CLARIN-LT repozitoriumas

Tomas Krilavičius

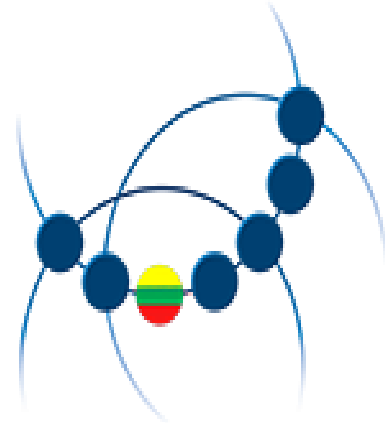https://clarin.vdu.lt/xmlui/

CLARIN-LT   **News**   **Repository**   **Legal information**

> " "There ought to be only one grand dépôt of art in the world, to which the artist might repair with his works, and on presenting them receive what he required... "
> — Ludwig van Beethoven, 1801

CLARIN-LT

🔍 [Search field]   Search

Advanced Search

Author	Subject	Language (ISO)
Rimkutė, Erika (2)	Lithuanian (3)	
Utka, Andrius (2)	wordlist (2)	
Aleksandravičiūtė, G ...(1)	corpus (1)	
Amilevičius, Darius (1)	Delfi (1)	
Bielinskienė, Agnė (1)	Lietuvos Rytas (1)	
... View More	... View More	

Lemmatised Wordlist of 1 m. Corpus of Contemporary Lithuanian

CLARIN-LT

✏ **Authors**	Utka, Andrius
🗓 **Date issued**	2016-11-18
🏷 **Type**	lexicalConceptualResource
⤢ **Size**	128,058 entries
⚑ **Language(s)**	Lithuanian
📄 **Description**	The lemmatised wordlist of 1 m. word Lithuanian corpus. The structure of the tab delimited text file (dazninis.txt): Headword<TAB>Part of Speech<TAB>Wordform<TAB>Frequency of Occurrence. The data is the basis for "Frequency Dictionary of Written Lithuanian - based on 1m word morphologically annotated corpus" (A_Utka-Dazninis_zodynas.pdf). Reference:

Ką turim šiandien?

- Repozitoriumas
  - Duomenų pateikimas
  - Duomenų paėmimas
  - Duomenų paieška
  - Autentifikacija
    - CLARIN ERIC
    - Lokalus
    - Integracija su EDUGAIN
  - Smulkmenos: citavimas, ….

- CLARIN ERIC atestavimas
  - Data Seal of Approval

Assessment

CLARIN ERIC

CLARIN ERIC B centrų atestavimas

- Formali procedūra

- Atlieka Centre Assessment Committee

- Procedūros tobulinamos

- Prieš pateikiant būtina gauti Data Seal of Approval

Show [all] entries   Search: [   ]

Centre	Shorthand	Type	Type status	City	Consortium	DSA
Collections de corpus oraux numeriques	COCOON	C	*none*	Paris	*none*	*none*
The ILC4CLARIN Centre at the Institute for Computational Linguistics	ILC4CLARIN	C	*none*	Pisa	CLARIN-IT (IT)	*none*
ASV Leipzig	ASV	B	Certified	Leipzig	CLARIN-D (DE)	seal
Bayerisches Archiv für Sprachsignale	BAS	B	Certified	München	CLARIN-D (DE)	seal
Berlin-Brandenburg Academy of Sciences and Humanities	BBAW	B	Certified	Berlin	CLARIN-D (DE)	seal
Center of Estonian Language Resources	CELR-EKK	C	Aiming for B.	Tartu	CELR (EE)	*none*
Centre for the Digital Foundation of Research in the Humanities, Social, and Educational Sciences	CEDIFOR	C	Aiming at B	Frankfurt, Germany	*none*	*none*

1. Entry in the center registry
2. B-center checklist
3. Data Seal of Approval application
4. Declaration of centre compliance (by NFC)

1. The centre candidate ensures

assessment@clarin.eu
2. The centre candidate applies by email

Standing Committee for CLARIN Technical C...

The Standing Committee for CLARIN Technical Centres (SCCTC) is a permanent Committee that reports to the Bo...
responsibility is to ensure consistence, coherence and stability of infrastructure services through decisions on in...
between the centres and members.

Its main tasks are:

- to serve as a forum for CLARIN centres to exchange ideas and experiences
- to coordinate between CLARIN centres by discussing and agreeing on measures needed for the impleme... infrastructure
- to review and update the Centre Types Document which was produced in the CLARIN Preparatory Phase
- to interact with the Centre Assessment Committee on relevant issues
- to advise the Board of Directors and the National Coordinators' Forum on any issues deemed relevant for... services provided by the CLARIN technical infrastructure.

Every year the SCCTC prepares an overview of the state of the technical infrastructure services for the General A...

All meetings, minutes and documents of the SCCTC can be accessed here.

| | Dieter Van Uytvanck | Chair Centre Committee & Technical D... |

Centre Assessment Committee

The Centre Assessment Committtee (CAC) is a permanent Committee that report...
the level and quality of CLARIN Centres of type A, B and E, as described in the Cer...

Its main tasks are:

- to assess the quality of candidate centres of type A, B and E;
- to regularly assess the quality of centres that have been recognized as typ...
- to maintain the set of centre requirements and adapt them to new develo...
- to advise the Board of Directors on the content of Service Level Agreemer...
- to advise the Board of Directors on requirements and recognition procedu...
- to advise the Board of Directors on any issues deemed relevant for mainta... infrastructure.

Assessment of a centre can take place

- in the case of a new candidate centre on the basis of a request submitted...
- in the case of existing centres on the basis of a request by the head of the...
- in the case of existing centres on the CAC's own initiative.
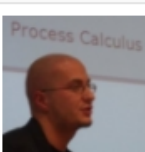
The draft assessment report is sent to the BoD, and to the head of the centre wit...
further action. CAC members cannot participate in the assessment of their own c...
chair.

Every year the CAC prepares an overview of the state of the centres for the Gene...
meeting takes place.

All relevant documents for the Centre Assessment Committee are available here.

| | Lene Offersgaard | |

Dieter & Lene

3. CAC & SCCTC provide timeline

	Lene Offersgaard	Chair
	Jozef Mišutka	Vice Chair
	Patrice Bellot	Member
	Tobias Blanke	Member
	Daan Broeder	Member
	Tomas Krilavicius	Member

4. CAC evaluates Centre

Announced
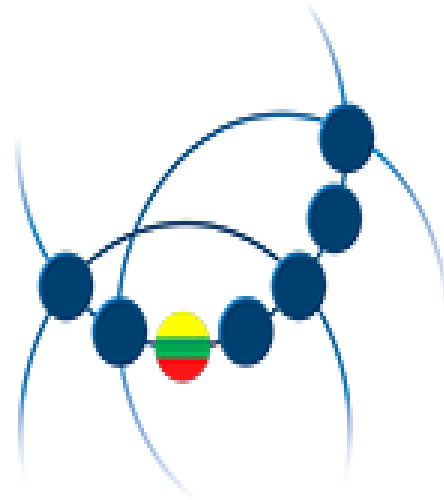
Situacija šiandien

1. Centre Compliancy (document)
2. Visibility of connection to CLARIN
3. Funding support

1. General Requirements

2. Intellectual Property Rights & Privacy

1. Data offering & IPR
2. Privacy statement (GÉANT Data Protection Code of Conduct)

Data Seal of Approval Benefits

Certifying your repository against the 16 Data Seal of Approval guidelines has many benefits, chief among them that the repository is recognized in the community as a trustworthy source of data.

Starting point for obtaining the Data Seal of Approval

Complete a self-assessment in the DSA online tool. The online tool takes you through the 16 guidelines and provides you with support.

Once you have completed your self-assessment you can submit it for peer review. The peer reviewer will go over your answers and documentation and will award the Data Seal of Approval (DSA) when you have met the requirements.

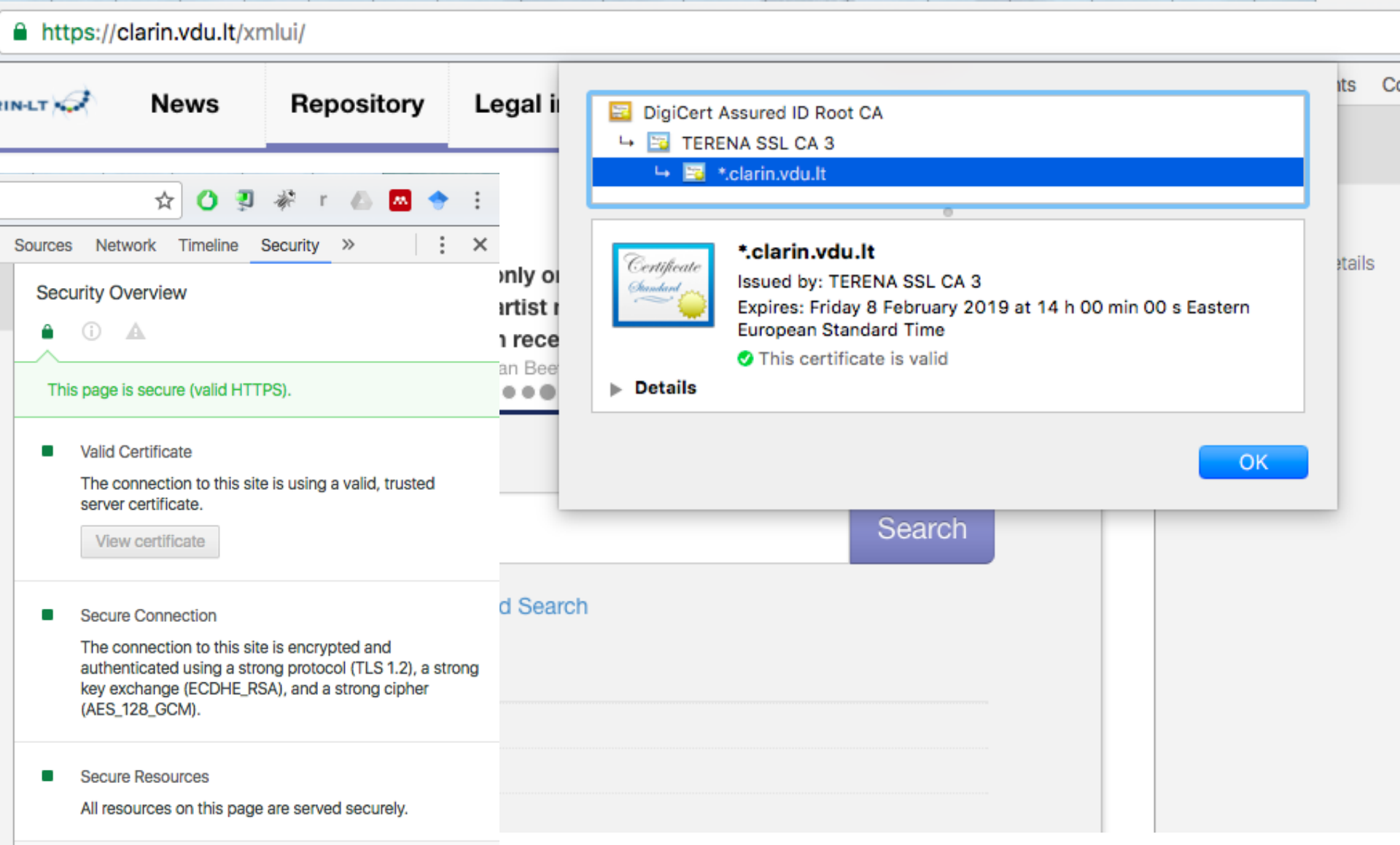Your self-assessment and review will not become public until the Data Seal of Approval is awarded.

Apply now!
or
Log in

Data Seal of Approval

3. Assessment

News Repository Legal i

DigiCert Assured ID Root CA
↳ TERENA SSL CA 3
↳ *.clarin.vdu.lt

*.clarin.vdu.lt
Issued by: TERENA SSL CA 3
Expires: Friday 8 February 2019 at 14 h 00 min 00 s Eastern European Standard Time
✅ This certificate is valid
▶ **Details**

OK

Sources Network Timeline Security »

Security Overview

🔒 ⓘ ⚠

This page is secure (valid HTTPS).

Valid Certificate

The connection to this site is using a valid, trusted server certificate.

View certificate

Secure Connection

The connection to this site is encrypted and authenticated using a strong protocol (TLS 1.2), a strong key exchange (ECDHE_RSA), and a strong cipher (AES_128_GCM).

Secure Resources

All resources on this page are served securely.

Search

4. Server Certificates

5. Federated Management

1. SAML2 Service provider
2. shib_test.pl
3. National Identity Federation
4. CLARIN IdP users support
5. CLARIN Service Providers Federation
6. EDUGAIN

Clarin.eu website account
1240 km

clarin

▸ **Please help, I cannot find my provider**

⚠ Error from IdP Discovery Service [clarin.vdu.lt]: No access to return content to [https://clarin.vdu.lt/xmlui/themes/UFAL/lib/html/disco-juice.html?cid=0]. Return parameter invalid.

⦾ Locate me and show nearby providers

Show providers in [all countries] show all countries

DiscoJuice © UNINETT

| Lithuania | LITNET fedi | Connection planned since 2015. |

1. OAI-PMH endpoint

2. Metadata harvesting by CLARIN ERIC

3. Visibility in VLO

4. CMDI files + profiles + ISOcat

4. Metadata

http://hdl.handle.net/99999/12
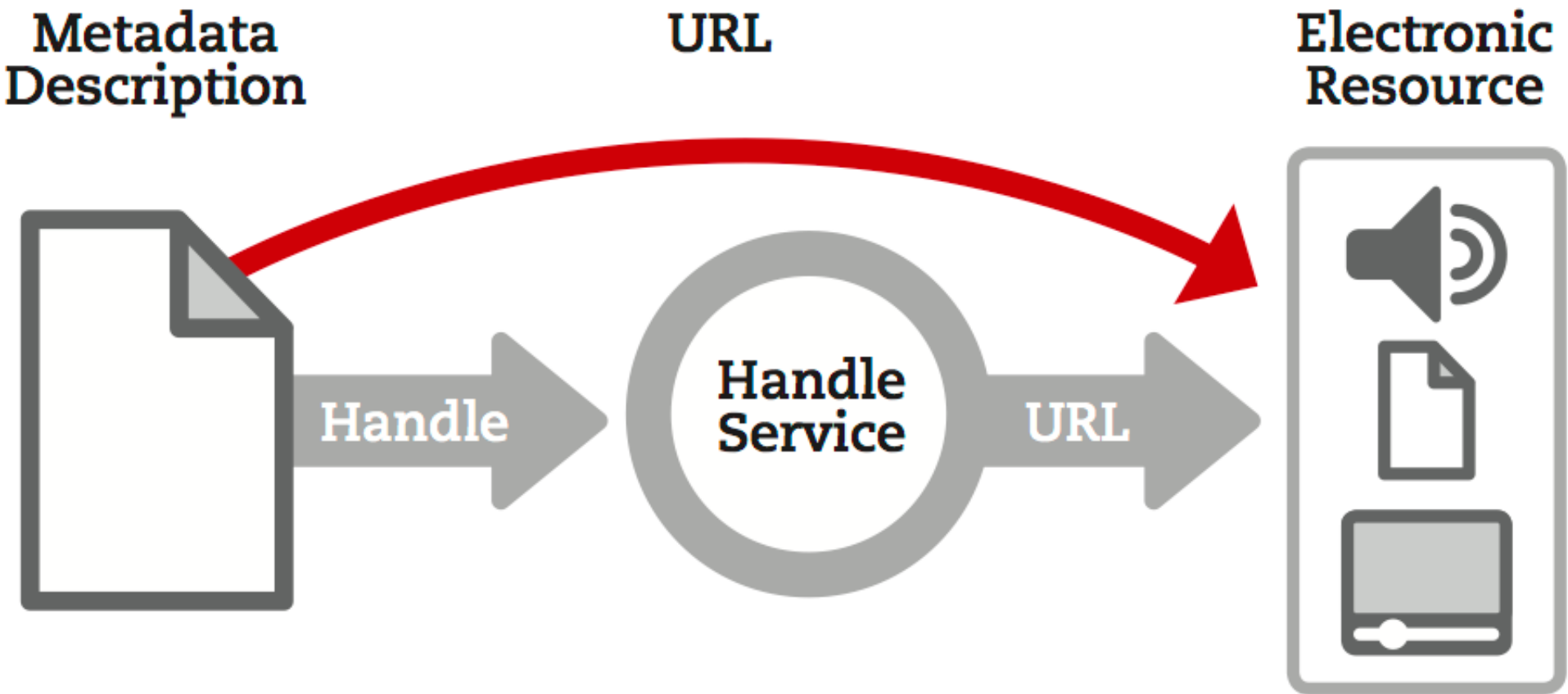
5. Persistent Identifiers

```
                          +----------+
                          |  Client  |
                          +----------+
                              /|\
                               |
                  ----------------------------
                  |     SRU / CQL            |
                  | w/CLARIN-FCS extensions  |
                  ----------------------------
                               |
                              \|/
       +------------------------------------------------+
       |   |          Endpoint        /|\              |
       |   |              |            |               |
       |   ----------------       -------------------  |
       |   | Translate CQL |      | Translate Result | |
       |   ----------------       -------------------  |
       |          |                      |             |
       |         \|/                     |             |
       +------------------------------------------------+
                              /|\
                               |
                              \|/
                  +----------------------------+
                  |       Search Engine        |
                  +----------------------------+
```
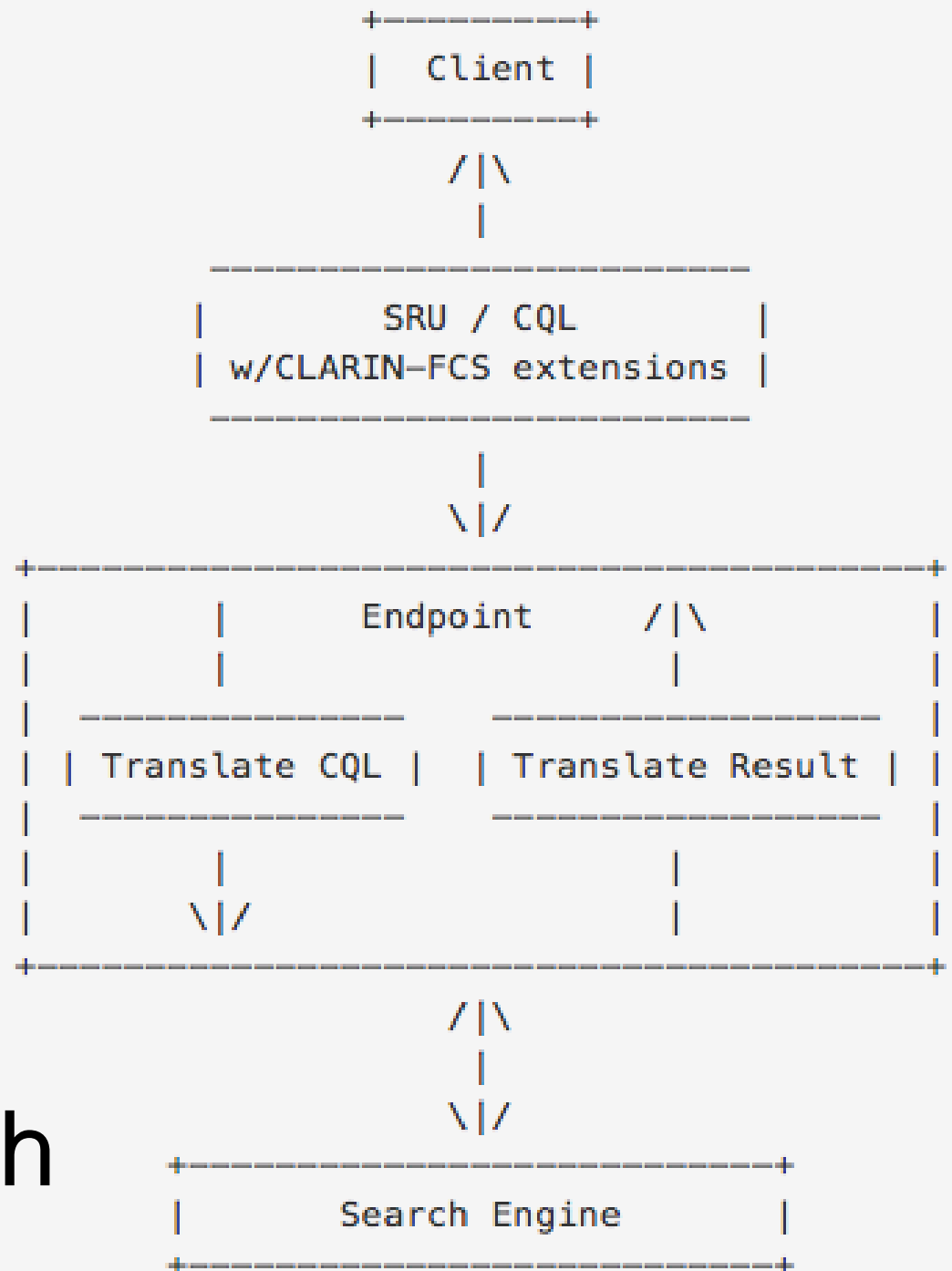
Skip for now

6. Federated Content Search

Not finished yet

Apibendrinimas

- Pasirinktas Lindat sprendimas

- Daugiau komplikacijų, negu buvo galima tikėtis
    - FEDI
    - Daugiakalbystė
    - Daug iki galo nedokumentuotų dalykų

- Pasirodė nauja Lindat versija – ką darysim?

- Veikia, bet yra trūkumų