

CLEF-HIPE-2020

Named Entity Recognition and Linking on Historical Newspapers

Maud Ehrmann (EPFL)
Matteo Romanello (EPFL)
Alex Flückiger (UZH)
Simon Clemenide (UZH)

CLEF 2020 (online), 23.09.2020

EPFL



Media Monitoring
of the Past

Universität Zürich^{UZH}

CLEF-HIPE-2020 in a nutshell

- **HIPE**: Identifying **H**istorical **P**eople, **P**laces and other **E**ntities
- 1st NE processing shared task on historical documents
- Tasks:
 - NE recognition and classification
 - NE linking
- Participating teams: 13



Why HIPE?

New data:

Emergence of large-scale
archives of digitized contents

New needs:

Content retrieval by
humanities scholars

Challenge: NLP on historical texts is hard

- Spelling variations
- Noisy OCR
- Multilingualism
- Data sparsity
- Limited resources or KB coverage

→ Objectives

1. strengthen the **robustness** of approaches;
2. enable **performance comparison**;
3. foster **efficient semantic indexing** of digitized cultural heritage collections.

Background

impresso project

mining 200 years of historical newspapers

EPFL

Digital Humanities Laboratory (DHLAB)
Ecole Polytechnique Fédérale de Lausanne,
Switzerland



University of Zurich
UZH

Institute of Computational Linguistics Zurich
University, Switzerland



Centre for Contemporary and Digital History
(C2DH) Luxembourg University,
Luxembourg.



Swiss National Library, SNL



National Library of Luxembourg,
BnL



State Archives of Valais, AEV



Swiss Economic Archives, SWA

LE TEMPS

Le Temps

NZZ

Neue Zürcher Zeitung, NZZ

Unil
UNIL | Université de Lausanne

History department, University
of Lausanne, UNIL

infoclio.ch

infoclio



SWISS NATIONAL SCIENCE FOUNDATION



Project: <https://impresso-project.ch>

Interface: <https://impresso-project.ch/app/>

Semantic indexation of historical newspapers

Search NEs
(among others)
over 47M articles

Media Monitoring of the Past

Search Newspapers Topics Inspect & Compare Text reuse Collections

FAQ

Maud Ehrmann STAFF

SEARCH ARTICLES SEARCH IMAGES NGRAMS

otto von Bism

...otto von Bism Search in article contents

...otto von Bism Search in article titles

Otto von Bismarck

prince Otto von Bismarck

Otto von Bismarck

Otto von Bismarck

Otto Christian Archibald von Bismarck

... otto von Bism More Persons ...

2 countries of publication

530,086 named entities disambiguated

More? Check on our [blog](#)

info @ impresso-project [dot] ch

project website: [impresso-project.ch](#)

github: [impresso](#)

twitter: [@impressoproject](#)

LINES: OFF

DARK MODE: ON

Mining 200 years of historical newspapers

How can newspapers help understand the past? How to explore them?

For legal reasons not all content is available in Open Access.
To **gain full access**:

DOWNLOAD NON-DISCLOSURE-AGREEMENT FORM

... and return the signed form to info@impresso-project.ch

Semantic indexation of historical newspapers

Visualize facsimile,
OCR and entity
mentions

impresso Media Monitoring of the Past Search Newspapers Topics Inspect & Compare Text reuse Collections FAQ Maud Ehrmann STAFF

TABLE OF CONTENTS

43 articles in 4 pages (Personal use (no export))

Add filters from your current search query
* 3 search filters can't be applied.

Search words...

show only matching articles (no results)

La Journée politique

M. de Brazza est arrivé lundi à bord du Stamboul, courrier de la côte occidentale d'Afrique. M. de Brazza est parti hier pour Algérie | article | short text | p.1

Continue reading: TRANSCRIPT

LOCATIONS France, Congo Free State, Westminster, Rome, Naples, Italy, Bulgari, Russia

PEOPLE Pierre Savorgnan de Brazza, John Redmond, Ferdinand I of Bulgaria

A la Chambre Séance du 12 février Présidence de M. de Wacquani, présid[...]

A la Chambre Séance du 12 février Présidence de M. de Wacquani, président. Décidément ces tains de nos députés trouvent qu'ils ont le temps pour s'occ | article | short text | pp.2,3 (2 pages)

Continue reading: TRANSCRIPT

LOCATIONS Laval, Mayenne

PEOPLE Victor de Tornaco

Sans titre

Postas. — Un député de notre ville a tonné il y a quelques jours à la Chambre contre la distribution des lettres, qui depuis plusieurs années est fait | article | short text | p.3

L'indépendance luxembourgeoise · Wednesday, February 13, 1895 | p.3 of 4

Faux départ

LOCATIONS Marcel Dallemagne

PEOPLE Otto von Bismarck, Guillaume Cail

TOPICS

42.5% FR gouvernement · parti · ministre · président · politique

18.6% FR roi · prince · empereur · comte · reine

12.2% FR question · point · pays · conseil · gouvernement

9.2% FR vie · esprit · pays · travail · effort

ADD TO COLLECTION ...

CURRENT SELECTION

Otto von Bismarck PERSON

1850 1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000

apply current search filters (3 filters)

317 results in total (from 1857 to 2010)

ADD AS SEARCH FILTER EXCLUDE FROM CURRENT SEARCH

Otto von Bismarck (Q8442)

HUMAN Schönhausen, 1815 - Friedrichsruh, 1898

German statesman and Chancellor (1815-1898)

SOURCE: WIKIDATA W/Q8442

MORE...

NO TEXT REUSE PASSAGES AVAILABLE

Faux

L'opinion publique assez vivement préoccupée de politique qu'impressionne en dépit des démentis

gement de personnes l'empire. Elle note les symptômes qui lui permettent de porter un jugement sur les tendances actuellement prévalantes en haut lieu. Les débats du Reichstag, ceux surtout de la commission chargée d'examiner le projet contre les menées subversives, ont déjà jeté un jour assez vif sur ce point.

en Allemagne est réoccupée du changement politique et qu'entraîne, entis officieux, le chan-

nes à la chancellerie de avec soin tous les i permettent de porter un jugement sur les tendances actuellement prévalantes en haut lieu. Les débats du Reichstag, ceux surtout de la commission chargée d'examiner le projet contre les menées subversives, ont déjà jeté un jour assez vif sur ce point. Bien que le prince de Hohenlohe, fidèle aux traditions d'une vie presque tout entière passée dans la diplomatie

Semantic indexation of historical newspapers

Overview of named entities

Otto von Bismarck

PERSON

OVERVIEW

317 RELATED ARTICLES

MENTIONED 334 TIMES

OPEN IN SEARCH PAGE...

"German statesman and Chancellor (1815-1898)" (wikidata)



W

Otto von Bismarck

German statesman and Chancellor (1815-1898)

Schönhausen, 1815 - Friedrichsruh, 1898

SOURCE: W/Q8442

17401760178018001820184018601880190019201940196019802000

12

COUNTRY

Switzerland 275

Luxembourg 42

LANGUAGE

German 176

French 141

PERSON

Otto von Bismarck 317

Wilhelm II, German Emperor 11

Herbert von Bismarck 10

Karl Marx 9

Adolf Hitler 7

Richard Wagner 7

Debbie Reynolds 5

François Duvalier 5

Frederick III, German Emperor 5

Frederick II of Prussia 5

LOCATION

Berlin 92

Germany 81

Auch 79

Paris 79

Switzerland 68

France 48

England 44

Lage 40

Marcel Dallemagne 37

Italy 36

TYPE

article 244

page article 66

unclassified content 5

advertisement 2

TOPIC

DE regierung · paris · frankreich · minister · kammer 113

FR roi · prince · empereur · comte · reine 87

DE könig · kaiser · königin · prinz · prinzeessin 68

DE verlag · buch · band · geschichte · werk 62

DE krieg · deutschland · frankreich · welt · friede 59

DE welt · leben · mensch · wort · art 43

DE mann · hand · kopf · nacht · gesicht 42

FR gouvernement · parti · ministre · président · politique 42

FR guerre · paix · pays · peuple · politique 40

FR vie · monde · mort · foi · peuple 39

PARTNER

Swiss National Library 157

NZZ 66

ACCESSRIGHT

Personal use 199

Personal use (no export) 80

COLLECTION

« 1 »



7 tide

Semantic indexation of historical newspapers

The screenshot displays the 'impresso' Media Monitoring of the Past interface. The top navigation bar includes 'Media Monitoring of the Past', 'Search', 'Newspapers', 'Topics', 'Inspect & Compare', 'Text reuse', 'Collections', 'FAQ', and a user profile 'Maud Ehrmann'. The main content area is titled 'Otto von Bismarck' with a 'PERSON' tag. It features tabs for 'OVERVIEW', '317 RELATED ARTICLES', and 'MENTIONED 334 TIMES'. Below these are buttons for 'OPEN IN SEARCH PAGE...' and 'ORDER BY PUBLICATION DATE, OLDEST FIRST'. The left sidebar lists various locations with their respective article counts and mention counts: Lausanne (2,918,317 articles, 4,642,691 mentions), Suisse, Moselle (2,561,532 articles, 4,268,837 mentions), Switzerland (2,390,778 articles, 4,727,170 mentions), Fribourg (2,286,590 articles, 4,264,647 mentions), Paris (2,132,568 articles, 3,551,470 mentions), France (2,082,458 articles, 4,036,841 mentions), Gare de Cornavin (1,858,798 articles, 2,822,990 mentions), Lake Neuchâtel (1,791,813 articles, 2,680,593 mentions), La Chaux-de-Fonds (1,411,002 articles, 2,104,854 mentions), Zürich (1,112,003 articles, 2,257,052 mentions), and Italy (1,037,830 articles, 1,607,500 mentions). The main content area shows a list of newspaper articles mentioning Otto von Bismarck, including 'Blättern liest man folgendes „ Gedicht an Otto von Bismarck , ehemaligen Reichskanzler : Der alte Otto ! Vr', 'Vermischtes Bündner Nachrichten', 'Hugo (p . n . k .) 1815 : Geburt Otto von Bismarcks in Schönhausen . Nachmaliger deutscher', 'EDIGENOSSENSCHAFT.Die Tat', 'von Bismarck , der jüngere Bruder dos Grafen Otto von Bismarck und Enkel des Eisernen Kanzlers , wegen', 'Im letzten Stadium.Die Tat', 'Im Schloß Friedrichruh bei Hamburg , wo Fürst Otto von Bismarck begraben liegt , hat Himmler dem schwedischen', 'daß nicht Karl Marx das Rennen macht , sondern Otto von Bismarck , sei es in Gestalt einer neuen Partei oder', 'thv. Paris. 9. Nov. Die nationale Konfe...', 'th. Paris , 9 . Nov . Die nationale Konferenz der Gewerkschaftsfraktion « Force ouvriere » , die am...', '(UP) Der Enkel des « Eisernen Kanzlers » Otto von Bismarck , Graf Gottfried von Bismarck , erlitt am', and 'ab. Der Verunfallte ist auf dem Transport...'. Each article entry includes the newspaper name, date, and page number. A pagination bar at the bottom shows the current page (1) and total pages (2).

Tasks

1. NERC

Recognition and classification of entity mentions with

- subtask 1: coarse types
- subtask 2: fine-grained types.

1	3 L. e	A ££ uroo	
2	M. Piclaon	à fflitadrid	
3	M. Pichon, ministre des affaires		
4	étrangères	, esl arrive' à Madrid	. Aujourd'hui,
5	il ^ dîne chez le roi. Diverses fêtes sont		
6	projetées en son honneur.		

Types	Sub-types	
NERC Coarse	pers	pers.ind pers.coll
	org	org.ent org.adm
	prod	prod.media prod.doctr
	date	time.date.abs
	loc	loc.adm
NERC Fine		
+ entity components		
+ metonymy		
+ nested entities		
		pers.ind.articleauthor
		org.ent.pressagency
	loc.adm	loc.adm.town loc.adm.reg loc.adm.nat loc.adm.sup
	loc.phys	loc.geo loc.hydro loc.astro
	loc.oro	
	loc.fac	
	loc.add	loc.add.phys loc.add.elec

Table 1: Entity types used for NERC tasks.

Tasks

2. Entity Linking

Towards Wikidata QID or NIL



- end-to-end EL: w/o mention boundaries
- EL-only: with mention boundaries



The screenshot shows a text snippet with several entity mentions highlighted in yellow. A pop-up window displays the details for the entity 'M. Pichon, ministre des affaires étrangères' (ID: 13255). The details include the Wikidata ID, the label, and the description: 'French politician (1857-1933)'.

Impresso Named Entity ID: 13255
wikidata_id: http://www.wikidata.org/entity/Q1069292
Label: Stephen Pichon
Description: French politician (1857-1933)

Participation bundles:

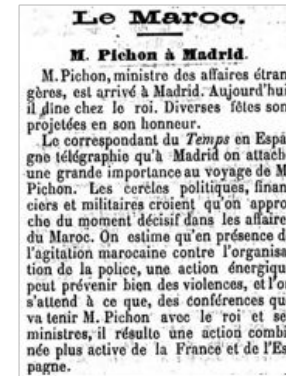
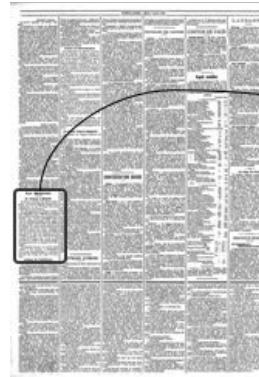
Bundle id	Associated tasks
bundle1	NERC-coarse and NERC-fine and NEL
bundle2	NERC-coarse and NEL
bundle3	NERC-coarse and NERC-fine
bundle4	NERC-coarse
bundle5	NEL

Participation guidelines:

[10.5281/zenodo.3677171](https://zenodo.org/record/3677171)

Corpus selection

- Digitized newspaper archives (CH, LU, US)
- Diachronic: from 1738 to 2019
- Multilingual: fr, de, en
- Sampling and manual triage:
 - journalistic content
 - no feuilleton, cross-words, meteo, etc.
 - exclusion of extreme OCR noise
 - no provision of different OCR → real-life setting



3 L. e A EE uroo.

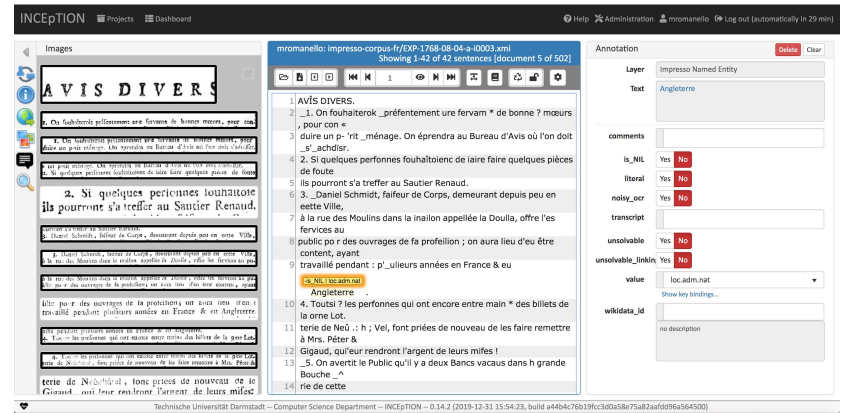
M. Pichon à Madrid.

M. Pichon, ministre des affaires étrangères, est arrivé à Madrid. Aujourd'hui, il a dîné chez le roi. Diverses fêtes sont projetées en son honneur.

Le correspondant du Temps en Espagne télégraphie qu'à Madrid on attache une grande importance au voyage de M. Pichon. Les cercles politiques, financiers et militaires croient qu'on approche du moment décisif dans les affaires du Maroc. On estime qu'en présence de l'agitation marocaine contre l'organisation de la police, une action énergique peut prévenir bien des violences, et l'on s'attend à ce que, des Conférences que va tenir M. Pichon avec le roi et ses ministres, il résulte une action combinée plus active de la France et de l'Espagne.

Corpus annotation

- Trilingual annotators, trained on a mini-ref
- INCEpTION platform
- NERC annotation difficulties:
 - NE mention boundaries
 - consideration of multiple languages
 - what is to be annotated or not
 - definition at time x
 - metonymy



*M. Curtoys d' Anduaga, doyen du corps diplojt d'elfsue
espagnol, et ministre plenipotentielle pendant 50 ans*

Zürichputsch, Baslerpropaganda

Commission imperiale, Die französische Regierung

Is Savoie or Moldavia a region or a country?

Annotation guidelines

[10.5281/zenodo.3604227](https://zenodo.org/record/3604227)

Corpus annotation

- Trilingual annotators, trained on a mini-ref
- INCEpTION platform
- NERC annotation difficulties:
 - NE mention boundaries
 - consideration of multiple languages
 - what is to be annotated or not
 - definition at time x
 - metonymy
- EL annotation difficulties:
 - Requires historical knowledge + Sherlock Holmes skills
 - Historical statuses of entities unequally represented in KB

Germany, Q183

962-1813: Holy Roman Empire, Q12548

1806-1813: Confederation of the Rhine, Q154741

1815-1866: German Confederation, Q151624

1867-1870: North German Confederation, Q150981

1871-1918: German Empire, Q43287

1918-1933: Weimar Republic, Q41304

1933-1945: Nazi Germany, Q7318

1949-1990: West Germany, Q713750

1949-1990: East Germany, Q16957

Annotation guidelines

[10.5281/zenodo.3604227](https://doi.org/10.5281/zenodo.3604227)

Corpus characteristics

newspaper articles	563
tokens	444,596
(linked) mentions	18,962
metonymy	1252
components	6,219
noisy mentions (test set)	10%
NIL	25.72%

mentions: 10,923 (Fr), 6584 (De), 1455 (En)

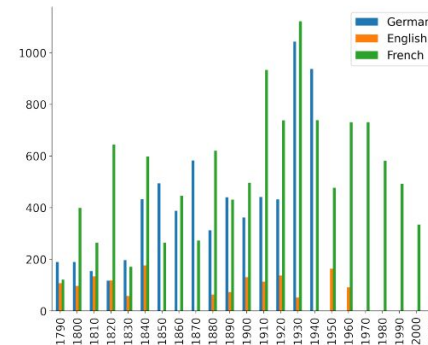


Fig. 2: Diachronic distribution of mentions across languages.

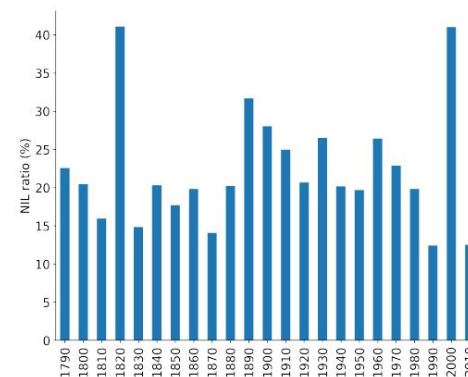


Fig. 3: Diachronic ratio of NIL entities.

Corpus release

- train/dev/test (70/15/15)
- no train set for English
- no sentence segmentation
- no sophisticated tokenization
- document metadata

TOKEN	NE-COARSE-LIT	NE-COARSE-METO	NE-FINE-LIT	NE-FINE-METO	NE-FINE-COMP	NE-NESTED	NEL-LIT	NEL-METO	MISC
# language = fr									
# newspaper = GDL									
# date = 1908-01-07									
# document_id = GDL-1908-01-07-a-i0009									
[...]									
# segment_iif_link = https://iif.dhlab.epfl.ch/iif_impresso/GDL-1908-01-07-a-p0002/296,3651,514,64/full/0/default.jpg									
M	B-pers	0	B-pers.ind	0	B-comp.title	0	Q1069292	-	NoSpaceAfter
.	I-pers	0	I-pers.ind	0	I-comp.title	0	Q1069292	-	-
Piclaon	I-pers	0	I-pers.ind	0	B-comp.name	0	Q1069292	-	-
à	0	0	0	0	0	0	-	-	-
ffltadrid	B-loc	0	B-loc.adm.town	0	0	0	Q2807	-	NoSpaceAfter
.	0	0	0	0	0	0	-	-	EndOfLine
# segment_iif_link = https://iif.dhlab.epfl.ch/iif_impresso/GDL-1908-01-07-a-p0002/213,3709,731,64/full/0/default.jpg									
M	B-pers	0	B-pers.ind	0	B-comp.title	0	Q1069292	-	NoSpaceAfter
.	I-pers	0	I-pers.ind	0	I-comp.title	0	Q1069292	-	-
Pichon	I-pers	0	I-pers.ind	0	B-comp.name	0	Q1069292	-	NoSpaceAfter
,	I-pers	0	I-pers.ind	0	0	0	Q1069292	-	-
ministre	I-pers	0	I-pers.ind	0	B-comp.function	0	Q1069292	-	-
des I-pers	0	I-pers.ind	0	I-comp.function	0	Q1069292	-	-	-
affaires	I-pers	0	I-pers.ind	0	I-comp.function	0	Q1069292	-	EndOfLine
# segment_iif_link = https://iif.dhlab.epfl.ch/iif_impresso/GDL-1908-01-07-a-p0002/171,3753,770,64/full/0/default.jpg									
étrangères	I-pers	0	I-pers.ind	0	I-comp.function	0	Q1069292	-	NoSpaceAfter
,	0	0	0	0	0	0	-	-	-
esl	0	0	0	0	0	0	-	-	-
arrive	0	0	0	0	0	0	-	-	NoSpaceAfter
'	0	0	0	0	0	0	-	-	-
à	0	0	0	0	0	0	-	-	-
Madrid	B-loc	0	B-loc.adm.town	0	0	0	Q2807	-	NoSpaceAfter
.	0	0	0	0	0	0	-	-	-
Aujourd	0	0	0	0	0	0	-	-	NoSpaceAfter
'	0	0	0	0	0	0	-	-	NoSpaceAfter
hui	0	0	0	0	0	0	-	-	NoSpaceAfter
,	0	0	0	0	0	0	-	-	EndOfLine

CC BY-NC 4.0

<https://github.com/impresso/CLEF-HIPE-2020/tree/master/data>
[10.5281/zenodo.3706857](https://zenodo.org/record/3706857)

Auxiliary resources

In-domain Fr, De, and En embeddings:

- fastText word embeddings (with and w/o subwords)
- flair character embeddings (now integrated into the flair [framework](#))

CC BY-SA 4.0

[https://files.ifi.uzh.ch/cl/siclemat/impresso/clef-hipe-2020/
10.5281/zenodo.3706808](https://files.ifi.uzh.ch/cl/siclemat/impresso/clef-hipe-2020/10.5281/zenodo.3706808)

Evaluation

- Entities (not tokens) as the unit of reference
- Macro & **Micro** Precision, Recall and F1 measure
- Evaluation scenarios:

	NERC	EL
Strict	exact mention boundaries	consideration of the top link only, (overlapping mention boundaries)
Fuzzy	overlapping boundaries	historical mapping, cut-offs @3 and @5 (overlapping mention boundaries)

HIPE Scorer: <https://github.com/impresso/CLEF-HIPE-2020-scorer>

HIPE Eval Toolkit: <https://github.com/impresso/CLEF-HIPE-2020-eval>

Participation

40

registrations

75 runs

42% French

31% German

26% English

6 teams work on all languages

13

participating teams

All participated to NERC-Coarse

3 to NERC-Fine

5 to EL-only and end-to-end EL

11

Working Notes

Participating systems' main features

- 11 teams applied **neural approaches** for NERC;
- Most of them worked with **contextualized embeddings**, esp. **BERT**;
- Experimentation with **various input embeddings** (char, subword, word, historical or contemporary, type-level or contextualized)
- Some attempted to improve the **newspaper line-based input format** with proper sentence segmentation and tokenization;

Results overview (NERC)

- Neural system with strong embedding resource prevail;
- Performances correlates with amount of train/dev data;
- BERT-based systems > Bi-LSTM;
- Great performances diversity, but results are better than expected (6 teams > .8);
- NERC fine with 12 classes more difficult;
- NE components show reasonable performances.

F1 scores	French		German		English	
	Strict	Fuzzy	Strict	Fuzzy	Strict	Fuzzy
NERC-Coarse literal						
Baseline	.646	.769	.476	.585	.405	.562
Median	.677	.808	.636	.766	.463	.645
Best system	.840	.921	.797	.878	.632	.806
NERC-Coarse metonymic						
Best system	.783	.783	.634	.694	-	-
NERC-Fine						
Best system	.784	.856	.668	.771	-	-
NE components						
Best system	.657	.751	.642	.707	-	-

Results overview (EL)

- EL performances are lower, and as diverse;
- NERC error propagation in end-to-end setting, but EL-only not a lot better;
- Performance increase with cut-offs @3 and @5.

Overall, what helps:

- BERT;
- actively tackling the problems of OCR noise, word hyphenation and sentence segmentation;
- in-domain resources.

F1 scores	French		German		English	
	Strict	Fuzzy	Strict	Fuzzy	Strict	Fuzzy
End-to-end Entity Linking (literal)						
Baseline	.257	.270	.180	.195	.239	.239
Best system	.598	.617	.534	.557	.531	.531
End-to-end Entity Linking (metonymic)						
Best system	.297	.462	.396	.469	-	-
Entity linking only (with mentions provided)						
Baseline	.498	.512	.418	.437	.506	.506
Best system	.639	.659	.582	.602	.658	.658

Time-based observations

Analysis of F1 score as a function of time.

Hypothesis: the older, the more difficult.

Observation: no strong correlation between article publication date and performance.

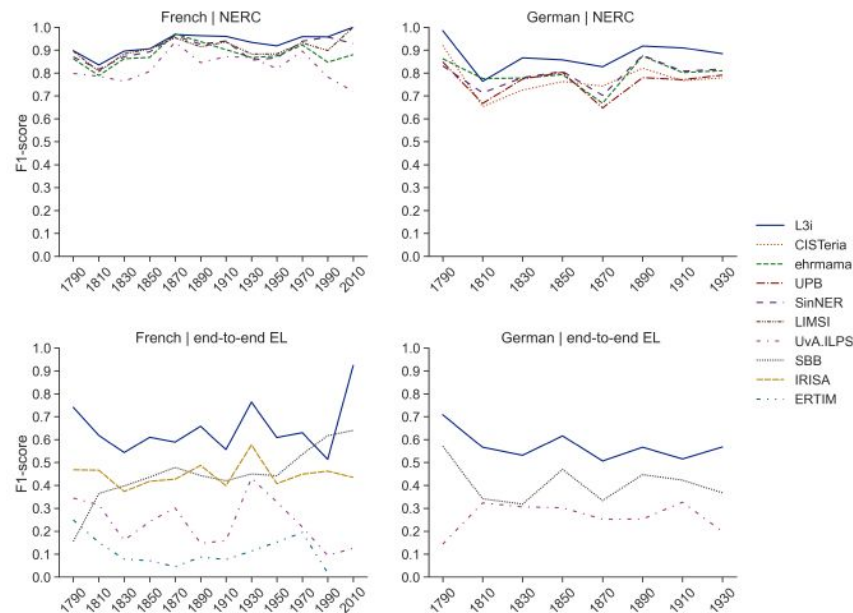
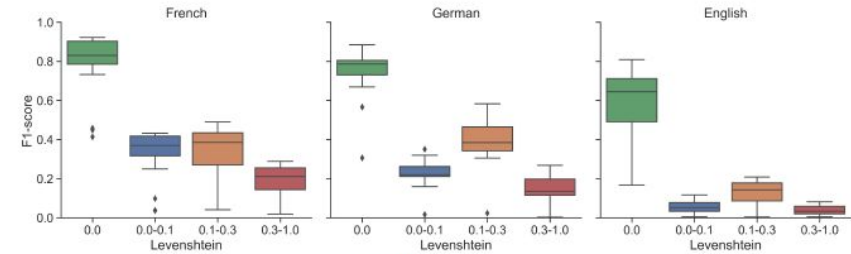


Fig. 4: F1 score as a function of time for the 5 best systems for NERC (top) and end-to-end EL (bottom) for the languages French (left) and German (right). The x-axis shows 20-years time buckets (e.g. 1790 = 1790-1809).

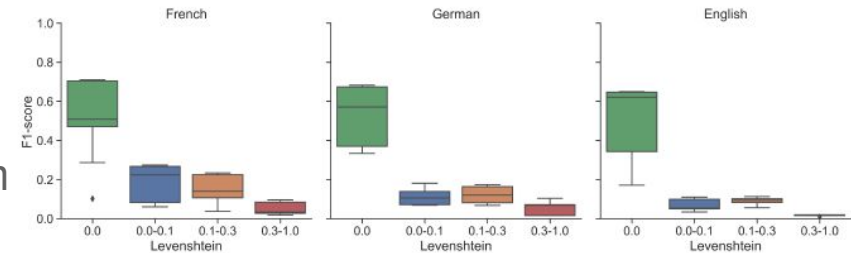
Impact of OCR noise

Evaluation on various noise levels

- noise: length-normalized Levenshtein distance between surface form and manual transcription;
- noisy vs non-noisy have remarkable differences on both NERC and EL;
- greatest performance variation at medium noise level



(a) NERC-Coarse.



(b) End-to-end EL with the relaxed evaluation regime and a cutoff @3.

Fig. 5: Impact of OCR noise: distribution of performances across systems on entities with different noise level severity for NERC (a) and end-to-end EL (b).

Conclusion (1/2)

- **Robustness test** for NERC and EL approaches on challenging historical material;
- New insights in **domain and language adaptation**;
- **Neural-based systems** with strong resources and proper segmentation are capable of dealing with historical and noisy inputs;
- EL, nested entities, entity components remain challenging;
- Performances are affected by OCR noise, but not by document publication date;

Discover more about systems this afternoon 3-6:30pm!

Conclusion (2/2)

Main Outcomes:

- Contribution to the advance of SoTA for NE processing on historical texts;
- Datasets;
- Scorer;
- Annotation guidelines;
- Step towards efficient semantic indexing of historical material.

Future Directions:

- Potential HIPE2 in 2021 with additional document types and languages.

Many thanks to

- CLEF organizers
- Participating teams (kudos!)
- *NZZ*, *Le Temps*, and the Swiss and Luxembourg national libraries
- Richard Eckart de Castillo, Clemens Neudecker, Sophie Rosset and David Smith
- INCEpTION project team
- Camille Watter, Gerold Schneider, Emmanuel Decker and Ilaria Comes
- SNSF (grant number CR-SII5 173719)

Thank you for your attention



HIPE : impresso.github.io/CLEF-HIPE-2020/



Scorer : impresso.github.io/CLEF-HIPE-2020/



Evaluation toolkit : impresso.github.io/CLEF-HIPE-2020/



Impresso project : impresso.github.io/CLEF-HIPE-2020/



@ImpressoProject