



HAL : Strengthening Connections Between Publications, Data, and Software in the French National Open Science Ecosystem

Open Repositories 2025, Chicago, USA

Yannick Barborini

Center for Direct Scientific Communication (CCSD) / CNRS



HAL+, a Research Infrastructure for Scholarly Communication



HAL Overview



- **Founded in 2001** by a physicist, inspired by the **arXiv** model
- **Open repository** for sharing and preserving scholarly publications in all disciplines
- Serves as the **national multidisciplinary open repository** for French research
- Supported by all French universities and research institutions



HAL
open science

Open and share the knowledge

+ Upload

			
An international scope	A common good for research	A large collaborative community	An archive, some services
Your publications are easy to find, will be referenced by search engines and interconnected with other services (ORCID, preprint servers)	The largest research organizations and the majority of French universities have chosen and support HAL, a public, sustainable and responsible infrastructure.	From the researchers to the specialists of the scientific information, HAL federates multiple skills to support the opening of the publications	HAL guarantees the long-term preservation of your publications. A set of services (CV, institutional portals, collections, documentary watch, APIs, identifiers) contribute to their valorisation.

Key Figures

✓ **1.5 M**

Full-text Documents

*Published & Unpublished Scholarly Works
(50% Journal Articles)*

 **70% English, 25% French**

Content in Several Languages

 **170,000**

New Deposits in 2024

Consistent Annual Growth

 **100,000+**

Active Users / Year

Growing Community of Researchers

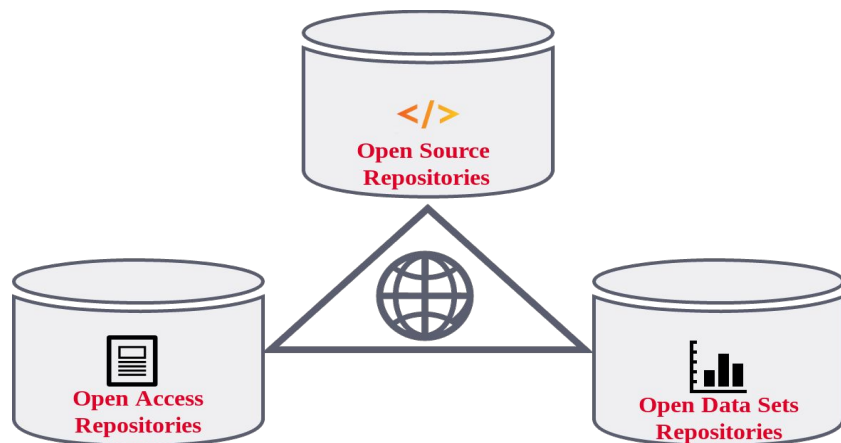
 **200+ Million**

Consultations & Downloads

Record Usage of Content

Enabling Reproducibility and Openness

To advance reproducible science and support knowledge sharing, it is essential to connect the **three pillars of Open Science**:



Three pillars of Open Science
Software Heritage CC-BY 4.0 2019



This work is carried out as part of the HALiance project, funded by the **French National Research Agency (ANR)**. PIA - Equipex+ HALiance - ANR-21-ESRE-0047

Second French Plan for **Open Science**



<https://www.enseignementsup-recherche.gouv.fr/sites/default/files/2021-10/second-french-plan-for-open-science-13715.pdf>

Linking Data to Publications: Creation Form

1 Choose the Type of Relationship

Based on a simplified **DataCite typology** (e.g., isPartOf, isSupplementedBy)

2 Select the Repository/Archive

Supports major French data repositories:

- HAL
- Recherche Data Gouv
- Nakala

3 Assisted Search

Helps locate datasets within selected repositories

4 Identifier Validation

Checks that provided DOIs or other persistent IDs are valid

5 Citation Retrieval

Automatically retrieves dataset citation from:

- Recherche Data Gouv
- Nakala
- Any repository with a DOI (via **citation.doi.org**)

Cites (Cites)
Is cited by (IsCitedBy)
Has version (HasVersion)
Is version of (IsVersionOf)
Is identical to (IsIdenticalTo)
✓ Is part of (IsPartOf)
Has part (HasPart)
Reviews (Reviews)
Is reviewed by (IsReviewedBy)
Documents (Documents)
Is documented by (IsDocumentedBy)
Is translation of (IsTranslationOf)
Has translation (HasTranslation)

HAL
Nakala
✓ Recherche Data Gouv
Software Heritage
Other source (other repository, archive, etc.)

Relationship
Is part of (IsPartOf)

Repository
Recherche Data Gouv

Search a deposit in Recherche Data Gouv (identifier, title)
olive tree

ROUSSEY, Catherine; AZIAN, Amira, 2024, "phenological scales of olive tree", <https://doi.org/10.57745/E33CXX>, Recherche Data Gouv, V1

BADEL Eric; Arnoul, Olivier; Bozonnet, Cyril; Conchon, Pierre; Duhamel, Yvan; Ruelle, Julien; Viguier, Joffrey; King, Andy, 2024, "Olivier / Olea europaea / Olive tree", <https://doi.org/10.57745/ME912Z>, Recherche Data Gouv, V2

SARAH, Gautier; Aqbouch, Laila, 2025, "Effects of chill and heat conditions in different years on the genetic determinism of flowering stages in the olive tree", <https://doi.org/10.57745/LPZODO>, Recherche Data Gouv, V1, UNF:6:pvLrftRq6gFgyXc7nN1qVg== [fileUNF]

Alletru, David; Gazel, François; Corradi, Michel; Greil, Marie-Laure; Delmas, Marine, 2020, "Phenological data of 26 varieties of cherry, apricot, peach and apple trees since 2016 in a French orchard of Toulence, Gironde, France", <https://doi.org/10.15454/TJTV6N>, Recherche Data Gouv, V3, UNF:6:HPzk9LBM04Yk5dx5fRsT6Q== [fileUNF]

Visibility and Display of Related Resources

0 Dedicated Widget on the Deposit Landing Page

- Clearly displays and previews associated datasets or software
- Improves visibility and accessibility of linked research outputs

0 Specific API fields

- New fields available to retrieve and expose relationships
- Facilitates integration with external platforms and tools

0 Display in export formats

- Relationships included in different export formats (Dublin Core, TEI, RIOXX)
- Ensures interoperability and consistent metadata across systems

Download

Towards emotion analysis for Alsatian theater

Grimy Liu (1), Pablo Ruiz Falo (2), Delphine Bernhart (2)

Show details

- USA - Université Grenoble Alpes
- LILPA - Linguistic, Langues et Parole

Dates and versions

hal-04213917, version 1 (05-01-2024)

License

Attribution

Identifiers

HAL ID : hal-04213917, version 1
DOI : 10.5281/zenodo.8492522

Cite

Grimy Liu, Pablo Ruiz Falo, Delphine Bernhart. Towards emotion analysis for Alsatian theater. Computational Humanities Research / CHRS 2023 Dec 2023, Paris, France, 2023. 10.5281/zenodo.8492522. hal-04213917.

Export

Bibtex XML TEI Dublin Core

DC Terms EndNote DataCite

Collections

USGA SITE+SLACS

122 View 59 Download

Altmetric

7

Share

Towards emotion analysis for Alsatian theater

Grimy Liu (1), Pablo Ruiz Falo (2), Delphine Bernhart (2)

Show details

- USA - Université Grenoble Alpes
- LILPA - Linguistic, Langues et Parole

Abstract en

The poster presents the first emotion analysis system for theater in Alsatian language varieties. It is a lexicon-based system, relying on an emotion lexicon which handles Alsatian's huge morphological variability. We describe the system's functioning, regarding configurable term-weighting schemes, scores aggregation based on plays' metadata (e.g. genre) and character related variables (gender, socio-professional group), and visualizations output by the tool, presenting some initial results. The system, still at an early stage, intends to contribute to computational literary studies of a tradition scarcely studied using computational methods.

Keywords en

Alsatian theater Emotion analysis

Dominants

Computer science [L] Linguistics Literature

Complete list of metadata

Files and preview

Related resources -

- [10.34847/nhl.40cwp9n \[Preprint\]](#)
 — Bernhart D. (2022). EAL : lexique de mots d'émotion en alsacien pour l'analyse de pièces de théâtre (Version 1) [Data set]. NAMALA - https://hal.archives-ouvertes.fr/hal-04213917. CNRS. https://doi.org/10.34847/nhl.40cwp9n
- [10.34847/nhl.398761tv \[Preprint\]](#)
 — Bernhart D. (2020). EAL : score d'activation pour les dialectes alsaciens (Version 1) [Data set]. NAMALA - https://hal.archives-ouvertes.fr/hal-04213917. CNRS. https://doi.org/10.34847/nhl.398761tv
- [10.34847/nhl.62rhvde \[Preprint\]](#)
 — Liu G. (2022). METNAL, Personalized Alsatian Theater Character Features (Version 1) [Data set]. NAMALA - https://hal.archives-ouvertes.fr/hal-04213917. CNRS. https://doi.org/10.34847/nhl.62rhvde

Pablo Ruiz Falo - Connect in order to contact the contributor

https://hal.archives-ouvertes.fr/hal-04213917

Last modification on : Friday, January 5, 2024 at 14:23 PM

Scheduled on : Monday, December 16, 2024 at 10:38:17 PM

Example of Data and Software Preview

[illegible]

<https://hal.science/hal-03336685v1>

<https://hal.science/hal-04213017v1>

<https://hal.science/hal-04344205v1>

Visibility and Display of Related Resources

0 Dedicated Widget on the Deposit Landing Page

- Clearly displays and previews associated datasets or software
- Improves visibility and accessibility of linked research outputs

0 Specific API fields

- New fields available to retrieve and expose relationships
- Facilitates integration with external platforms and tools

0 Display in export formats

- Relationships included in different export formats (Dublin Core, TEI, RIOXX)
- Ensures interoperability and consistent metadata across systems

Download

Dates and versions
 hal.04213017 - version 1 [05-01-2024]

Licence
 Attribution

Identifiers
 Hal Id : hal.04213017 - version 1
 DOI : 10.5281/zenodo.840522

Cite
 Qinys Liu, Pablo Ruiz Fabo, Delphine Bernhart. Towards emotion analysis for Alsatian theater: Computational/Humanist Research (CHoV) 2023, Dec 2023, Paris, France, 2023, 15.0281/zenodo.840522 - hal.04213017.

Export
 BibTex XML TeX Dublin Core
 DC Terms Endnote DataCite

Collections
 UGA SALSACE

122 Views
 59 Downloads

Altmetric

Share

Towards emotion analysis for Alsatian theater

Qinys Liu (1), Pablo Ruiz Fabo (2), Delphine Bernhart (2)

Show details

- USA - Université Grenoble Alpes
- LLPL - Linguistique, Langues et Parole

Abstract

The poster presents the first emotion analysis system for theater in Alsatian language varieties. It is a lexicon-based system, relying on an emotion lexicon which handles Alsatians' huge morphosyntactic variability. We describe the system's functionality, regarding configurable term-weighting schemes, score aggregation based on plays' metadata (e.g. genre) and character-related variables (gender, socio-professional group), and visualizations output by the tool, presenting some initial results. The system, still at an early stage, intends to contribute to computational literary studies of a tradition scarcely studied using computational methods.

Keywords

Alsatian theater Emotion analysis

Dominants

Computer Science LINGUISTICS Literature

Complete list of metadata

Files and preview

RELATED RESOURCES

- [hal.04213017](#) **Document**
 — Bernhart, D. (2023). E&A : Inviser de mout d'effection en alsacien pour l'analyse de piéces de théâtre (Version 1) [Data set]. NAKALA - https://nakala.fr/hal.04213017.
- [hal.3987617v1](#) **Document**
 — Bernhart, D. (2023). E&A : Inviser d'effectuations pour les dialectes alsaciens (Version 1) [Data set]. NAKALA - https://nakala.fr/thume-hum - CNRS. https://doi.org/10.34847/NAL.3987617V
- [hal.6287618v1](#) **Document**
 — Liu, Y. (2023). M&T&P: Perceptron-based Alsatian Theater Character Features (Version 1) [Data set]. NAKALA - https://nakala.fr/thume-hum - CNRS. https://doi.org/10.34847/NAL.6287618V

Pablo Ruiz Fabo - Connect in order to contact the contributor

https://hal.archives-ouvertes.fr/04213017

Submitted on : Friday, January 5, 2024 10:42:33 PM

Last modification on : Monday, December 16, 2024 3:38:17 PM

Example of Data and Software Preview

H-I	Périodicité	création	Discipline(s)	éditoriales	bibliographiques	contribution	intégraux	clés	clés	Résumés
308	Annuelle	2011 - 2020	Art Littérature Linguistique	Yes	Yes	Ponctuellement	French (FR, FRIE)	Yes	French (FR, FRIE) English (EN, ENG) Spanish (ES, SPA)	
349	Semestrielle	2011 - 2020	Droit	Yes	Yes	Jamais	French (FR, FRIE)	Yes	French (FR, FRIE)	
367	Semestrielle	2011 - 2020	Information et communication	Yes	Yes	Pour tous les numéros.	French (FR, FRIE)	Yes	English (EN, ENG)	140 bytes

<https://hal.science/hal-03336685v1>

<https://hal.science/hal-04213017v1>

<https://hal.science/hal-04344205v1>

Automated Exchange of Relationships Between Data and Publications

0 Reciprocal Linking Made Simple

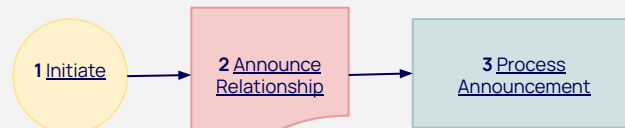
The relationship is shared with the data repository to establish a two-way link

0 Enabled by the COAR Notify protocol

- **International initiative** to develop and promote a standardized, decentralized, and interoperable protocol for connecting research outputs and services across repositories and platforms
⇒ <https://coar-notify.net/>
- **Built on Open Web Standards**
 - Linked Data Notifications (LDN)
 - ActivityStreams



“Repository Announces Relationship” Workflow



```

{
  "@context": [
    "https://www.w3.org/ns/activitystreams",
    "https://coar-notify.net"
  ],
  "actor": {
    "id": "https://research-organisation.org",
    "name": "Research Organisation",
    "type": "Organization"
  },
  "context": {
    "id": "https://another-research-organisation.org/repository/datasets/item/201203421/",
    "set:item": "https://doi.org/10.5555/999555666",
    "set:item": {
      "id": "https://another-research-organisation.org/repository/datasets/item/201203421/data.archive.zip",
      "mediaType": "application/zip",
      "type": [
        "Object",
        "sorg:Dataset"
      ]
    }
  },
  "type": [
    "Page",
    "sorg>AboutPage"
  ],
  "id": "urn:uuid:94ecae35-dcfd-4182-8558-22c7164fe29f",
  "object": {
    "as:object": "https://another-research-organisation.org/repository/datasets/item/201203421/",
    "as:relationship": "http://purl.org/vocab/frbr/core#supplement",
    "as:subject": "https://research-organisation.org/repository/item/201203421/",
    "id": "urn:uuid:74FFB356-6632-44D9-B176-8880A857580C",
    "type": "Relationship"
  },
  "origin": {
    "id": "https://research-organisation.org/repository",
    "inbox": "https://research-organisation.org/inbox/",
    "type": "Service"
  },
  "target": {
    "id": "https://another-research-organisation.org/repository",
    "inbox": "https://another-research-organisation.org/inbox/",
    "type": "Service"
  },
  "type": [
    "Announce",
    "coar-notify:RelationshipAction"
  ]
}
  
```


Progress Update: Linking Publications, Data, and Software

0 New Relationship Creation System

- **Operational since early 2025**
- Enables authors to establish explicit relationships between Publications, Research data and Software/ source code
- **+33,000 relationships** already created

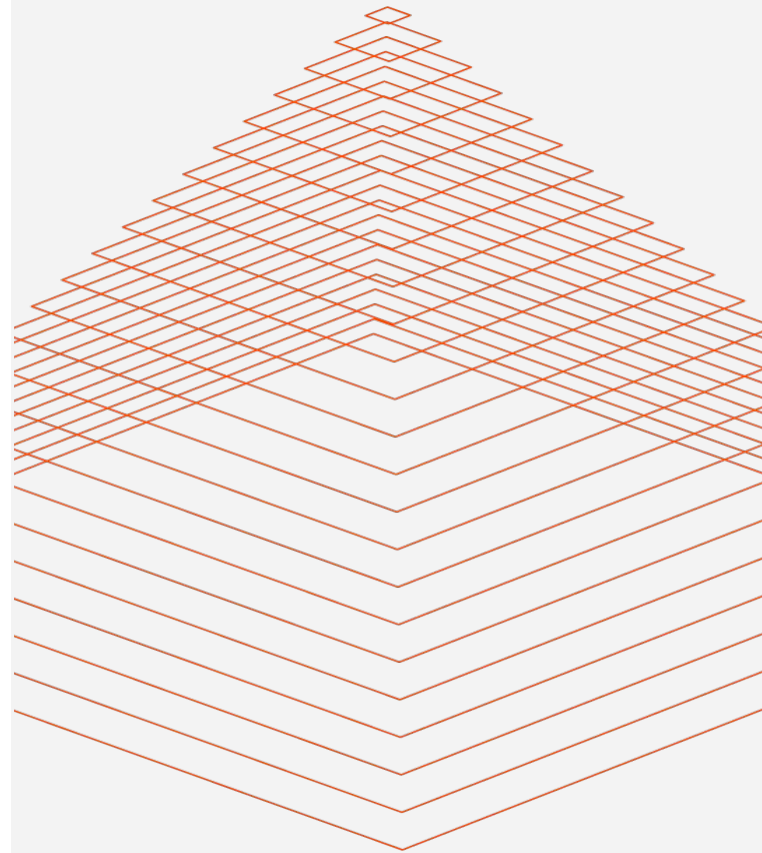
0 Notification Transfer System

- Currently undergoing testing
- Built using the **COAR Notify Protocol** for interoperability and decentralization

0 Communication Fully Operational



0 Upcoming Integrations (in progress)



What's Next ?

1 Expand Interoperability

- Identify and collaborate with new repositories interested in exchanging relationships via the **COAR Notify protocol**
⇒ Toward broader adoption across disciplines and platforms

2 Leverage Machine Learning for Discovery

- Apply machine learning tools to detect mentions of data and software in publications (**GROBID, SoftCite, DataStet**)
- Propose author validation of these automatically detected relationships



Making Software FAIR: A machine-assisted workflow for the research software lifecycle

3 Enrich HAL with External Services

- Integrate external services to supplement HAL metadata with pre-existing relationships



- Accepts validated metadata about links between publications and data
- Enhances discoverability and metadata completeness





HAL
open science

Thank you !

Yannick Barborini

yannick.barborini@ccsd.cnrs.fr



**MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR
ET DE LA RECHERCHE**

*Liberté
Égalité
Fraternité*



Inria

INRAE