

FIAS REGISTRY OF RESEARCH SPECIFICATIONS

French Judicial Decision Research-Citation Profile

REGISTRY IDENTIFIER

FOAR-CON2X-ED0JF68GZ

OBJECT TYPE

Identifier or citation scheme

DOCUMENT MATURITY

Stable

PUBLISHED

12 February 2021, 12:00 UTC

VERSION

1.1

SUBJECT AREA

Information and library science

SUBMITTED

30 November 2020, 12:00 UTC

LICENCE

Creative Commons Attribution 4.0

AUTHORITATIVE VERSION

Registration records the outcome of the Registry review for this version: the submission was sufficiently identified, documented, and internally determinate for the stated procedure. Scope, adoption, and competing conventions are documented separately where relevant.

REGISTRY STATUS**Registered**

12 February 2021, 12:00 UTC

Core description

SUBMISSION KIND

New version of a registered specification

EARLIER SUBMISSION OR REGISTRY IDENTIFIER

FIAS-SUB-2016-LAW-01

FULL TITLE

French Judicial Decision Research-Citation Profile

SHORT TITLE OR ACRONYM

FJDRCP

OBJECT TYPE

Identifier or citation scheme

STATUS CLAIMED BY THE SUBMITTER

Project standard

MATURITY OF THIS VERSION

Stable

VERSION

1.1

PRINCIPAL SUBJECT AREA

Information and library science

KEYWORDS

- French case law
- ECLI
- open justice data
- representation versioning
- pseudonymisation provenance
- reproducible corpus

PROBLEM, SCOPE, AND EXCLUSIONS

Version 1.1 retains the 2016 research-citation wrapper but addresses the new legal and technical reality of large-scale public dissemination of French judicial decisions. It records

the publication channel, retrieval timestamp, full digest, content media type, and whether the representation is marked as subject to basic or additional occultation. It enables a researcher to distinguish successive public representations without attempting to reverse occultation. It remains outside ECLI assignment, legal authority ranking, personal-data enrichment, and profiling of judges, clerks, parties, lawyers or litigants.

INTENDED USERS AND CONTEXTS

- Empirical-law teams downloading public French judicial decisions
- Corpus curators comparing releases of the same decision
- Peer reviewers checking the evidential basis of computational legal studies

RELATION TO EXISTING PRACTICE

Replaces an earlier specification

SCIENTIFIC CONTEXT, SOURCES, AND SCHOLARLY DISCUSSION

This section describes the scientific background, underlying work, and sources of the specification, together with the research groups and scholarly settings in which the proposed solution or its formalisation was presented or discussed. It documents the context and nature of these scholarly exchanges.

Reproducible research on judicial decisions requires a distinction between the decision, a particular published representation and the event through which that representation enters a corpus. ECLI supplies a stable decision identifier. The French framework for opening judicial data, developed through the 2016 law and the 2020 decree, also makes the publication and occultation state of the consulted text relevant to the interpretation of a dataset. CNIL's deliberation of 6 February 2020 addresses this publication context and the treatment of identifying information.

The dissemination and anonymisation of judicial decisions were discussed at the online conference *Convergences du droit et du numérique*, held on 12–14 October 2020. L. Béziaud's contribution examined the difficulties of anonymising decisions and the relationship between access to case law, transparency and privacy. These discussions form part of the contemporary context for documenting the provenance and successive public representations of research corpora.

Version 1.1 extends the earlier research-citation profile by recording the publication channel, retrieval time, media type and full SHA-256 digest in a manifest. A representation-state field captures the basic or additional occultation indicated for the published text. These elements allow researchers to compare successive releases and reproduce the documentary basis of an extraction or coding operation.

The profile belongs to the methodological discussion of judicial open data shared by legal-information research at CERSA and the implementation work of the Cour de cassation. D. Bourcier's work on open data and privacy contributes to this intellectual setting, alongside the development of documentary practices for public judicial corpora. The resulting citation

structure connects a stable reference to the decision with evidence of the exact representation used in a study.

References:

- Council ECLI Conclusions 2011/C 127/01: Source: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:C:2011:127:FULL>
- Loi n° 2016-1321 du 7 octobre 2016, arts. 20-21: Source: <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000033202746>
- CNIL, Délibération n° 2020-021 du 6 février 2020 on the draft open-data decree: Source: <https://www.legifrance.gouv.fr/cnil/id/CNILTEXT000042096094/>
- Décret n° 2020-797 du 29 juin 2020 relatif à la mise à la disposition du public des décisions: Source: <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000042055251>
- GDPR, Regulation (EU) 2016/679, Recital 26 and Articles 4-5: Source: <https://eur-lex.europa.eu/eli/reg/2016/679/oj>
- Open Law. Report on Convergences du droit et du numérique, online conference, 12-14 October 2020, including L. Béziaud's contribution on anonymisation. Conference report: <https://openlaw.fr/notre-actualite/retour-sur-levenement-convergences-du-droit-et-numerique/>

EXTERNAL STANDARDISATION STATUS

None intended

LICENCE FOR THE SUBMITTED SPECIFICATION

Creative Commons Attribution 4.0

PUBLICATION DECLARATION

Yes

Identifier or citation scheme

WHAT RECEIVES AN IDENTIFIER?

A versioned public representation of one French judicial decision as retrieved for a research corpus. The representation record is distinct from the ECLI-identified decision and from any private or non-occulted judicial source.

IDENTIFIER SYNTAX OR GRAMMAR

Display key:

FJDRCP/1.1/<ECLI-or-NOECLI>/<publisher>/<retrieved-UTC>/<digest-prefix>

Manifest canonical fields:

profile_version, ecli, fallback_id, publisher, source_url,
retrieved_at, media_type, sha256, occultation_notice,
source_modified_at, supersedes_representation.

retrieved-UTC := YYYYMMDDThhmmssZ.

digest-prefix := first 20 lowercase hexadecimal characters of sha256.

occultation_notice := NOT_STATED | BASIC | ADDITIONAL | SOURCE_DECLARED.

ASSIGNMENT AUTHORITY AND COLLISION PREVENTION

A corpus curator creates the record only from a publicly accessible source. The ECLI is copied exactly after syntax validation. The full source URL and SHA-256 are mandatory. A uniqueness constraint on (source_url, sha256) prevents duplicate representation records. The curator MUST NOT infer or reconstruct redacted personal information.

PERSISTENCE AND REASSIGNMENT POLICY

Released representation records are immutable. A corrected or newly occulted publication receives a new digest and record linked through supersedes_representation. The record for an earlier public artefact is retained when lawful, or its metadata is retained with access set to restricted if continued storage would be unlawful.

RESOLUTION OR LOOKUP MECHANISM

The ECLI resolves the decision concept through official search facilities. The source_url attempts retrieval of the representation. If bytes differ from sha256, the resolver reports SOURCE_CHANGED and directs the reader to the retained research package, subject to its access conditions; it never silently treats new bytes as identical.

CONCEPT, VERSION, EDITION, AND FRAGMENT GRANULARITY

Decision concept: ECLI. Public representation: source URL plus digest. Corpus inclusion event: representation key plus dataset release ID. Paragraph or page locator: a non-

identifying fragment selector attached to a citation occurrence and kept separate from decision identity. Researcher labels and model outputs remain separate annotations.

ALIAS, MERGE, SPLIT, AND DEPRECATION BEHAVIOUR

A corrected ECLI deprecates the wrapper containing the transcription error while preserving it in the record. A later occultation produces a successor representation with distinct bytes. A split PDF is represented by child artefacts with part numbers; a merged file has a new digest and links to all source parts.

VALID, INVALID, AND EDGE-CASE EXAMPLES

Valid display key:

FJDRCP/1.1/ECLI:FR:CCASS:2016:C101245/COURDECASSATION/20201130T103000Z/
0123456789abcdef0123

Minimal manifest fragment:

```
profile_version: "1.1"
ecli: "ECLI:FR:CCASS:2016:C101245"
publisher: "COURDECASSATION"
source_url: "https://example.invalid/real-url-redacted-in-test"
retrieved_at: "2020-11-30T10:30:00Z"
media_type: "text/html"
sha256: "0123456789abcdef0123456789abcdef0123456789abcdef0123456789abcdef"
occultation_notice: "SOURCE_DECLARED"
```

Invalid: occultation_notice: "ANONYMOUS". Profile rule: curators keep a source's occultation separate from any general claim of anonymity.

Registry review record

Submission received

30 November 2020, 12:00 UTC

Submission received.

Registered

12 February 2021, 12:00 UTC

Accepted as an incremental replacement of version 0.9. The decisive improvement is the separation of decision identity, published representation and corpus inclusion event. The term ANONYMOUS was removed from controlled values because occultation or pseudonymisation alone leaves irreversible anonymisation unproven.