

EFRO: An Ontology for Education Funding Regulations

Umair Arshad*, David Corsar and Ikechukwu Nkisi-Orji

School of Computing, Engineering and Technology, Robert Gordon University, Aberdeen AB10 7GJ, Scotland, United Kingdom

Abstract

Ensuring compliance with education funding rules is a resource-intensive task for education providers. We introduce the Education Funding Regulations Ontology (EFRO) - a formal vocabulary designed to represent knowledge related to education funding. By modelling this domain, EFRO aims to serve as a foundational knowledge source for intelligent decision support systems aimed at assisting education providers. The ontology is validated through competency questions derived from official funding guidance and demonstrated using an early progression and funding retention use case. All resources are available at GitHub.

Keywords

Ontology, Educational Funding Regulations, Compliance Management, Knowledge Representation, SPARQL

1. Introduction

The financial ability to provide or access education is dependent on providers and/or students evidencing their compliance with the regulations that govern funding allocation. These regulations are frequently updated and defined as lengthy unstructured texts, requiring continual investment by education providers to ensure their ongoing compliance. Current compliance processes rely heavily on manual interpretation of guidance, making consistent decision support and auditability difficult at scale. We introduce the Education Funding Regulations Ontology¹ (EFRO), which provides a formal, structured vocabulary for representing funding regulations and associated knowledge. EFRO is intended to provide a foundational knowledge source for intelligent decision support systems providing, for example, automated checks on compliance with funding regulations, student/course funding eligibility, and evidence suitability.

EFRO was developed using the METHONTOLOGY framework [1]. The specification phase derived a set of 16 competency questions defining the ontology's purpose and scope based on frequently asked questions of existing education funding regulations, including [2, 3]. Knowledge Acquisition subsequently involved manual analysis of regulations, creating a glossary of core domain concepts, which was expanded during conceptualisation to document entities covering education programmes and stakeholders, funding, and obligations, along with their interrelationships. While the regulations used are specific to the UK, the upper-level concepts (shown in Figure 1) in the defined hierarchies are abstract and should be applicable to other countries. During the integration phase, existing vocabularies were specialised where their native classes did not capture regulation-specific constraints: for example, Academic Institution Internal Structure Ontology (AIISO)² Programme class was specialised through `efro:StudyProgramme` to attach properties such as `regulatedBy` and `qualifying-period` constraints, which are required to represent funding compliance rules but are absent from AIISO itself. The ontology was implemented using Protégé³ and OWL-2, and evaluated using the competency questions encoded as SPARQL queries.

K-CAP 2025 Companion Volume

*Corresponding author.

✉ u.arshad1@rgu.ac.uk (U. Arshad); d.corsar1@rgu.ac.uk (D. Corsar); i.nkisi-orji@rgu.ac.uk (I. Nkisi-Orji)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

¹<http://www.w3id.org/efro/efro>

²<http://purl.org/vocab/aiiso/schema>

³<https://protege.stanford.edu/>

2. EFRO Schema

EFRO follows FAIR principles [4] and it is openly available¹. The EFRO comprises 187 classes, 80 object properties, and 55 data properties. All concepts are documented with human-readable labels and comments, and published with stable IRIs to facilitate reuse. To support interoperability, the EFRO reuses the: Funding, Research Administration and Projects Ontology⁴ modelling constructs for funding bodies and applications; the AIISO² concepts for education providers and courses; and Friend-of-a-Friend⁵ concepts for people and organisations.

Figure 1 shows the EFRO core schema. Subclasses of `efro:EducationalEntity` model organisations and governing bodies that are responsible for delivering funded education provision and for complying with regulatory requirements. These entities are linked to `efro:Obligation` instances, which represent duties derived from funding regulations that providers are required to meet. `efro:Learning` models learning modes and temporal aspects related to that provision, such as planned teaching hours, representing the learning provision delivered by an educational entity and associated with specific programmes. Educational programmes are represented by reusing `aiiso:Programme`, which EFRO specialises through `efro:StudyProgramme` and links to funding regulations via `regulatedBy`. `efro:FundingElement` captures how people, organisations, and education programmes receive funding linked to `efro:TimeRelated` temporal conditions, periods, and important time points for achieving and evidencing compliance. `efro:Person` models staff, students, education providers, auditors, etc. that participate in one or more `efro:Activity`, which define the tasks, processes, and events for administrating and managing education programmes and funding compliance. `efro:Status` defines the state or condition of a learner or programme, with the `efro:Reason` class modelling justifications for changes or actions within some funded provision.

EFRO defines domain-specific subclasses to specialise its core concepts for funding scenarios. Learners are represented using `efro:Student`, progression is modelled through `efro:ProgressionActivity` and its subclasses, and educational provision is captured using `efro:StudyProgramme` and its specialisations such as `efro:Traineeship`. Outcomes are represented using `efro:Status` specialisations such as `efro:CompletedEarlyStatus`, while temporal constraints are captured using `efro:TimeRelated` concepts including `efro:QualifyingPeriod`. These subclasses support the instance-level modelling presented in Section 3.1.

3. Evaluation

Iterative refinement of EFRO was informed by the Ontology Pitfall Scanner (OOPS)⁶, OntoMetrics⁷, Pellet [5] and Hermit [6], which were used to identify and resolve modelling issues and logical consistency and to assess structural and metadata quality characteristics. These tools supported improvements to consistency, clarity, and maintainability throughout the ontology development process.

3.1. Competency Questions

Evaluation further focused on assessing EFRO’s ability to support domain-relevant reasoning by executing SPARQL queries corresponding to the 16 competency questions defined during ontology specification. Briefly, these questions cover: withdrawal scenarios across different learning modes; adjustments of planned teaching hours within qualifying periods; learner status changes from full-time to part-time; early progression from traineeships to employment or further education; temporal constraints such as determining withdrawal dates based on attendance records; expressing funding eligibility rules and checking compliance with those rules for multi-year learning programmes, students completing their programme before term end, and funding retention when students transition between

⁴<http://purl.org/cerif/frapo>

⁵<http://xmlns.com/foaf/0.1/>

⁶<https://oops.linkeddata.es/index.jsp>

⁷<https://ontometrics.informatik.uni-rostock.de/ontologymetrics/>

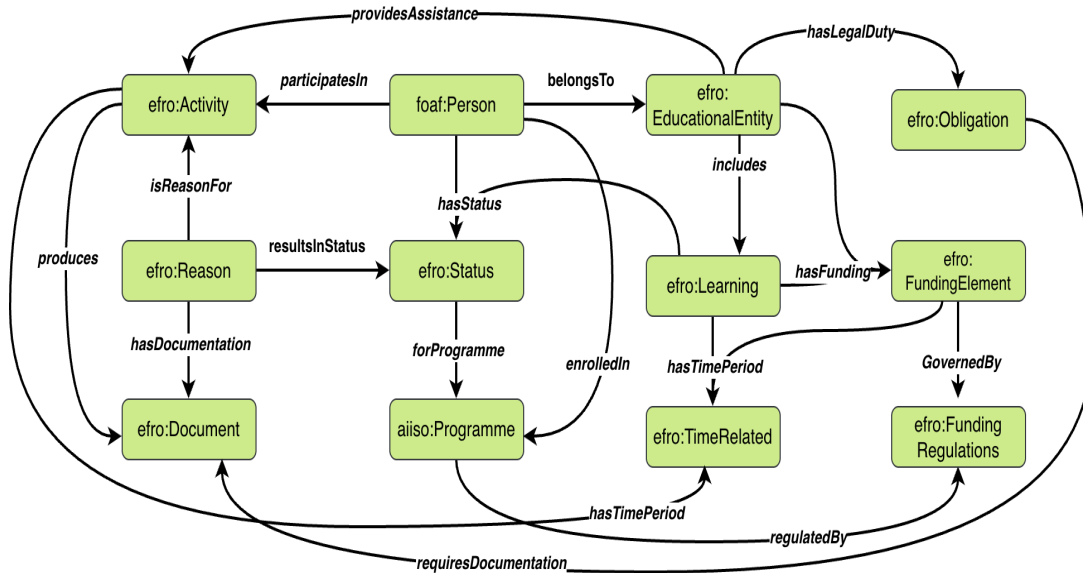


Figure 1: Core EFRO structure; all properties are defined in the efro namespace, <http://www.w3id.org/efro/efro>.

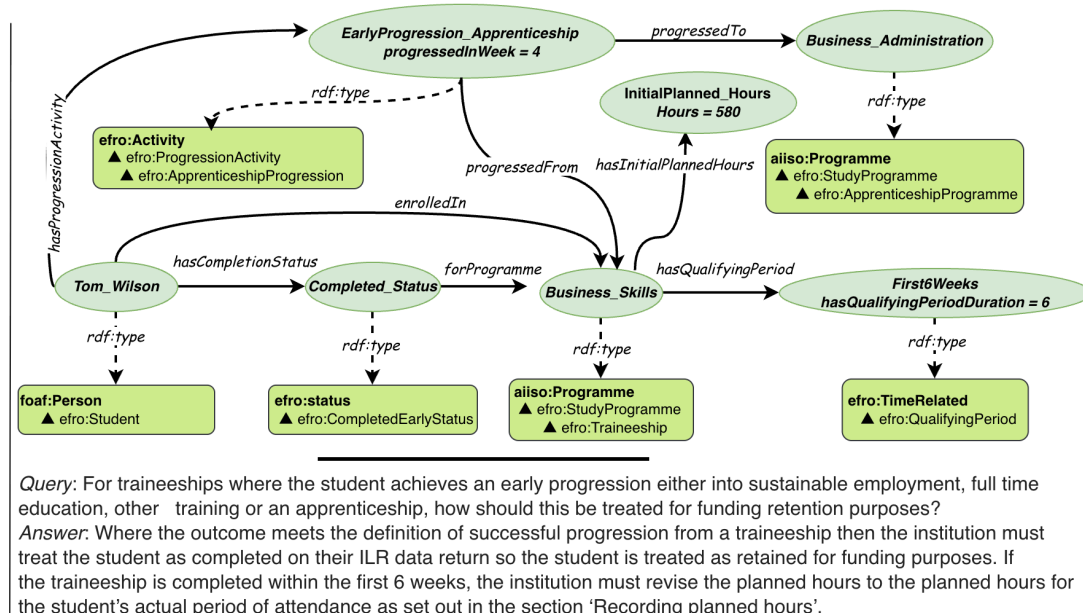


Figure 2: An instance-level EFRO use case illustrating early progression in a traineeship. Rectangular boxes represent EFRO classes, while oval shapes represent individuals. Dotted arrows indicate *rdf:type* relations linking individuals to their corresponding classes. Solid arrows represent object property assertions between individuals. The ▲ symbol denotes subclass relationships within the EFRO class hierarchy, for example efro:Student is a subclass of foaf:Person. In this example, updated planned hours are not yet recorded following early progression; therefore, the data shown represents a non-compliant state with respect to this rule.

programmes within the same institution. Together, these questions assess EFRO's ability to represent regulatory conditions, evidence requirements, and temporal eligibility constraints in a unified model. All competency questions, sample individuals, and corresponding SPARQL queries are publicly available at [GitHub](#).

3.2. Illustrative Use Case

Figure 2 shows an instance-level application of EFRO to represent data captured by an education provider about a student as necessitated by a funding rule. This rule defines the conditions under which an education provider can receive funding when a learner progresses early from a traineeship. In this situation, the provider must record the programme the learner progressed from, the destination programme, and the number of weeks into the programme that progression occurred. The timing of progression is significant, as progression within the programme’s qualifying period, typically two or six weeks, may require revisions to the planned study hours recorded for the learner.

Figure 2 illustrates how this information is represented for a specific learner, `Tom_Wilson`. `Tom_Wilson` is enrolled in a traineeship, `Business_Skills`, modelled as `efro:StudyProgramme` with an associated qualifying period of six weeks. While qualifying periods normally define the minimum duration required for funding retention, the guidance specifies exceptions where early progression to certain destinations is permitted, including apprenticeships, sustainable employment, fulltime education, or other approved training programmes. In this case, `Tom_Wilson` progresses from the traineeship to a `Business_Administration` apprenticeship before the end of the qualifying period, and the learner’s status on the original programme is updated to reflect completion.

This modelling supports the evaluation of regulatory conditions, compliance with which can be assessed via SPARQL queries. The associated SPARQL query is available at [GitHub CQs](#).

3.3. Application-Oriented Validation

As an application-orientated validation, EFRO has been used in NeuReg⁸, a neuro-symbolic question-answer (QA) generation framework that combines regulatory text with EFRO-aligned knowledge graph triples to guide large language models (LLMs) QA generation [7]. In NeuReg, EFRO provides the schema for extracting triples from funding guidance, which are incorporated alongside source text in structured prompts to generate factual, relational, comparative, and inferential QA pairs. Provenance is preserved by linking extracted triples to their source text via chunk identifiers. Evaluations, including expert human assessment and multiple LLM judges, indicate that combining EFRO-aligned structured knowledge with regulatory text supports improvements in the accuracy, completeness, and semantic alignment of generated QA pairs compared to using text or knowledge graph inputs alone.

4. Conclusions and Future Work

EFRO provides a reusable vocabulary for representing education funding regulations and supports compliance-related reasoning and assessment of eligibility and supporting evidence. As shown through the early student progression use case (Section 3.1), EFRO allows institutions to structure their information required for evidencing compliance with funding regulations, formulate and execute queries over their data, and follow reporting requirements defined by regulatory guidance.

Beyond this immediate use for compliance reasoning, EFRO is intended to act as a foundational knowledge structure for intelligent systems operating in regulatory domains. Its modelling of obligations, evidence artefacts, temporal constraints, and learner status conditions offers a structured basis for combining symbolic representations with data-driven approaches in a controlled manner.

While EFRO’s regulatory content is specific to the UK, its upper-level concepts, such as `EducationalEntity`, `Obligation`, and `FundingElement`, are not tied to a specific jurisdiction’s rules and may provide a starting point for modelling education funding regulations elsewhere. Applying EFRO to another country’s regulatory context would require re-deriving the domain-specific subclasses (e.g., `Traineeship`, `QualifyingPeriod`) and competency questions from that jurisdiction’s own guidance, since these reflect UK-specific rules rather than general structural patterns.

Future work will investigate the integration of EFRO within Retrieval-Augmented Generation (RAG) pipelines, where the ontology may be used to organise retrieved regulatory text, support query formu-

⁸<https://github.com/RGU-Computing/NeuReg>

lation, and constrain generated responses. By making the ontology openly available, we encourage the community to build on and adapt EFRO for other domains and use cases. All resources are available at GitHub.

Declaration of use of Generative AI

During the preparation of this work, the author used ChatGPT for grammar and spelling checks, paraphrasing, and rewording. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

- [1] M. Fernández-López, A. Gómez-Pérez, N. Juristo, Methontology: From ontological art towards ontological engineering, in: AAAI Conference on Artificial Intelligence, 1997.
- [2] UK Department for Education, Apprenticeship funding rules, 2025. URL: <https://www.gov.uk/guidance/apprenticeship-funding-rules>, accessed: 2025-09-06.
- [3] UK Department for Education, Funding for 16 to 19 year olds in schools, 2025. URL: <https://www.gov.uk/education/funding-for-16-to-19-year-olds-in-schools>, accessed: 2025-09-06.
- [4] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, et al., The fair guiding principles for scientific data management and stewardship, *Scientific data* 3 (2016) 1–9.
- [5] E. Sirin, B. Parsia, B. Cuenca Grau, A. Kalyanpur, Y. Katz, Pellet: A practical owl-dl reasoner, *SSRN Electronic Journal* (2007). URL: <http://dx.doi.org/10.2139/ssrn.3199351>. doi:10.2139/ssrn.3199351.
- [6] B. Glimm, I. Horrocks, B. Motik, G. Stoilos, Z. Wang, Hermit: An owl 2 reasoner, *Journal of Automated Reasoning* 53 (2014) 245–269. URL: <http://dx.doi.org/10.1007/s10817-014-9305-1>. doi:10.1007/s10817-014-9305-1.
- [7] U. Arshad, D. Corsar, I. Nkisi-Orji, Neureg: Neuro-symbolic qa generation from regulatory compliance, in: *Proceedings of the 13th Knowledge Capture Conference 2025, K-CAP '25*, Association for Computing Machinery, New York, NY, USA, 2025, p. 232–235. URL: <https://doi.org/10.1145/3731443.3771375>. doi:10.1145/3731443.3771375.