

# *Contributions à l'épidémiosurveillance végétale*

*Constitution d'un lexique multilingue  
pour l'extraction d'information à partir de textes*

Elisa Lubrini (INRAE Bibliome MaIAGE)

Robert Bossy (Bibliome, MaIAGE)

Marie Grosdidier (INRAE, BioSP et Plateforme ESV)

Claire Nédellec (Bibliome, MaIAGE)

# PESV

Plateforme d'Épidémiosurveillance en Santé  
Végétale (ESV)

One Health\* = balancing and optimising the  
health of people, animals and the environment.

Objectives:

Active platform supporting the agricultural  
industry



\* [WHO](#)

# Ressource Lexicale

EPPO_CODE	EFSA_PHT	NCBI_TAXID	preferred_name	synonym	re	ds_language
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	peach borer		en
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	cerambicide dal collo rosso	cerambicid\w dal collo rosso	it
CRSPAN	AnisogrammaAnomala-PHT		Anisogramma anomala	blight of hazel	blight of hazel	en
CRSPAN	AnisogrammaAnomala-PHT		Anisogramma anomala	eastern blight of filbert		en
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	Aromia bungii		
XYLEFA		ncbi:2371	Xylella fastidiosa	Xylella fastidiosa subsp. piercei		

# Assisting Technical Translations

. Out of the six species that were analyzed, **crotalaria**, **forage radish** and **giant stock-beans** were identified as potential hosts of the pathogen. 1

Document à annoter. *Giant  
stock-beans* absent de la  
taxonomie

# Assisting Technical Translations

. Out of the six species that were analyzed, crotalaria, forage radish and giant stock-beans were identified as potential hosts of the pathogen. 7

Document à annoter. *Giant stock-beans* absent de la taxonomie

usando-se a bananeira, também conhecida como mar-de-afonso, planta mais destrutiva dessa cultura no mundo. Das seis espécies avaliadas, crotalaria, nabo forrageiro e feijão de porco foram apontadas como potenciais hospedeiras do patógeno. A descoberta pode contribuir para otimizar as estratégias de manejo de

Document d'origine. *Feijão de porco* est le terme en portugais

# Assisting Technical Translations

Out of the six species that were analyzed, **crotalaria**, **forage radish** and **giant stock-beans** were identified as potential hosts of the pathogen. T

Document à annoter. *Giant stock-beans* absent de la taxonomie

usando a canavalia, também conhecida como mar-dei-anã, que é uma planta destrutiva dessa cultura no mundo. Das seis espécies avaliadas, crotalária, nabo forrageiro e **feijão de porco** foram apontadas como potenciais hospedeiras do patógeno. A descoberta pode contribuir para otimizar as estratégias de manejo de

Document d'origine. *Feijão de porco* est le terme en portugais

NCBI_TAXID	preferred_name	synonym	ds_language
ncbi:3823	Canavalia ensiformis	Feijão de porco	pt

...

# Helping Annotators

&quot;A case of **Xylella fastidiosa** on a **wild almond tree**&quot; - Tusciaweb.eu



"A case of **Xylella fastidiosa** on a **wild almond tree**". Canino - The mayor Giuseppe Cesetti and the president of the **fruit** and **vegetable** cooperative Giovanni Pettinari: "The **tree** is located in the area of the **Vulci** castle"



# NLP Pipelines

- 1** Assembling 3 resources into a single complete LR
- 2** Translation via LR
- 3** Annotation via LR



## Pipeline 1 Assembling the LR



EPPO

European and Mediterranean Plant Protection Organization

**Resource Type** - Glossary

**Information** - synonyms, languages, scientific name



EFSA

European Food Safety Authority

**Resource Type** - Glossary

**Information** - synonyms



NCBI

National Center of Biotechnology Information

**Resource Type** - Ontology

**Information** - synonyms, scientific name, hierarchy

EPPO			
EPPO_CODE	preferred_name	synonym	ds_language
AROMBU	Aromia bungii	peach borer	en
AROMBU	Aromia bungii	cerambicide dal collo rosso	it
CRSPAN	Anisogramma anomala	blight of hazel	en
CRSPAN	Anisogramma anomala	eastern blight of filbert	en
AROMBU	Aromia bungii	Aromia bungii	
XYLEFA	Xylella fastidiosa	Xylella fastidiosa subsp. piercei	

NCBI		
Scientific name	synonym	TAXID
Xylella fastidiosa	X . fastidiosa	2371
Aromia bungii	Aromia bungii	320466
Aromia bungii	A . bungii	320466
Acalymma vittatum	striped cucumber beetle	133117

EFSA			
EFSA_PHT	preferred_name	re	synonym
AromiaBungii-PHT	aromia bungii	cerambicid\w dal collo rosso	
AnisogrammaAnomala-PHT	anisogramma anomala	blight of hazel	
AcalymmaVittatum-PHT	acalymma vittatum	striped cucumber beetle\w*	striped cucumber beetle

## EPPO (+ NCBI)

EPPO_CODE	preferred_name	synonym	ds_language	NCBI_TAXID
AROMBU	Aromia bungii	peach borer	en	<a href="#">ncbi:320466</a>
AROMBU	Aromia bungii	cerambicide dal collo rosso	it	<a href="#">ncbi:320466</a>
CRSPAN	Anisogramma anomala	blight of hazel	en	
CRSPAN	Anisogramma anomala	eastern blight of filbert	en	
AROMBU	Aromia bungii	Aromia bungii		<a href="#">ncbi:320466</a>
XYLEFA	Xylella fastidiosa	Xylella fastidiosa subsp. piercei		<a href="#">ncbi:2371</a>

## EFSA (+ NCBI)

EFSA_PHT	preferred_name	re	synonym	NCBI_TAXID
AromiaBungii-PHT	aromia bungii	cerambicid\w dal collo rosso		<a href="#">ncbi:320466</a>
AnisogrammaAnomala-PHT	anisogramma anomala	blight of hazel		
AcalymmaVittatum-PHT	acalymma vittatum	striped cucumber beetle\w*	striped cucumber beetle	<a href="#">ncbi:133117</a>

## EPPO (+ NCBI)

EPPO_CODE	preferred_name	synonym	ds_language	NCBI_TAXID
AROMBU	Aromia bungii	peach borer	en	ncbi:320466
AROMBU	Aromia bungii	cerambicide dal collo rosso	it	ncbi:320466
CRSPAN	Anisogramma anomala	blight of hazel	en	
CRSPAN	Anisogramma anomala	eastern blight of filbert	en	
AROMBU	Aromia bungii	Aromia bungii		ncbi:320466
XYLEFA	Xylella fastidiosa	Xylella fastidiosa subsp. piercei		ncbi:2371

## EFSA (+ NCBI)

EFSA_PHT	preferred_name	re	synonym	NCBI_TAXID
AromiaBungii-PHT	aromia bungii	cerambicid\w dal collo rosso		ncbi:320466
AnisogrammaAnomala-PHT	anisogramma anomala	blight of hazel		
AcalymmaVittatum-PHT	acalymma vittatum	striped cucumber beetle\w*	striped cucumber beetle	ncbi:133117

EPPO + EFSA + NCBI

EPPO_CODE	EFSA_PHT	NCBI_TAXID	preferred_name	synonym	re	ds_language
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	peach borer		en
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	cerambicide dal collo rosso	cerambicid\w dal collo rosso	it
CRSPAN	AnisogrammaAnomala-PHT		Anisogramma anomala	blight of hazel	blight of hazel	en
CRSPAN			Anisogramma anomala	eastern blight of filbert		en
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	Aromia bungii		
XYLEFA		ncbi:2371	Xylella fastidiosa	Xylella fastidiosa subsp. piercei		
	AcalymmaVittatum-PHT	ncbi:133117		striped cucumber beetle	striped cucumber beetle\w*	

EPPO + EFSA + NCBI

EPPO_CODE	EFSA_PHT	NCBI_TAXID	preferred_name	synonym	re	ds_language
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	peach borer		en
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	cerambicide dal collo rosso	cerambicid\w dal collo rosso	it
CRSPAN	AnisogrammaAnomala-PHT		Anisogramma anomala	blight of hazel	blight of hazel	en
CRSPAN			Anisogramma anomala	eastern blight of filbert		en
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	Aromia bungii		
XYLEFA		ncbi:2371	Xylella fastidiosa	Xylella fastidiosa subsp. piercei		
	AcalymmaVittatum-PHT	ncbi:133117		striped cucumber beetle	striped cucumber beetle\w*	

# EPPO + EFSA + NCBI

EPPO_CODE	EFSA_PHT	NCBI_TAXID	preferred_name	synonym	re	ds_language
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia			en
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	cerambicide dal collo rosso	cerambicid\w dal collo rosso	it
CRSPAN	AnisogrammaAnomala-PHT		Anisogramma anomala	blight of hazel	blight of hazel	en
CRSPAN			Anisogramma anomala	eastern blight of filbert		
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia Bungii	Aromia bungii		
XYLEFA		ncbi:2371	Xylella fastidiosa	Xylella fastidiosa subsp. mercei		
	AcalymmaVittatum-PHT	ncbi:133117		striped cucumber beetle	striped cucumber beetle\w*	

RegEx match between EFSA and EPPO

cerambicide dal collo rosso

cerambicid\w dal collo rosso

Same NCBI ID in EFSA and EPPO

SAME ENTITY IF EITHER...



## EPPO + EFSA + NCBI

EPPO_CODE	EFSA_PHT	NCBI_TAXID	preferred_name	synonym	re	ds_language
-----------	----------	------------	----------------	---------	----	-------------

AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	peach borer		en
--------	------------------	-------------	---------------	-------------	--	----

AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	cerambicide dal collo rosso	cerambicid\w dal collo rosso	it
--------	------------------	-------------	---------------	-----------------------------	------------------------------	----

**SAME ENTITY  
DIFFERENT ROWS FOR SEPARATE TERMS**

CRSPAN						en
CRSPAN						en

AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	Aromia bungii		
--------	------------------	-------------	---------------	---------------	--	--

XYLEFA		ncbi:2371	Xylella fastidiosa	Xylella fastidiosa subsp. piercei		
--------	--	-----------	--------------------	-----------------------------------	--	--

	AcalymmaVittatum-PHT	ncbi:133117		striped cucumber beetle	striped cucumber beetle\w*	
--	----------------------	-------------	--	-------------------------	----------------------------	--



## EPPO + EFSA + NCBI

EPPO_CODE	EFSA_PHT	NCBI_TAXID	preferred_name	synonym	re	ds_language
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	peach borer		en
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	cerambicide dal collo rosso	cerambicid\w dal collo rosso	it
CRSPAN	AnisogrammaAnomala-PHT		Anisogramma anomala	blight of hazel	blight of hazel	en
CRSPAN			Anisogramma anomala	eastern blight of filbert		en
AROMBU						
XYLEFA		ncbi:2371	Xylella fastidiosa	Xylella fastidiosa subsp. piercei		
	AcalymmaVittatum-PHT	ncbi:133117		striped cucumber beetle	striped cucumber beetle\w*	

**DIFFERENT ENTITIES  
DIFFERENT ROWS FOR SEPARATE TERMS**

EPPO + EFSA + NCBI

EPPO_CODE	EFSA_PHT	NCBI_TAXID	preferred_name	synonym	re	ds_language
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	peach borer		en
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	cerambicide dal collo rosso	cerambicid\w dal collo rosso	it
CRSPAN	AnisogrammaAnomala-PHT		Anisogramma anomala	blight of hazel	blight of hazel	en
CRSPAN			Anisogramma anomala	eastern blight of filbert		en
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	Aromia bungii		
XYLEFA		ncbi:2371	Xylella fastidiosa	Xylella fastidiosa subsp. piercei		
	AcalymmaVittatum-PHT	ncbi:133117		striped cucumber beetle	striped cucumber beetle\w*	

# EPPO + EFSA + NCBI

EPPO_CODE	EFSA_PHT	NCBI_TAXID	preferred_name	synonym	re	ds_language
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	peach borer		en
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	cerambicide dal collo rosso	cerambicid\w dal collo rosso	it
CRSPAN	AnisogrammaAnomala-PHT		Anisogramma anomala	blight of hazel	blight of hazel	en
CRSPAN	AnisogrammaAnomala-PHT		Anisogramma anomala	eastern blight of filbert		en
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	Aromia bungii		
XYLEFA		ncbi:2371	Xylella fastidiosa	Xylella fastidiosa subsp. piercei		
	AcalymmaVittatum-PHT	ncbi:133117		striped cucumber beetle	striped cucumber beetle\w*	

CRSPAN  
CRSPAN



EPPO + EFSA + NCBI

EPPO_CODE	EFSA_PHT	NCBI_TAXID	preferred_name	synonym	re	ds_language
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	peach borer		en
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	cerambicide dal collo rosso	cerambicid\w dal collo rosso	it
CRSPAN	AnisogrammaAnomala-PHT		Anisogramma anomala	blight of hazel	blight of hazel	en
CRSPAN	AnisogrammaAnomala-PHT		Anisogramma anomala	eastern blight of filbert		en
AROMBU	AromiaBungii-PHT	ncbi:320466	Aromia bungii	Aromia bungii		
XYLEFA		ncbi:2371	Xylella fastidiosa	Xylella fastidiosa subsp. piercei		
	AcalymmaVittatum-PHT	ncbi:133117		striped cucumber beetle	striped cucumber beetle\w*	

# LR Usage

## Corpus

[...] Il cerambicide dal collo rosso è originario del Nord est asiatico, da dove si è diffuso raggiungendo anche l'Europa. [...]



## Lexical Resource

NCBI_TAXID	preferred_name	synonym	ds_language
ncbi:320466	Aromia bungii	cerambicide dal collo rosso	it
	...		

## Pipeline 2 - Annotation

[...] Il ncbi:320466 cerambicide dal collo rosso è originario del Nord est asiatico, da dove si è diffuso raggiungendo anche l'Europa. [...]

## Pipeline 3 - Translation

[...] The **Aromia bungii** is native to Northeast Asia, from where it has spread reaching Europe as well. [...]

# Future Work

- Allow multiple languages
- Detect more languages via term search on bigger corpus
- Translation with glossary

# Thank You

Get in touch: [elisa.lubrini@inrae.fr](mailto:elisa.lubrini@inrae.fr)

Check out the pipelines: [gitlab.com/elisa.lubrini/quarantine-organisms-kr-nlp](https://gitlab.com/elisa.lubrini/quarantine-organisms-kr-nlp)