CLARIN CENTRE K

INTELE

Infraestructura de Tecnologías del Lenguaje

DARIAH-EU

INTELE:

Tecnologías del Lenguaje para eHumanidades y ciencias sociales (y más allá)

RED2018-102797-E

www.clarin.eu
www.dariah.eu

German Rigau
INTELE - UPV/EHU

Objetivos de INTELE

INTELE: red estratégica para la participación en infraestructuras europeas

- **Lograr** la participación oficial de España en CLARIN y DARIAH.
- **Contribuir** al avance de la investigación española en humanidades y ciencias sociales
- **Posicionamiento estratégico** en proyectos y programas **internacionales**, fundamentalmente en el contexto del Espacio Europeo de Investigación.

CLARIN

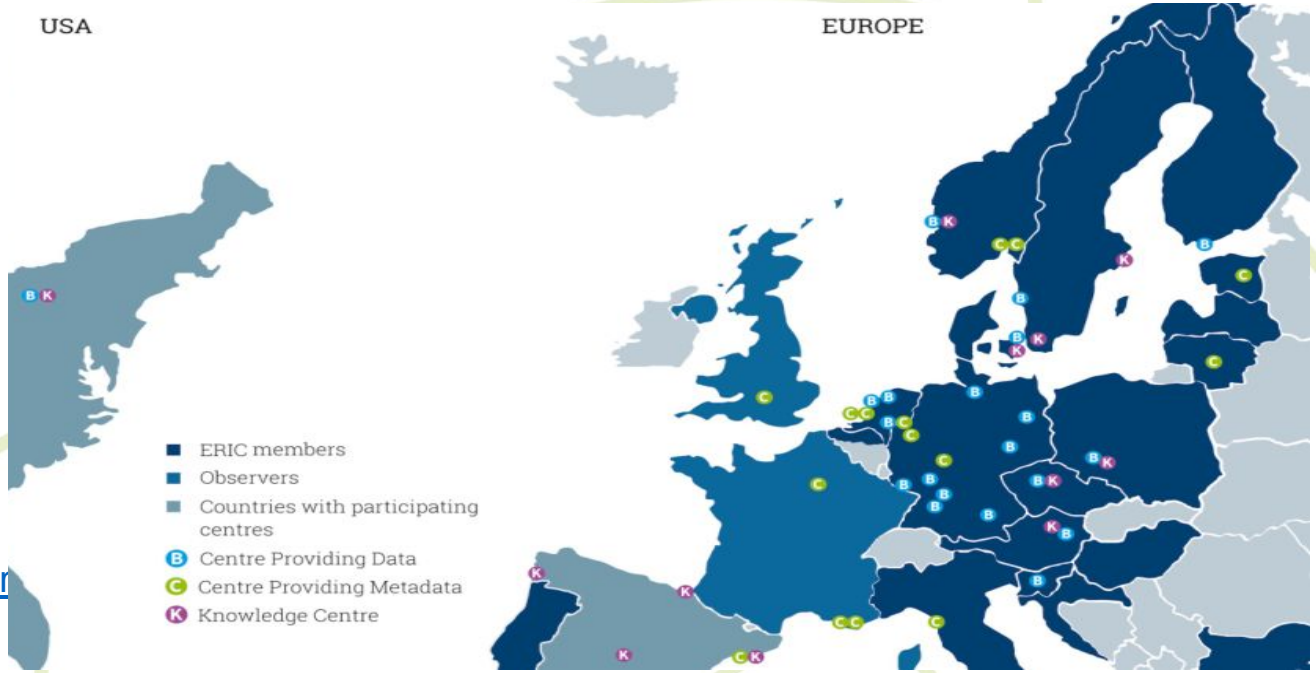
ESFRI European Strategy Forum on Research Infrastructures

- **C**ommon **LA**nguage **R**esources and Technology **INfrastructure**
  - ESFRI ERIC (2012) y ESFRI Landmark (2016)
- El objetivo es el de **digitalizar** las Humanidades y Ciencias Sociales (y más allá) y **dar acceso confederado a la tecnología lingüística** de todas las lenguas europeas desde una única web, para ayudar a los investigadores de estos ámbitos
  - Datos lingüísticos digitales (textos, voz, multimodales)
  - Herramientas avanzadas: describir, etiquetar, analizar, comparar, visualizar y compartir

Strategy Report on Research Infrastructures ROADMAP 2018

CLARIN en cifras (en 2018)

- 52 centros de servicios, datos y herramientas (C, B, K)
- 20 centros-B certificados
- 6 (+7) proyectos europeos
- 192 eventos con 10.000 participantes
- Más de 1 millón de recursos disponibles



FUENTE:
https://www.clarin.eu/content/overview-clarin-centres

DARIAH-EU

ESFRI European Strategy Forum on Research Infrastructures

DARIAH, “Digital Research Infrastructure for the Arts and Humanities”

- Aspira a fomentar y dar soporte a la **investigación y docencia digital** en las **artes y humanidades**.
- Desde 2019 en fase operacional desarrollando infraestructura con métodos/herramientas digitales para realizar análisis e interpretaciones innovadoras.

August 2014-August 2019: Implementation phase of DARIAH ERIC

Since September 2019: Operational phase of DARIAH ERIC

https://www.dariah.eu/about/history-of-dariah/

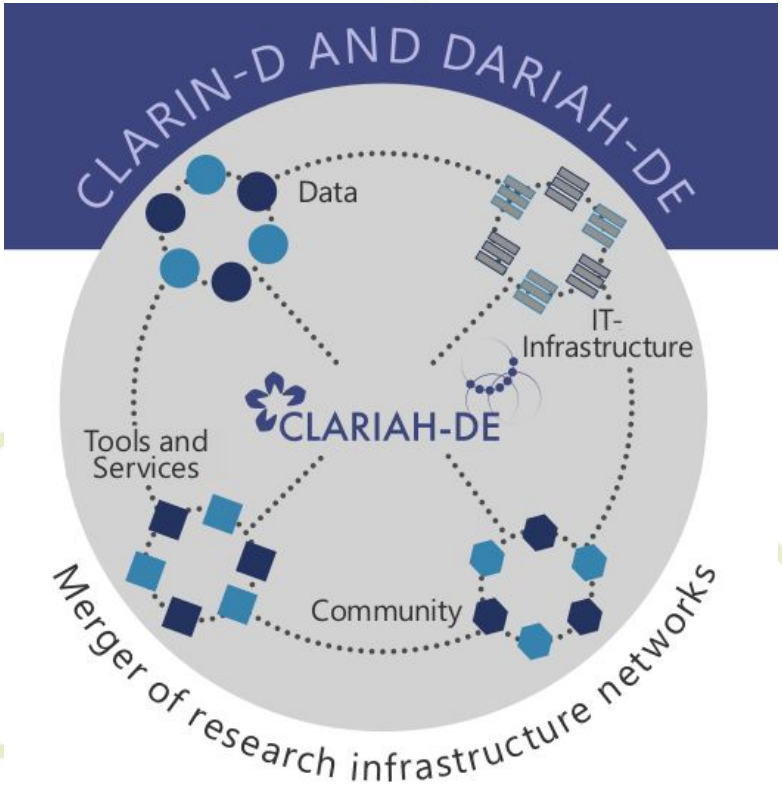
Strategy Report on Research Infrastructures ROADMAP 2018

Modelo de infraestructura: CLARIAH (CLARIN + DARIAH)

ESFRI European Strategy Forum on Research Infrastructures

THE NETWORK IN EUROPE

The CLARIN, DARIAH and CLARIAH research infrastructures are active on a national and the European level. CLARIN and DARIAH both have the status of a European Research Infrastructure Consortium (ERIC).



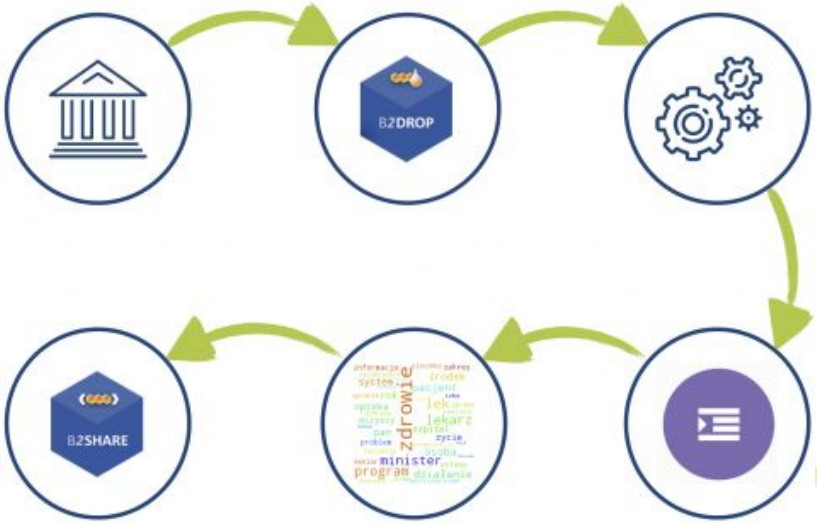
The content-related and technological foundations created by CLARIN-D and DARIAH-DE will be aligned, integrated, further developed and jointly maintained in CLARIAH-DE.

Investigación de principio a fin: FAIR principles

CLARIN era un caso FAIR, avant la lettre

Demo de caso de uso en EOSC

- Pregunta inicial:
  - Do parliamentary speeches of female and male members of parliament differ?
  - If they do, what are typical topics for each group?
- Los datos
  - Polish Parliamentary Corpus
- Análisis en la nube
- Publicación de todo



https://www.clarin.eu/showcase/eosc-portal-demonstration

Cobertura por lenguas de la tecnología CLARIN

https://vlo.clarin.eu hay más de 1 millón de componentes (herramientas, corpus …)

- English (154.928)
- German (143.271)
- Dutch (117.312)
- Danish (109.962)
- Slovenian (73.495)
- Polish (40.466)
- French (24.432)
- Afrikaans (7.870)
- ...

Recursos en las lenguas de CLARIN-es

- Spanish; castilian (14.444)
- Catalan; valencian (1.364)
- Basque (498)
- Galician (216)

El caso de Holanda y del holandés

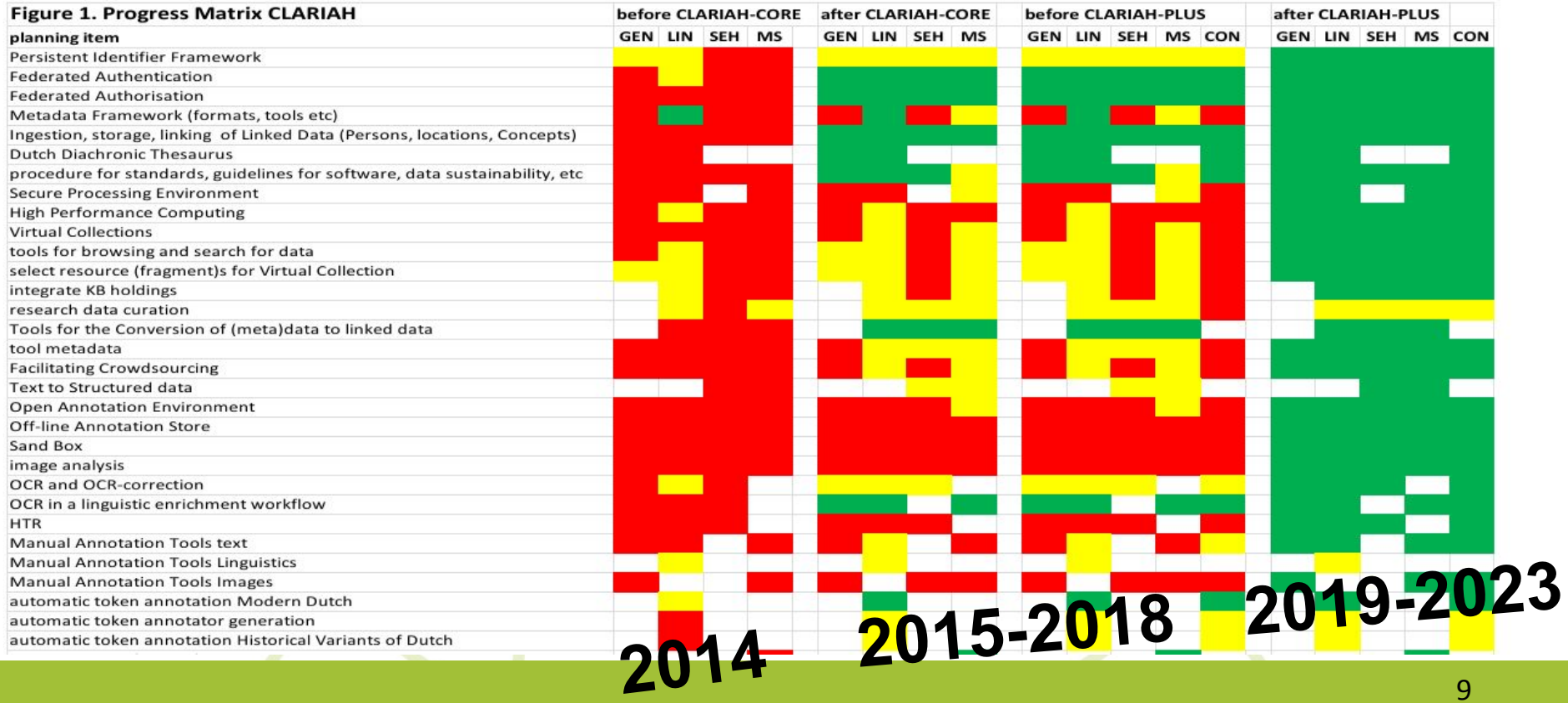
National Roadmap for Large-Scale Research Infrastructure

1 General information

GENeric functionality, LINguistics, Socio-Economic History and Media Studies, CONtent of texts: history, literary

FUENTE:

https://www.clariah.nl/over/bestanden/downloads/send/10-folders/166-clariah-plus



Propuesta de CLARIAH-ES

- Unión de CLARIN y DARIAH (modelo holandes, alemán)
  - Fee de CLARIN y DARIAH
- Organización (modelo alemán)
  - Universidades
  - CCAAs, diputaciones, etc.
  - Estado
- Alcance de la futura red CLARIAH-ES
  - a) **Servicios a investigadores**
    - > 2 personas tiempo completo por lengua + coordinación
  - b) Servicios + **difusión + formación**
  - c) Servicios + difusión + formación + **desarrollo de nueva tecnología**
    - Según los proyectos a desarrollar

Posibles escenarios de financiación anual de CLARIAH-ES

Proporcional al PIB			
	Como Dinam.	**Como Eslov.**	**Como Suecia***
CLARIN fee (2020)	110,605.00	110,605.00	110,605.00
CLARIN Total (2016)	1,072,308.38	2,122,796.13	4,744,998.15
DARIAH fee (2020)	82,000.00	82,000.00	8,200.00
DARIAH Total (2015)	491.000.00	491.000.00	49.100.00
CLARIAH-ES Total approx	1,755,913.38	2,806,401.13	4,912,903.15

* Suecia tiene estatus de país observador en DARIAH

Red INTELE y workshop

- Más de 130 grupos de más de 40 universidades
  - FFI2017-90738-REDT GREDISOR El Grupo Estudios del Discurso Oral (8 grupos)
  - RED2018-102418-T PLN.Net (13 grupos)
- También BNE, RAE, etc.

Formulario

View | Results

Submissions | Analysis | Table | Download

Showing 1 - 50 of 65. Show 10 | 50 | All results per page.

#	Submitted	User
65	11/04/2020 - 13:33	Anon
64	11/04/2020 - 12:21	Anon
63	11/04/2020 - 11:31	Anon
62	11/04/2020 - 11:28	Anon
61	11/04/2020 - 11:24	Anon
60	11/04/2020 - 11:21	Anon
59	11/04/2020 - 11:04	Anon
58	11/04/2020 - 10:31	Anon
57	11/04/2020 - 09:24	Anon
56	11/03/2020 - 17:50	Anon
55	11/03/2020 - 17:30	Anon
54	11/03/2020 - 17:27	Anon

You have already submitted this form. View your previous submissions.

Rellene este formulario

View | Results

Roadmap de INTELE*

- 11-2020 – **Catálogo de grupos** que apoyan la entrada de España en ambas infraestructuras europeas (borrador).

- 12-2020 – Informe sobre las **necesidades.**

- 09-2021 – Reunión anual CLARIN en Madrid.

- 09-2021 – Catálogo de grupos.

- 09-2021 – Contribuciones **in-kind**

- 12-2021 – **Evaluación** para su inclusión en ESFRI:
  - Consorcio español para CLARIN y DARIAH
  - Contribuciones in-kind para las infraestructuras CLARIN y DARIAH.

* Tentativo

CLARIN CENTRE K

INTELE

Infraestructura de Tecnologías del Lenguaje

DARIAH-EU

INTELE:

Tecnologías del Lenguaje para eHumanidades y ciencias sociales (y más allá)

RED2018-102797-E

www.clarin.eu
www.dariah.eu

German Rigau
INTELE - UPV/EHU