

# Session 2

## CorText Manager et ses méthodes

CorText Training Sessions, Janvier 2021

Lionel VILLARD

Marc BARBIER

Alexandre HANNUD ABDO

LISIS, IFRIS, INRAE, CorText, ESIEE Paris

# Présentation de CorText

Constituer une **plateforme scientifique et technique** avec ses exigences propres de positionnement liées à son existence et à son devenir pour soutenir un **espace de recherche sur les infrastructures, les traces et les usages numérique de la sciences et de l'innovation en société.**

2008	Equipes INRA Praxis et TSV		GIS IFRIS Projet PHARE CNRS
2010	UR INRA SenS	Projet SAD ESR sur Plateforme	GIS IFRIS LABEX SITES
2015	UMR LISIS	Ingénieur INRA	LABEX SITES
2019	UMR LISIS	Ingénieur INRA Ingénieur ESIEE	RISIS-1 RISIS-2

# Présentation de CorText





**Marc Barbier**

Member of CorText platform,  
Researcher at LISIS, Head of IFRIS



**Antoine Schoen**

Member of CorText platform,  
Researcher at LISIS, Senior lecturer  
at ESIEE Paris



**Lionel Villard**

Head of CorText platform, Researcher  
at LISIS, lecturer at ESIEE Paris



**Patricia Laurens**

Member of CorText platform,  
Researcher at CNRS and LISIS



**Philippe Breucker**

IT engineer from INRAE, LISIS.  
Technical Director of the CorText  
Digital Platform. Web Designer and  
developer.



**Bilel Benbouzid**

Researcher, Senior lecturer at LISIS



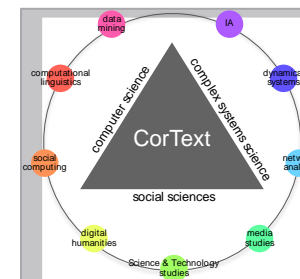
**Alexandre Hannud Abdo**

Post-doctorant, LISIS



**Juan-Pablo Ospina Delgado**

IT Engineer, Cortext



**Luis-Daniel Medina**

IT Engineer, Cortext



**Pierre-Yves Bulot**

IT Engineer Assistant, Cortext



**Diego-Fernando Gómez Peña**

IT Engineer, Cortext



**Tatiana Andrea Sánchez Castaño**

IT Engineer, Cortext



**Joenio Marques da Costa**

Research Software Engineer, Cortext



**Antoine Mazières**

Research scientist in the Computation  
Social Science team at Centre Marc  
Bloch



**Constance De Quatrebarbes**

Fondatrice Présidente - DRISS  
(Digital Research In Science &  
Society)



**Chloé Duloquin**

Web Designer, Graphiste, Intégratrice  
web



**Jean-Philippe Cointet**

Associate Professor, Sciences Po  
Paris, Medialab



**Guillaume Orsal**

Computer engineer, data mining, web  
development and SEO



**Cristian Martinez**

PhD Engineer in Computer Science,  
NLP/Data Senior Consultant at  
Cogniteva



**Nicolas Turenne**

Assistant professor in data science,  
Beijing Normal University & Hong  
Kong Baptist University United  
International College



**Tam Kien Duong**

Data & design, Etalab



**Nicolas Baya-Laffite**

STSLab, Université de Lausanne



**Loïc Boudoulec**

IT Engineer



**Bertha Brenes**

IT Engineer, Trainee, Cortext



**Anis Arabi**

Big data engineer



**Nicolas Ricci**

Web developer and data



**Audrey Baneyx**

Project manager Data science,  
Sciences Po Paris - medialab



**Andrei Mogoutov**

Bullescence



**Élise Tancoigne**

Researcher, University of Geneva,  
Switzerland

# Sessions : programme



Session 1 : de la scientométrie à CorText Manager

Session 2 : CorText Manager et ses méthodes

Session 3 : Sashimi et corpus des participants

# Session 2 : programme

## Plénière welcome 30'

## Plénière, 14h

*Corpus science Chloroquine sur 7 ans – vos – articles 2014-2020*

*Travailler le vocabulaire*

*Réseaux homogènes et hétérogène, Mesures de proximité*

*Co-citations*

*Evolution du paysage sémantique*

## Sessions parallèles, 15h30

### Distribution(s) et graphique(s) (Ale)

- Temps : area rank charts
- Temps : repérer des périodes
- Dégager la structure d'ensemble
- Thèmes : topic modeling

### Communautés épistémiques (Marc)

- Auteurs, comptage des distributions et communautés d'auteurs
- Eclairages par pays et fundings
- Contrast et profiling d'auteurs
- Clusters de termes
- Réseaux hétérogènes
- Réseaux socio-sémantiques

### Les lieux de la recherche (Lionel)

- Adresses d'auteurs et géocoding
- Travailler les localisations
- Agrégation géographique, aires urbaines et aires rurales
- Cartographier les résultats

## Plénière, 18h

*Questions ? Réponses !*

*Préparer la session 3*

# Les mesures de proximité

proximity measures	type of network	normalisation	special properties
raw	interaction network (e.g. social network)	no	-
$\chi^2$	homogeneous & heterogeneous	yes	normalization tend to create links toward higher degree nodes
MI	homogeneous & heterogeneous	yes	Inspired from information theory
Cramer	homogeneous & heterogeneous	yes	-
cosine	homogeneous network (eg. semantic)	yes	Classical measure (originating from scientometrics)
distributional	homogeneous network (eg. semantic)	yes	very robust measure (coming from computational linguistics)
cosine_het	affiliation network (eg. users sharing the same hashtags )	yes	two fields are required but the final network is homogeneous
dot_product_het	affiliation network (eg. users sharing the same hashtags )	no	two fields are required but the final network is homogeneous

# A vous de jouer

<https://docs.cortext.net/trainings/cortext-lisis/>

Et

<https://managerv2.cortext.net/project/101440003210>