar as it gives a tool to enhance interoperability and thus, implement the Once-only principle[8] but it is not its specific purpose.

- **Inclusion and accessibility**
  The purpose of ShEx is not related to e-accessibility. Therefore, this criterion is considered not applicable to this specification.

- **Security and privacy**
  Understanding data exchange as the publication by the administration of data for its later consumption by citizens, the specification facilitates trustworthy data exchange and processing in RDF format by providing the means to validate RDF data graphs.

- **Multilingualism**
  The purpose of ShEx is not related to multilingualism; therefore, this criterion does not apply to the specification.

The specification supports the foundation principles for cooperation among public administrations:

- **Administrative Simplification**
  ShEx can simplify the delivery of those European public services related to the semantic web. While RDF data graphs can be validated more easily with the specification's functionality, interoperability is triggered given that ShEx reduces possible syntax errors, and thus, simplifies the delivery of European public services. Moreover, Digital service delivery channels like those requiring Linked data, or semantic technologies may be enhanced by the use of ShEx, given that it allows for the validation and modelling of RDF structures as well as a better understanding of data representation details

- **Preservation of information**
  The purpose of ShEx is not related to the long-term preservation of information; therefore, this criterion does not apply to this specification.

- **Assessment of effectiveness and efficiency**
  There are several studies assessing ShEx effectiveness and efficiency when it comes to modelling and validating RDF graphs. An interesting study assesses the effectiveness of ShEx when describing and validating FHIR[9] (Fast Healthcare Interoperability Resources) RDF data; pointing out the benefits of using ShEx RDF validator.

[8] Once Only principle: Once Only Technical System (europa.eu)

[9] Modeling and validating HL7 FHIR profiles using semantic web Shape Expressions (ShEx): Modeling and validating HL7 FHIR profiles using semantic web Shape Expressions (ShEx) - ScienceDirect

2.2. EIF 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 Specification supports the implementation of digital public services complying with the EIF interoperability model:

- **Interoperability governance**
  ShEx can be mapped into the EIRA Library of Interoperability Specifications (ELIS) Technical Application view, as part of the Data Quality Component, Data Quality Service, and Machine to Machine ABBs, and in the Semantic View, as part of the Semantic Interoperability Agreement ABB. Although the ShEx is not included in any catalogue, neither at the national nor European level, it is nonetheless contributing to the development of the cross-border European Union Knowledge Graph[10] project. ShEx is itself a tool to assess the conformity of RDF data graphs against correspondent schemas. Moreover, W3C provides the ShEx Simple Online Validator[11], an online tool that allows performing RDF data validation using Shape Expressions.

- **Legal Interoperability**
  ShEx is not a European Standard, and nor is yet in the W3C standards track.

- **Organisational interoperability**
  ShEx can be used in model development, as well as for the documentation of data models to the extent that it provides a human-readable representation that helps developers to better understand the model and its semantics. Organisational Interoperability agreements also can be achieved by ShEx as is the case of the Hercules ASIO project[12], which uses ShEx to enhance interoperability between university knowledge technologies.

- **Semantic Interoperability**
  ShEx specification and schemas are publicly available for implementation and use on Github. The GitHub repository hosts a community of developers involved with ShEx, where they can share information, raise issues and propose solutions for the ShEx development.

[10] EU Knowledge Graph: The EU Knowledge Graph - EU Knowledge Graph (linkedopendata.eu)

[11] ShEx Simple Online Validator: ShEx2 — Simple Online Validator (rawgit.com)

[12] Hercules ASIO project: Semantic Architecture and Ontological Infrastructure - Hércules (um.es)

3. ASSESSMENT RESULTS

This section presents an overview of the results of the CAMSS assessments for **ShEx**. The CAMSS "Strength" indicator measures the reliability of the assessment by calculating the number of answered (applicable) criteria. On the other hand, the number of favourable answers and the number of unfavourable ones is used to calculate the "Automated Score" per category and an "Overall Score".

Category	Automated Score	Assessment Strength	Compliance Level
Principle setting the context for EU actions on interoperability	20/100 (20%)	100%	Ad-hoc
Core interoperability principles	2140/2200 (97%)	91%	Seamless
Principles related to generic user needs and expectations	420/500 (84%)	60%	Seamless
Foundation principles for cooperation among public administrations	500/500) (100%)	80%	Seamless
Interoperability layers*	660/1100 (60%)	100%	Sustainable
Overall Score	3240/3900 (83%)[13]	87%	

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

With an 87% of assessment strength, this assessment can be considered representative of the specification compliance with the EIF principles and recommendations.

The Overall Automated Score of 83% (3240/3900) demonstrates that the specification supports the European Interoperability Framework in the domains where it applies.

[13] See the "results interpretation" section of the CAMSS Assessment EIF Scenario Quick User Guide:

https://joinup.ec.europa.eu/collection/common-assessment-method-standards-and-specifications-camss/solution/camss-assessment-eif-scenario/results-visualisation-and-interpretation