



HAL
open science

Consignes d'annotation du corpus EPOP (Epidemiomonitoring Of Plant)

Claire Nédellec, Catalina Garcia, Elisa Lubrini, Marie Grosdidier, Louise Deléger, Robert Bossy, Sandy Dupérier, Clara Sauvion, Isabelle Pieretti

► **To cite this version:**

Claire Nédellec, Catalina Garcia, Elisa Lubrini, Marie Grosdidier, Louise Deléger, et al.. Consignes d'annotation du corpus EPOP (Epidemiomonitoring Of Plant). 2024. hal-04744335

HAL Id: hal-04744335

<https://hal.science/hal-04744335v1>

Submitted on 18 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Consignes d'annotation du corpus EPOP (*Epidemiomonitoring Of Plant*)

Version - 1.3

18 octobre 2024

Auteurs

Claire Nédellec, Catalina Garcia, Elisa Lubrini, Marie Grosdidier, Louise Deléger,
Robert Bossy, Sandy Duperier, Clara Sauvion, Isabelle Pieretti

Destinataires

Experts annotateurs impliqués dans la campagne d'annotation
Utilisateurs du corpus annoté EPOP, notamment pour l'extraction d'information

Copyright 2024 by INRAE.



Le document *Consignes d'annotation du corpus EPOP (Epidemiomonitoring Of Plant)* est distribué sous les termes de la Creative Commons Attribution-ShareAlike 4.0 License (CC-BY-SA). Vous trouverez le texte officiel (en anglais) ici : <http://creativecommons.org/licenses/by-sa/4.0/>

Sommaire

1. Contenu du document	2
2. Annotation des mentions d'entités.....	3
2.1. Listes des types d'entité	3
2.2. Consignes générales d'annotation des entités.....	3
2.2.A. Bornes des entités.....	3
2.2.B. Enumération	Erreur ! Signet non défini.
2.2.C. Discontinuité.....	4
2.2.D. Acronymes	5
Guillemets, quotes	5
2.2.E. Zone d'annotation.....	6
2.3. Consignes d'annotation par type d'entité	6
2.3.A. Bornes des mentions d'espèce	6
2.3.B. Organisme nuisible.....	7
2.3.C. Plante hôte	9
2.3.D. Vecteur	12
2.3.E. Maladie.....	14
2.3.F. Voie de dispersion	15
2.3.G. Lieu.....	16
2.3.H. Date.....	19
3. Normalisation	21
3.1. Normalisation des organismes	21
3.2. Normalisation de la Voie de Dispersion.....	24
3.3. Normalisation de Lieu.....	24
4. Annotation des relations	28
4.1. Relation de Coréférence.....	28
4.2. Relations binaires	29
4.2.A. Cause	30
4.2.B. Trouve_sur	30
4.2.C. Vecte_par	31
4.2.D. S_exprime_sur	31
4.2.E. Disperse_par	31

4.2.F. Localise_en	32
4.2.G. Detecte_le	32
4.3. Relations n-aires.....	33
4.4. Modalité	34
4.4.A. Négation.....	34
4.4.B. Hypothèse	34

1. Contenu du document

Ce document détaille les consignes d'annotation manuelle du corpus EPOP (*Epidemiomonitoring Of Plant*). Il est destiné aux experts qui annotent ces documents et à l'évaluation et la mise au point les méthodes d'extraction d'information qui prédisent automatiquement ces annotations. Le document est divisé en trois parties dédiées à l'annotation des entités et à l'annotation des relations entre les entités. Chaque instruction est détaillée et illustrée par des exemples.

2. Annotation des mentions d'entités

2.1. Listes des types d'entité

Les types d'entité sont les suivants.

Organisme_nuisible	Voie_de_dispersion	Lieu
Vecteur	Maladie	Date
Plante		Quantité

Dans ce qui suit, nous utilisons ce code de couleurs dans les exemples pour mettre en évidence les mentions d'entités et faire référence aux entités annotées.

2.2. Consignes générales d'annotation des entités

2.2.A. Bornes des entités

On appelle borne, les limites de l'annotation des mentions d'entités. Les annotations doivent s'étendre à tous les éléments pertinents, mais exclure ceux qui n'apportent pas d'informations pour identifier l'entité. Notamment, les mots qui n'ont pas d'intérêt hors contexte ne sont pas annotés. Les annotations sont restreintes à un groupe nominal. Elles incluent éventuellement une proposition relative si elle contribue à la description de l'entité.

Des consignes détaillées sont données ci-dessous dans les sections dédiées aux différents types. Cette section présente la philosophie générale illustrée par quelques exemples.

Exemples

	Correct	Incorrect
(1) Entités distinctes	for the control of HLB; in the genetic improvement of citrus varieties with resistance to <i>Candidatus Liberibacter asiaticus</i>	for the control of HLB; in the genetic improvement of citrus varieties with resistance to <i>Candidatus Liberibacter asiaticus</i>
(2) Type de la zone géographique	Basque Country	Basque Country

- | | | | |
|-----|------------------------------|---|---|
| (3) | Type de l'entité | de Huanglongbing disease | Huanglongbing disease |
| (4) | description pas localisation | found on the steep Mont Blanc massif | found on the steep Mont Blanc massif |

2.2.B. Énumération

Si les éléments énumérés sont tous pertinents indépendamment les uns des autres, ils sont annotés séparément (exemple (5)).

Si une énumération de termes qualifie la même entité, les termes sont annotés séparément (exemple (6)). Si l'énumération est composée de termes de plus en plus spécifiques ou généraux, seule l'entité la plus spécifique participe aux relations.

Exemples

	Correct	Incorrect
(5) séparés	The bacteria can be found in both Italy and France	The bacteria can be found in both Italy and France
(6) région, pays	During the 2014 - 2015 harvest, field samples of wilting chickpea plants were taken to the laboratory at the CENSA, Mayabeque Province, Cuba .	During the 2014 - 2015 harvest, field samples of wilting chickpea plants were taken to the laboratory at the CENSA, Mayabeque Province, Cuba .

2.2.C. Discontinuité

Les annotations d'entités peuvent être discontinues, en particulier quand un même adjectif ou un même nom qualifie deux entités distinctes. Dans l'exemple (7) deux ensembles distincts de provinces de Toscane sont mentionnés, ceux qui sont au nord-ouest et ceux qui sont au sud. Ils sont annotés comme deux entités distinctes avec les mots *of Tuscany* en commun.

La discontinuité ne sera pas utilisée lorsque l'entité est représentée par l'union des mots ou phrases comme dans l'exemple (8), qui désigne une période entre deux dates précises. Ce moyen ne doit pas être utilisé avec l'objectif de rapprocher des expressions dans le texte de manière arbitraire comme dans l'exemple (9), où l'intention pourrait être d'exclure "fungus called". Ici, il faut annoter seulement "tropical race 1".

Exemples

	Correct	Incorrect
(7) 2 entités distinctes	North-Western provinces and Southern provinces of Tuscany	North-Western provinces and Southern provinces of Tuscany
(8) Durée : une seule entité	between August 15 and September 12	between August 15 and September 12
(9) Discontinuité	a strain of Fusarium fungus called tropical race 1	a strain of Fusarium fungus called tropical race 1

Exceptions

Dans le cas de texte contenant des caractères non signifiants, comme des balises, l'annotation d'entités inclut les caractères des balises en cause, et elle n'est pas discontinuée (exemple (10)).

Exemple

- (10) infected by different subspecies of `Xylella` (`multiplex`

2.2.D. Acronymes

Les acronymes qui suivent une mention d'entité et qui désignent la même entité sont annotés séparément. Les deux entités sont connectées par une relation de coréférence, voir la [section correspondante](#).

Exemples

- (11) `Fusarium oxysporum f. sp. Cubense` fungi (`Foc`)

- (12) `Huanglongbing` (`HLB`)

Guillemets, quotes

Les guillemets ou apostrophes éventuellement présents autour de l'entité ne sont pas inclus dans l'annotation. Inversement, s'ils se trouvent à l'intérieur de l'entité, ils sont inclus selon la règle ci-dessus pour éviter la discontinuité.

	Correct	Incorrect
(13) Guillemets externes	"Bois noir"	"Bois noir"
(14) Quotes externes	'Candidatus Phytoplasma solani'	'Candidatus Phytoplasma solani'
(15) Guillemets internes	"tropical" race 4	"tropical" race 4

2.2.E. Zone d'annotation

L'ensemble du texte du document sera annoté par défaut. Toutefois, si nécessaire, une partie du texte peut être sélectionnée comme "zone_exclue", afin d'exclure une zone du texte pour des raisons telles que : le texte est trop long, répétitif, non pertinent, etc.

Les titres des "références" peuvent être annotés. Les entités à l'intérieur du titre de la référence peuvent être reliées entre elles, mais il ne doit pas y avoir de relation qui lie des entités du texte du document avec des entités des références, ni entre des entités de références distinctes.

2.3. Consignes d'annotation par type d'entité

Les entités de types Organisme nuisible, Plante hôte et Vecteur sont toutes des espèces. Ces types partagent les mêmes règles d'annotation.

2.3.A. Bornes des mentions d'espèce

Lorsqu'ils qualifient une entité qui n'est *pas* un nom scientifique, les noms des types d'entité (insect, pest, bacteria,...) et les adjectifs pertinents sont inclus dans l'annotation pour contribuer à préciser l'entité (exemples (16) à (19)).

Les noms scientifiques de taxon sont suffisants pour désigner les espèces. Les entités sont donc annotées en excluant le nom du type de l'entité (ex. *insect*) et les qualificatifs (exemples (20), (21) et (22)).

Exemples

	Correct	Incorrect
(16) Adjectif non pertinent	Field samples of wilting chickpea plants with chlorotic and necrotic roots were taken to the laboratory at CENSA.	Field samples of wilting chickpea plants with chlorotic and necrotic roots were taken to the laboratory at CENSA.

(17)	Inclure ce qui aide à identifier	PWD occurs when the host plant is infected with pinewood nematodes .	FPWD occurs when the host plant is infected with pinewood nematodes .
(18)	Annoter tout le nom	yellow mottle virus	yellow mottle virus
(19)	Inclure pest	fall armyworm pest	fall armyworm pest
(20)	Nom scientifique	Bactrocera dorsalis insect	Bactrocera dorsalis insect
(21)	Nom scientifique	Xylella fastidiosa bacterium	Xylella fastidiosa bacterium
(22)	Adjectifs non pertinents	Gram-negative bacterium Candidatus Liberibacter species	Gram-negative bacterium Candidatus Liberibacter species

À remarquer, la différence entre adjectifs qui ajoutent de l'information pour spécifier l'entité (e.g. **fall armyworm pest**) et les adjectifs qui décrivent l'entité sans donner plus d'information sur son identité (e.g. (1) **wilting** chickpea).

2.3.B. Organisme nuisible

Définition

Un organisme nuisible est un **organisme vivant** ou un groupe d'organismes, dont **la nuisance est explicite dans le document**. Ce sont des ravageurs ou agents pathogènes, des genres bactéries, virus, insectes, champignons, oomycètes, nématodes, gastéropodes ou arachnides qui affectent les plantes. Plus rarement, l'organisme nuisible peut aussi être une plante parasite.

Ils sont dénotés par leur nom scientifique ou leur nom vernaculaire (nom commun) en entier, ou sous la forme de métaphore, d'acronyme ou d'abréviation.

Les mentions d'organismes **non nuisibles** ne doivent pas être annotées et les genres dont certaines espèces ne sont pas pathogènes ne doivent pas non plus être annotés (ex *Fusarium* sp.).

Lorsque la mention désigne un organisme nuisible par son taxon, seules les mentions de taxon inférieur ou égal au niveau taxonomique du genre sont annotées, c'est-à-dire, les genres, les espèces et les sous-espèces (e.g. souches, cultivars, pathovars). Les taxa de niveau supérieur au genre ne sont pas annotés.

Les entités de type organisme nuisible annotées doivent être normalisées (voir la [section sur la normalisation](#)).

Exemples

- | | | |
|------|-----------------------------|---|
| (23) | nom scientifique | The Spodoptera frugiperda is a species in the order Lepidoptera. |
| (24) | nom vernaculaire | The name " fall armyworm " can refer to several species. |
| (25) | Nom vernaculaire (acronyme) | An urgent response to the rapid spread of FAW is needed. |

Bornes

La mention organisme nuisible inclut tous les qualificatifs utiles à sa détermination et exclut les autres. Notamment, le nom scientifique inclut les mots désignant la **sous-espèce**, l'**autorité** et la **date** le cas échéant.

Exemples

	Correct	Incorrect
(26) Sous-espèce	X. fastidiosa subspecies pauca ST53	X. fastidiosa subspecies pauca ST53
(27) L'autorité	Spodoptera frugiperda (JE Smith)	Spodoptera frugiperda (JE Smith)
(28) Date	Spodoptera frugiperda, Smith, 1797	Spodoptera frugiperda, Smith, 1797

Les organismes nuisibles désignés par leur stade de développement ne sont pas annotés.

Exemples

	Correct	Incorrect
(29) trop général pour être une entité	The larvae develop in 10 months	The larvae develop in 10 months

(30) informations non nécessaires à la détermination de l'entité adult **Popillia japonica** adult **Popillia japonica**

Exceptions

Certaines mentions d'organisme nuisible sont très générales. Certaines de ces mentions sont tellement fréquentes qu'elles n'apportent aucune information.

Les mentions suivantes sont dans ce cas. Elles ne sont pas annotées si elles sont employées seules, sans autre précision, c'est-à-dire sans complément ni adjectif.

- | | | |
|-------------|-----------------|------------|
| - pest | - microbial | - snail |
| - specimens | - microorganism | - nematode |
| - biotope | - population | - pathogen |
| - carrier | - subject | - parasite |
| - host | - vector | |
| - microbe | - virus | |

2.3.C. Plante hôte

Définition

Une plante hôte est une plante ou un groupe de plantes, dont le rôle d'hôte est explicite dans le document. Le rôle d'hôte peut être exprimé par **une relation avec un organisme nuisible** ou **par des termes caractérisant son rôle**.

Le rôle d'hôte doit être compris ici au sens large **d'organisme hôte qui est habitat d'un organisme nuisible**, il inclut les plantes de repos, les plantes réservoirs, et les plantes dans lesquelles l'organisme nuisible se reproduit.

Les entités annotées doivent être normalisées (voir [section sur la normalisation](#)).

Exemple

(31) [...] shared by two hosts, **Crepis foetida** and **Daucus carota**.

Les espèces et genres peuvent être dénotés par leur **nom scientifique**, leur **nom vernaculaire** mais aussi par le mot qui désigne l'ensemble d'une même espèce cultivée (exemple (33)).

Exemples

(32) **Pitch pine** (**Pinus rigida** Mill.)

(33) around the vineyards in the Tuscan region.

Le type plant hôte n'inclut pas les plantes désignées par leur usage ou par leur filière, par exemple *ornamental, cultivated, garden plant, etc.*
Le type plant hôte n'inclut pas les parties de plantes, par exemple les fruits, ou les tiges.

Contre-exemples

- (34) common plant species such as olive, blackberry, plum, avocado, citrus, cypress and many other woody species". "as major pest of many fruit crops", "the oriental fruit fly is known to target over 230 different types of products, including pome, stone fruits, citrus, dates, avocados, peppers and tomatoes
- (35) a fungus called TR1 began destroying banana crops

Lorsque la mention désigne une plante par son taxon, seules les mentions inférieures ou égales au genre sont annotées, c'est-à-dire, les genres, les espèces et les sous-espèces (e.g. cultivar). Cependant, les génotypes entre parenthèses ne seront pas annotés (ex 337). Par contre, les organismes génétiquement modifiés (OGM) seront annotés comme plantes, en incluant les mentions qui désignent leur nom (ex 38).

Exemples

	Correct	Incorrect
(36)	Cultivar Asplenium sp. CV 961 NCBI ID: 218655 (Asplenium sp. CV 961)	NCBI ID: 32071 (Asplenium)
(37)	Cultivar et genotype banana cultivar Maça (Silk, AAB)	banana cultivar Maça (Silk, AAB)
(38)	Plante OGM Bt maize has been granted environmental release in Kenya. TELA maize has the potential to transform Nigeria	Bt maize has been granted environmental release in Kenya. TELA maize has the potential to transform Nigeria

Bornes

La mention de type plante hôte inclut tous les qualificatifs utiles à sa détermination et exclut les autres. Notamment, le nom scientifique inclut les mots désignant la sous-espèce, l'autorité et la date le cas échéant. Ainsi que des qualificatifs indiquant qu'il s'agit d'une espèce cultivée (crops, farms, plantation, et autres) ou d'une variété (variety).

Exemples

-
- (39) *Plectranthus hadiensis* var. *tomentosus* plants with stem and root rot symptoms
- (40) A sentinel plantation of *Prunus domestica* cv. Opal, *Quercus petraea* and *Salix alba* were established .
- (41) *Pinus koraiensis* Siebold & Zucc
- (42) it is still largely unknown which African banana varieties are susceptible
- (43) in addition, certain varieties of plane trees have been selected
- (44) banana variety called Cavendish
-

Exceptions

Les noms de plante qui sont inclus dans des noms de maladie ne sont pas annotés. Par exemple, dans "Tomato brown rot virus", "Tomato" ne doit pas être annoté comme une plante.

Certaines mentions désignent des entités très générales et n'apportent aucune information. Les mentions suivantes ne sont pas annotées si elles sont employées **seules, sans autre précision**, c'est à dire sans complément ni adjectif :

- host
- crop

Dans certains cas, le nom vernaculaire de la plante est utilisé pour caractériser la filière, l'industrie, etc. Dans ce cas, la mention de la plante ne sera pas annotée. Cependant, les mentions de plantes associées à la ferme, à la plantation ou aux agriculteurs sont annotées.

Exemples

Correct

Incorrect

(45)	Fruit	The global supply of bananas has previously been threatened	The global supply of bananas has previously been threatened
(46)	Fruit	Bananas are Britain's favourite fruit	Bananas are Britain's favourite fruit
(47)	Fruit	They often grow the Cavendish banana that is sold in supermarkets	They often grow the Cavendish banana that is sold in supermarkets
(48)	Travailleurs	The Fusarium fungus is well-known within the world of banana growers	The Fusarium fungus is well-known within the world of banana growers
(49)	Travailleurs	where the dominant TR4 strain is affecting banana farmers	where the dominant TR4 strain is affecting banana farmers
(50)	Plantations	save banana plantations with a well-known fungicide	save banana plantations with a well-known fungicide
(51)	Fermes	This pathogen affects tomato farms	This pathogen affects tomato farms

2.3.D. Vecteur

Définition

Un vecteur est un organisme ou un groupe d'organismes, dont le rôle de vecteur ou de transporteur d'organisme nuisible ou de maladie à une plante est **explicite** dans le document. Ce rôle peut être représenté **par une relation** avec un organisme nuisible ou une plante, ou une maladie **ou par des termes** caractérisant son rôle, par exemple le mot *vector*.

Il peut s'agir d'insectes, d'araignées, de nématodes de champignons et potentiellement d'autres organismes.

Le rôle de vecteur doit être compris ici au sens large, il inclut les actions de transport externe (ex. poil du thorax) ou interne (salive, pièces buccales). La capacité de transmission d'un organisme nuisible à une plante donnée n'est pas nécessairement mentionnée.

Lorsque la mention désigne le vecteur par son taxon, seules les mentions inférieures ou égales au genre sont annotées, c'est-à-dire, les genres, les espèces et les sous-espèces.

Les entités de vecteur annotées doivent être normalisées (voir [section sur la normalisation](#)).

Exemples

- (52) **ACP** vectors the lethal bacterium (**Candidatus Liberibacter asiaticus**) that causes **Huanglongbing**

- (53) with respect to the Citrus Leafhopper plague (*Diaphorina citri*. kuwayama), for which prevention
- (54) The African citrus psyllid, *Trioza erytrae* (Hemiptera : Triozidae) is a vector of citrus greening disease (Huanglongbing - HLB) caused by the bacterium *Candidatus liberibacter*

Les espèces et genres de vecteurs sont dénotés par leur nom scientifique, leur nom vernaculaire, des acronymes ou des expressions métaphoriques.

Exemples

- (55) Asian citrus psyllid
- (56) *T. erytrae*, *Diaphorina citri*

Bornes

La mention de Vecteur inclut tous les qualificatifs utiles à sa détermination et exclut les autres. Notamment, le nom scientifique inclut les mots désignant la sous-espèce, l'autorité et la date le cas échéant. Dans l'exemple (57) les mentions (*Coleoptera* : *Cerambycidae*) entre parenthèses après (*Gebler*) sont les rangs taxonomiques supérieurs.

Exemples

- (57) *Monochamus alternatus* Hope and *M. saltuarius* (Gebler) (*Coleoptera* : *Cerambycidae*) transmit *Bursaphelenchus xylophilus* causing pine wilt disease

Exceptions

Certaines mentions désignent des entités très génériques. Ces mentions sont tellement fréquentes que généralement elles n'apportent aucune information. Les mentions suivantes ne sont pas annotées si elles sont employées **seules, sans autre précision**, c'est-à-dire sans complément ni adjectif :

- vector
- insect
- insect vector

2.3.E. Maladie

Définition

Une maladie peut être la conséquence de la présence d'un organisme nuisible sur ou dans une plante hôte. Elle est désignée par un terme reconnu. Le type d'entité Maladie exclut les symptômes ou signes seuls.

Exemples

- | | |
|------|---|
| (58) | Huanglongbing, citrus greening disease, pine wilt disease, beech bark disease, Olive Quick Decline Syndrome (OQDS), almond leaf scorch disease (ALSD) |
| (59) | Common Florida Citrus Disease |

Les termes désignant une plante ou une partie de plante, suivis du mot "disease" ne sont pas annotés s'ils ne représentent pas le nom habituel donné à la maladie, par exemple, *apple tree disease* ou *citrus disease* ne doivent pas être annotés.

Dans certains cas, ces expressions désignent néanmoins des maladies précises, bien identifiées et elles sont à annoter. Par exemple,

- banana disease désigne fréquemment Panama disease

Bornes

La mention de Maladie est restreinte aux seuls mots qui la désigne. Les qualificatifs qui représentent la gravité ou l'étendue de la maladie ne sont pas inclus dans l'annotation. Le nom de la plante est inclus dans l'annotation de la maladie s'il fait partie de son nom.

Exemples

	Correct	Incorrect
(60) gravité	deadly pine wilt disease	deadly pine wilt disease
(61) étendue	widespread pine wilt disease	widespread pine wilt disease
(62) nom de la plante fait partie du nom de la maladie	pine wilt disease	pine wilt disease

Exceptions

Les noms de maladie qui sont inclus dans le nom de l'organisme ne sont pas annotés. Par exemple, dans la phrase "*tomato brown rot virus has been found in a tomato farm*", "*tomato brown rot virus*" est annoté comme un organisme nuisible. La maladie "tomato brown rot" ne doit pas être annotée.

Les noms de groupes ou de familles de maladies ne sont pas annotés, s'ils font référence à plusieurs maladies partageant un certain symptôme ou affectant une certaine espèce de plante. Par exemple : "yellowing diseases", "vine disease", "grapevine yellows", ou "citrus disease".

Certaines mentions désignent une entité très générique. Ces mentions sont tellement fréquentes qu'elles n'apportent aucune information. Les mentions suivantes ne sont pas annotées si elles sont employées **seules, sans autre précision**, c'est à dire sans complément ni adjectif :

- disease
- illness

2.3.F. Voie de dispersion

Définition

Une voie de dispersion est une voie **abiotique** ou **biotique**. Son rôle est explicite dans le document. Ce rôle peut être représenté **par une relation** avec l'organisme nuisible ou la maladie transmise ou la plante hôte, ou **par des termes** caractérisant son rôle, par exemple le mot *pathway*. Ce sont généralement des parties de plantes (ex. graines, fruits) transformées ou non (ex. palettes de transport en bois), ou plus largement des voies abiotiques, aériennes, maritimes, aquatiques, routières, etc.

La voie de dispersion qui permet à un organisme nuisible d'affecter une plante n'est pas nécessairement mentionnée dans le document.

Les vecteurs sont des organismes vivants et des voies de dispersion. Ils sont annotés comme vecteurs, et non pas comme voies de dispersion.

Les pratiques agricoles ne sont pas annotées, cependant les outils ou instruments agricoles sont annotés si la transmission grâce à eux est explicite.

Les entités de type Voie de Dispersion sont normalisées (voir la [section](#) Normalisation de la Voie de Dispersion).

Exemples

-
- (63) Imports are vital to Florida's economy and can include hitchhikers within the **cargo** or on/in the vessel (e.g. **woodboring beetles** in **wooden pallets**, organisms in **ship ballast water**).
- (64) It is mainly transmitted mechanically by **contact between plants**, with cultivation care (tying, weeding, picking, etc.).
- (65) It is transmitted by **hands** that have been in contact with an infected plant, on **contaminated clothing**, on **tools** used on infected plants
-

Bornes

La mention de voie de dispersion inclut tous les qualificatifs utiles à sa détermination et exclut les autres.

Exemples

-
- (66) The most common pathway for these pests to enter the State is by "hitchhiking" in **fruits** and **vegetables** brought back inadvertently by **travelers** as they return.
- (67) The main route of entry would be the importation of **fruits** from affected areas as well as **contaminated plant material** (**plants** or **twigs** for grafting not subject to quarantine)
-

Exceptions

Certaines mentions désignent une entité très générique ou abstraite. Ces mentions sont tellement fréquentes que, la plupart des fois, elles n'apportent aucune information. Les mentions suivantes ne sont pas annotées si elles sont employées **seules, sans autre précision**, c'est à dire sans complément ni adjectif qui caractérise la voie de dispersion :

- transport
- pathway
- trade contamination patterns

Normalisation

Les entités de type voie de dispersion sont normalisées (voir [section sur la normalisation](#)).

2.3.G. Lieu

Définition

Les entités de type lieu désignent des lieux géographiques, et doivent être annotées seulement s'ils référencent des éléments à caractère géographique.

Les lieux peuvent être désignés par des adjectifs, par exemple *Apulian région* ou *Brazilian agriculture* (exemple (68)). Ils sont annotés comme des lieux quand ils font partie d'une autre entité à la condition que le lieu mentionné soit bien le lieu physique et non pas une entité abstraite.

Les adjectifs qui caractérisent une culture ou une plante, indiquant l'emplacement géographique de ces cultures, sont annotés comme lieu (exemples (69) et (70)). Les nationalités de personnes (exemples (71), (72))) ne sont pas annotées.

Par contre en cas de métonymie l'adjectif serait annoté comme dans l'exemple (74) où "growers" réfère vraiment à des plantes cultivées par les agriculteurs.

Les entités annotées doivent être normalisées (voir [section sur la normalisation](#)).

Exemples

	Correct	Incorrect
(68)	Australian vineyards	Australian vineyards
(69)	Brazilian agriculture	Brazilian agriculture
(70)	Mediterranean vegetation	Mediterranean vegetation
(71)	The Spanish citrus employers	The Spanish citrus employers
(72)	Belgian et Dutch farmers	Belgian et Dutch farmers
(73)	Panama disease	Panamadisease
(74)	"A total of 57 Dutch growers have been infected since the first outbreak"	"A total of 57 Dutch growers have been infected since the first outbreak"

Bornes

Les annotations incluent toutes les précisions utiles à l'identification du lieu. Voir les exemples (75) à (82).

Exemples

	Correct	Incorrect
(75)	Entité discontinue has colonized African and Asian countries (African countries et Asian countries)	has colonized African and Asian countries
(76)	précision utile 38 infected plants, 16 are in the countryside of Polignano	38 infected plants, 16 are in the countryside of Polignano
(77)	précision utile advance of Xylella in the province of Bari	advance of Xylella in the province of Bari
(78)	précision utile was also detected near the French border, in Basel (Switzerland)	was also detected near the French border, in Basel (Switzerland)

(79)	précision utile	is now advancing –as it already did further north in Portugal- inland, towards Huelva	is now advancing –as it already did further north in Portugal- inland, towards Huelva
(80)	précision utile	this nematode is widespread in the Far East and Siberian regions of Europe and the Russian Federation.	this nematode is widespread in the Far East and Siberian regions of Europe and the Russian Federation.
(81)	ensemble	The bacteria was found in an ancient administrative region that now encompasses territories across Italy and France	The bacteria was found in an ancient administrative region that now encompasses territories across Italy and France
(82)	précision utile	city of Zürich	city of Zürich

Exceptions

Les expressions portant sur des entités de type institutionnel, politique ou économique ne sont pas annotées (voir exemples (83) à (85))

Certains noms d'organisations, de journaux ou d'instituts comprennent des noms de lieux géographiques (exemples (86) à (89)), ainsi que certaines maladies ou nom d'espèces (scientifique ou vernaculaire), (exemples (90) à (92)). Ces noms de lieux géographiques ne doivent pas être annotés.

Exemple

	Correct	Incorrect
(83)	The head of the Balearic Islands Agriculture Service, Andreu Juan, also participated in the reception for the Corsican delegation	The head of the Balearic Islands Agriculture Service, Andreu Juan, also participated in the reception for the Corsican delegation
(84)	China's Agriculture Ministre	China's Agriculture Ministre
(85)	commitments set forth in the European strategy (...) do not overlap with European production	commitments set forth in the European strategy (...) do not overlap with European production
(86)	A team of researchers from Wageningen University	A team of researchers from Wageningen University

Commenté [1]: à vérifier

(87)	organized by the Inter-American Institute for Cooperation on Agriculture (IICA)	organized by the Inter-American Institute for Cooperation on Agriculture (IICA)
(88)	The Florida Department of Agriculture and Consumer Services (FDACS)	The Florida Department of Agriculture and Consumer Services (FDACS)
(89)	the Union believes that ratifying or signing agreements by the EU, such as with Mercosur or South Africa	the Union believes that ratifying or signing agreements by the EU, such as with Mercosur or South Africa
(90)	Panama disease	Panama disease
(91)	efficiency of the control of the African Psila	efficiency of the control of the African Psila
(92)	also known as the giant African snail	also known as the giant African snail

Exceptions

Les mentions suivantes sont trop générales pour être utiles et ne sont pas annotées :

- worldwide
- all southern countries

Les adresses des auteurs d'un article scientifique ne sont pas annotées comme lieux. Elles sont en général situées juste après le titre.

Exemples

	Correct	Incorrect
(93)	The vector is found worldwide.	The vector is found worldwide.
(94)	in many countries, such as France and Spain.	in many countries, such as France and Spain.

2.3.H. Date

Définition

Les mentions de type date sont des expressions temporelles qui peuvent être une date entière ou partielle (année-mois-jour ou juste année) ou une période / saison. Elles sont absolues ou relatives.

Exemples

	Correct	Incorrect
(95) date d'observation	In 1948, the pine wilt disease was found on 3 pine trees in the region.	Spodoptera frugiperda, Smith, 1797
(96) Période relative	This last decade, the disease spread	

Bornes

L'annotation doit inclure la date la plus précise (heure, jour, année, mois, saison, décennie, etc), ainsi que les prépositions qui sont rattachés aux dates. Par exemple, "from", "over", "between", "during", "more than" "since", "begins in", "within", "during", "in", "at", "for", "by", "of", "on", etc. Les dates relatives ou dépendantes sont annotées aussi (ex 101).

Exemples

	Correct	Incorrect
(97) Date d'observation	In March 1948, the pine wilt disease was found on 3 pine trees in the region.	In March 1948, the pine wilt disease was found on 3 pine trees in the region.
(98) Date relative	In March 1948, the disease was discovered. The following year its presence was recorded	In March 1948, the disease was discovered. The following year its presence was recorded
(99) Période vague	over several years	over several years
(100) Décennie	During the 1940s and 1950s (une annotation en continue et une annotation en discontinue: During the 1940's / During the 1950's)	During the 1940s and 1950s
(101) Saison	During the spring	During the spring

(102)	Jour	The National Bank for Economic and Social Development (BNDES) launched this Tuesday (30) the second public notice	The National Bank for Economic and Social Development (BNDES) launched this Tuesday (30) the second public notice
(103)	Depuis une date	available from the beginning of 2024	available from the beginning of 2024

Les intervalles temporels sont annotés comme une seule entité.

Exemples

	Correct	Incorrect
(104)	Période	Période
	In the period January 2017 - March 2020 , 60480 news items on 494 pests were selected.	In the period January 2017 - March 2020 , 60480 news items on 494 pests were selected.
(105)	Période	Période
	between August 15 and September 12	between August 15 and September 12

Exceptions

Les dates mentionnées dans les citations bibliographiques (qu'elles soient dans le corps du texte ou dans la section référence) et les dates dans les noms d'espèces ne sont pas annotées.

3. Normalisation

3.1. Normalisation des organismes

Les entités de type organisme (organismes nuisibles, vecteurs et plantes) sont associées à leur taxon le plus précis possible, qui est désigné ("normalisé") par son identifiant taxonomique.

Le référentiel choisi est la taxonomie du NCBI dont le navigateur est à : <https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi>

Dans le cas où l'organisme n'est pas dans la taxinomie du NCBI, il devra être annoté par l'identifiant de *Encyclopedia of Life* (EOL) : <https://eol.org/>

Dans le cas où l'organisme n'est dans aucun des deux référentiels, il est normalisé par la classe de la version d'octobre 2021 de l'ontologie OntoBiotope, désignée par son identifiant.

<https://agroportal.lirmm.fr/ontologies/ONTOBIOTOPE>

Les taxons désignés par une abréviation, un acronyme ou une ellipse sont annotés par l'identifiant du taxon référencé.

La normalisation des taxa prend en compte le contexte dans le document, ce qui nécessite une interprétation précise de l'expert. Par exemple, une entité ON "Xylella" sans contexte pourrait être normalisée par le genre *Xylella*; néanmoins, elle peut représenter la coréférence d'une mention plus précise, ou *Xylella* peut être également le nom vernaculaire, le nom commun, de la bactérie, dans ce cas *Xylella* est normalisée sous la dénomination *Xylella fastidiosa*, ou une sous-espèce plus précise, selon le sens du contexte.

Dans l'exemple (106) la mention *Xylella* désigne *Xylella fastidiosa* et doit donc être annoté par l'identifiant NCBI de l'espèce *Xylella fastidiosa*.

Exemple

(106) In the event of an outbreak of *Xylella fastidiosa*, all plants sensitive to *Xylella* must be destroyed

Quand un document mentionne une espèce par son nom vernaculaire, et qu'il pourrait correspondre à plusieurs espèces d'un même genre, la normalisation sera faite par le genre, sauf si l'espèce est clairement mentionnée dans le document ou si l'expert annotateur le considère ainsi.

Par exemple, l'entité *citrus trees* sera normalisée avec l'identifiant du genre *Citrus sp.* car ce genre inclut toutes les espèces de citrus cultivés et non cultivés aussi (oranges, citrons, clémentines, etc) (exemple (107)).

Alors que la normalisation serait faite au niveau espèce quand la différenciation pourrait se faire au niveau de sous-espèces. Par exemple l'entité *vineyards*, serait normalisé avec l'identifiant de *Vitis vinifera*, car d'une part l'entité nous indique qu'il s'agit d'un culture et d'autre part les vignes cultivés se différencient au niveau de cultivar, par conséquent *Vitis vinifera* les inclut toutes (exemple (108)).

Cependant, si l'entité ou contexte ne suffisent pas pour savoir s'il s'agit d'une plante cultivée ou sauvage, la normalisation serait faite avec le taxon qui inclut à la fois les plantes cultivées et non cultivées. Par exemple, *olive trees* sans précision sera normalisé avec l'identifiant de *Olea europaea* (exemple (109)), tandis que l'entité *olive groves* (cultivé) serait normalisé avec l'identifiant de la sous espèce cultivée, *Olea europaea subsp. Europaea* (exemple (110)).

Exemples

(107) *Citrus trees*  *Citrus sp.* ID **2706**

(108) *Vineyards*  *Vitis vinifera* ID **29760**

-
- (109) [Olive trees](#)  Olea europaea ID [4146](#)
- (110) [Olive groves](#)  Olea europaea subsp. europaea ID [158383](#)
-

En cas d'erreur de l'auteur qui confondrait deux espèces, il faudra normaliser en prenant cette erreur comme correcte (exemple (111)). Dans cet exemple, *Fusarium oxysporum f.sp. cubense tropical race 4* est le nom scientifique complet de l'espèce de cet ON, mais dans cet exemple "*Fusarium oxysporum f.sp. Cubense*" et "race four tropical" sont présentés comme synonymes. Dans ce cas les deux mentions sont normalisées selon l'intention de l'auteur, par l'identifiant de *Fusarium oxysporum f.sp. cubense tropical race 4*.

Exemples

-
- (111) [Fusarium oxysporum F. sp. Cubense](#),  [Fusarium odoratissimum](#) ID 2502994
also known as [race four tropical](#)
-

Normalisation avec 2 ou plus identifiants

La normalisation permet aussi d'inclure 2 identifiants ou plus pour une seule entité. Comme dans l'exemple (112) de l'annotation de l'entité Foc. L'identifiant 61366 *Fusarium oxysporum f. sp. Cubense* du NCBI inclut tous les tropical races, sauf la 4, qu'a été renommé *Fusarium odoratissimum*. La conjonction des deux est donc nécessaire pour normaliser *Foc dans sa globalité*.

Exemple

	Correct	Incorrect
(112) Foc	 ID 61366 (<i>Fusarium oxysporum f. sp. cubense</i>)  ID 2502994 (<i>Fusarium odoratissimum</i> , homotypic synonym: <i>Fusarium oxysporum f. sp. cubense tropical race 4</i>)	 ID 61366 (<i>Fusarium oxysporum f. sp. cubense</i>)

3.2. Normalisation de la Voie de Dispersion

Les entités de type Voie de dispersion sont normalisées par une classe de l'ontologie OntoBiotope (version d'octobre 2021 <https://agroportal.lirmm.fr/ontologies/ONTOBIOTOPE>). Il faudra choisir la catégorie la plus pertinente, en fonction de la nature de la voie de dispersion.

Exemples

(113)	Water	environment water OBT:000047
(114)	Coniferous wood	wood OB :001383
(115)	Wood Packaging Material	wooden pallet OBT:003753
(116)	nonhost species	living organism OBT:000010
(117)	farm equipment	agricultural equipment OBT:000032
(118)	Citrus fruits	citrus fruits OBT:002971
(119)	contaminated tools	agricultural tool OBT:000185

3.3. Normalisation de Lieu

Les lieux sont normalisés par les identifiants de la version 2022 du référentiel GeoNames : <https://www.geonames.org/>

La sélection de l'identifiant de Geonames est conditionnée par le choix judicieux du 'Feature class'. Par défaut, la *Feature class* choisie est la division administrative et la *political entity* pour le pays. Pour annoter les divisions administratives, il faut soigneusement choisir le bon niveau (*first to fourth order*) quand plusieurs entrées sont possibles, en privilégiant les identifiants qui représentent des polygones (des surfaces) et en écartant les options de siège administratif (*seat of*) ou de lieu habité (*populated place*).

Les règles générales et les différents cas d'exception sont décrits ci-dessous, avec des exemples illustratifs :

Règles générales

- Pour une entité de lieu de type **continent** → choisir le *Feature class* “continent”
- Pour une entité de lieu de type **pays** → choisir le *Feature class* “independent political entity” ou “dependent political entity” dans le cas des territoires litigieux (ex. *Puerto Rico*).
- Pour une entité de lieu de type **région administrative** → choisir le *Feature class* “first / second order administrative division” (selon le niveau de division administrative du pays)
- Pour une entité de lieu de type **ville** ou **commune** → choisir le *Feature class* “second / third / fourth-order administrative division” (selon la division administrative du pays)

Lorsqu'un lieu n'est pas défini dans Geonames, il est annoté par l'identifiant de la **région englobante la plus spécifique** (exemples (120) et (121)).

Exemples

(120) **central-southern Puglia**  Puglia Geonames ID 3169778

(121) **western Mediterranean basin**  Mediterranean Basin Geonames ID 12217088

Les lieux ne doivent pas être normalisés par l'identifiant de type *Feature class* “seat of a X-order administrative division” qui représente simplement le lieu du siège administratif (exemple (122)), ni par les identifiants de type *Feature class* “populated place” (exemple (123)).

Exemples

		Correct	Incorrect
(122)	city of Freiburg	 Freiburg im Breisgau Geonames ID 6555728 (fourth-order administrative division)	 Freiburg im Breisgau Geonames ID 2925177 (seat of a second-order administrative division)
(123)	Mogro	 Miengo Geonames ID 6360690 (third-order administrative division)	 Mogro Geonames ID 3116630 (populated place)

Exceptions aux divisions administratives

- Région

Les lieux qui correspondent à une zone géographique particulière sans désignation administrative, peuvent être annotés avec la feature class “*region*” (exemple (124)).

Exemple

Correct		
(124)	Central Italy	Central Italy Geonames ID 12089032 (Feature class: region)

- Zone

Lorsque l’entité désigne des pays unifiés par un accord politique comme “*Union Européenne*”, elle doit être annotée avec la feature class “*zone*” (exemple (128)).

Exemple

	Correct	Incorrect
(125)	European union European union Geonames ID 6695072 (zone)	Europe Geonames ID 6255148 (continent)

- Géographie physique

Si l’entité désigne un élément de géographie physique (ex. montagne, lac, port) qui ne correspond pas à une division administrative, alors une autre valeur de *Feature class* peut être sélectionnée.

Exemple

Correct	
(126)	Rwenzori Mountains Massif du Ruwenzori Geonames ID 206041 (mountains)

- Union d’identifiants

Si l'entité désigne une région composée de plusieurs divisions administratives et qu'elle n'existe pas comme telle dans Geonames, alors l'entité est annotée soit par les tous les identifiants des divisions incluses si elles sont identifiables (exemple (127)), soit par la division administrative ou zone/région englobante la plus spécifique (exemple (128)). La normalisation avec plus d'un identifiant Geonames signifie l'union de ces identifiants.

Exemples

		Correct
(127)	Iberia	<ul style="list-style-type: none"> • Espagne Geonames ID 2510769 • Portugal Geonames ID: 2264397
(128)	Latin America	<ul style="list-style-type: none"> • South America Geonames ID 6255150 • Central America Geonames ID 7729892

- Plusieurs identifiants possibles

Dans le cas où il est possible de normaliser une entité avec plusieurs identifiants de Geonames et que les polygones de ces identifiants coïncident (voir carte dans Geonames), la règle indique que nous devons donner la priorité à la "division administrative" (exemple (129)), à moins que l'entité ne contienne un terme de géographie physique qui caractérise l'entité, comme *island* (exemple (130)).

Exemples

		Correct	Incorrect
(129)	Corsica	<ul style="list-style-type: none"> • Corse Geonames ID 3023519 (first-order administrative division) 	<ul style="list-style-type: none"> • Corse Geonames ID 3023517 (region)
(130)	Corsican island	<ul style="list-style-type: none"> • Corsican island Geonames ID 3023518 (island) 	<ul style="list-style-type: none"> • Corse Geonames ID 3023519 (first-order administrative division)

4. Annotation des relations

Les relations sont annotées dans le texte seulement lorsque ces dernières sont explicitées dans le texte. Si la relation est connue, mais non explicitée, elle n'est pas annotée. Les entités qui sont reliées par une relation sont appelées les arguments de la relation. Les relations sont dites binaires lorsqu'elles relient deux arguments. Elles sont dites ternaires pour trois arguments, quaternaire pour quatre. On appelle plus généralement n-aires les relations qui ont plus de deux arguments.

Distance

Une entité donnée est reliée à une autre entité si elles sont distantes au maximum de deux phrases avant et deux phrases après la phrase de l'entité. Les entités plus lointaines ne sont pas reliées malgré leur intérêt, sauf si elles font partie du titre principal. Les relations n-aires peuvent s'étendre sur plus de deux phrases, de proche en proche.

Type de relation

Il y a deux types de relations, (1) les relations d'équivalence qui relient deux entités qui signifient la même chose. Elles sont appelées relations de coréférence. (2) Les relations thématiques relient des entités en fonction de leur rôle dans l'interaction.

4.1. Relation de Coréférence

La relation de coréférence relie des termes, deux, ou plus, qui font référence à la même entité. Elle représente une équivalence entre plusieurs mentions.

Exemples

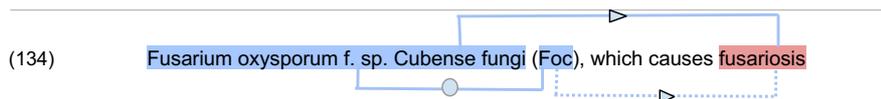
(131)	coréférence entre maladies	Panama disease, also called banana wilt, a devastating disease of bananas
(132)	coréférence entre 3 ONs	Apple proliferation phytoplasma (APP) (Candidatus Phytoplasma mali) is a plant pest
(133)	coréférence entre lieu	the virus could have been spread from the Democratic Republic of Congo (DRC)

Simplification de l'annotation grâce aux coréférences

Lorsque deux entités E1 et E2 sont reliées par une relation de coréférence, chaque entité hérite des relations thématiques de l'autre entité. Il n'est pas nécessaire alors d'annoter pour E1 les relations impliquant déjà E2.

Dans l'exemple ci-dessous, *Fusarium oxysporum f. Sp. Cubense fungi* et *Foc* sont reliés par une relation de coréférence. Il n'est pas nécessaire de créer une relation de cause entre le pathogène *Foc* et la maladie *fusariosis*. La relation entre *Fusarium oxysporum f. Sp. Cubense fungi* et

fusariosis et la relation de coréférence suffisent pour établir que *Foc* et la maladie *fusariosis* sont en relation également.



Lorsque le terme coréférent n'est pas synonyme de l'antécédent, il n'est pas annoté, notamment lorsqu'il est plus général.

Par exemple "bacteria" ou "nematode" ne sont pas annotées comme coréférence dans les exemples (135) et (136).

(135) Apple proliferation phytoplasma, this bacteria

(136) B. xylophilus. This nematode is widespread

La coréférence peut être utilisée pour relier des entités ayant une forme de surface différente mais ayant la même signification (ex 135 à 137), mais aussi pour relier des entités ayant une forme de surface identique avec l'objectif de leur faire hériter les relations d'une entité à une autre à la condition qu'elles aient un rôle équivalent dans l'argumentation (ex. 137).

Exemple

(137) coréférence surface identique Bursaphelenchus xylophilus, is a notorious parasitic nematode of pine trees that causes pine wilt disease (PWD), leading to extensive mortality of different pine trees around the world and considerable economic losses

Dans l'exemple (137), les relations sont les suivantes:

[Bursaphelenchus xylophilus → cause → pine wilt disease] → trouvé sur pine trees

pine trees – coreference – pine trees

pine wilt disease – coreference – PWD

4.2. Relations binaires

La relation binaire relie deux entités. Elle est orientée, elle a un sens. L'entité argument de départ est appelée source et l'entité argument d'arrivée est appelée cible. Les relations ont une étiquette associée qui caractérise le type de la relation par exemple, *cause* entre organisme nuisible et maladie. Il en existe 8 différentes.

Les sections ci-dessous sont organisées par type de relation binaire. Pour chacune le type des entités sources et cibles possibles sont indiquées.

->	ON	Plante	Vecteur	Maladie	Voie de Dispersion	de	Date	Lieu
----	----	--------	---------	---------	--------------------	----	------	------

ON	-	Trouvé_sur	Vecte_par	Cause	Trouvé_sur Dispersé_par	Detecte_le	Localise_en
Plante	-	-	-	-	-	Detecte_le	Localise_en
Vecteur		Trouvé_sur	-	-	Trouvé_sur	Detecte_le	Localise_en
Maladie	-	S- exprime_sur	-	-	Dispersé_par	Detecte_le	Localise_en
Voie	-	-	-	-	-	Detecte_le	Localise_en
Date	-	-	-	-	-	-	-
Lieu	-	-	-	-	-	-	-

Table 1: Relations valides selon le type d'entité source et cible

4.2.A. Cause

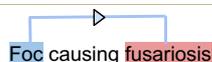
Entité source **ON**

Entité cible **Maladie**

Définition Relation entre un organisme nuisible et la maladie qu'il cause.

Exemples

(138)



4.2.B. Trouve_sur

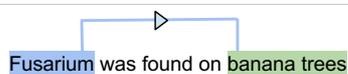
Entité source **ON** ou **Vecteur**

Entité cible **Plante** ou **Voie_de_dispersion**

Définition Relation entre un organisme nuisible ou un vecteur et l'entité sur ou dans laquelle on peut le trouver, soit une plante ou une voie de dispersion.

Exemple

(139)



4.2.C. Vecte_par

Entité source **ON**
Entité cible **Vecteur**
Définition Relation entre un **organisme nuisible** et un **vecteur**

Exemple

(140) **Candidatus Liberibacter asiaticus** is vectored by **Diaphorina citri Kuwayama**

4.2.D. S_exprime_sur

Entité source **Maladie**
Entité cible **Plante**
Définition Relation entre une plante et la maladie qui s'y exprime.

Exemple

(141) they help manage **fusariosis** in **banana trees**.

4.2.E. Disperse_par

Entité source **ON** ou **Maladie**
Entité cible **Voie_de_dispersion**
Définition Relation entre l'organisme nuisible ou la maladie, et sa voie de dispersion. De fait les maladies ne sont pas dispersées en elles-mêmes, mais cet

abus de langage qui identifie la maladie à l'agent est fréquent dans les textes.

Exemple

(142) common pathway for **Oriental fruit fly** to enter the State is by "hitchhiking" in **fruits**



4.2.F. Localise_en

Entité source **ON** ou **Vecteur** ou **Plante** ou **Maladie** ou **Voie_de_dispersion**

Entité cible **Lieu**

Définition Relation entre une entité observée et le lieu de l'observation.
Un organisme pathogène qui provoque une épidémie est localisé dans le lieu de l'épidémie.

Exemple

(143) **Fusariosis** is common in **Brazil**



4.2.G. Detecte_le

Entité source **ON** ou **Vecteur** ou **Plante** ou **Maladie** ou **Voie_de_dispersion**

Entité cible **Date**

Définition Relation entre l'entité observée et la date de l'observation.

Exemple



4.3. Relations n-aires

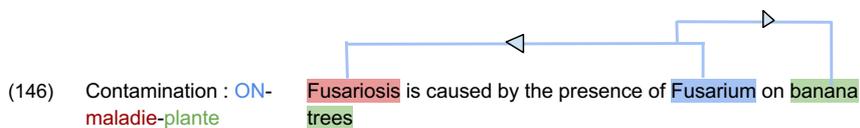
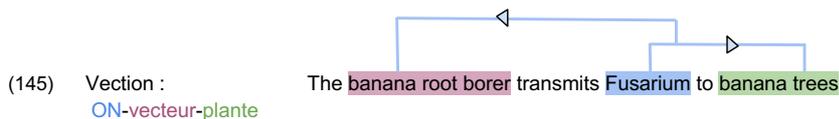
Définition

Les relations n-aires sont des relations qui relient plus de deux arguments. Elles sont aussi appelées "événements". Une première relation binaire relie les deux entités principales. Les autres relations relient la première relation avec les entités supplémentaires. Le sens de la relation est le même que celui des relations binaires

Le choix de la première relation est effectué selon un **ordre de priorité fixe** :

1. Cause [ON -> Maladie]
2. Trouve_sur [ON -> Plante]
3. Vecte_par ou Disperse_par [ON ou Maladie -> Vecteur ou Voie_de_dispersion]
4. Localise_en [? -> Lieu]
5. Detecte_le [? -> Date]

Exemples



4.4. Modalité

La propriété modalité est associée à certaines relations. Elle peut prendre la valeur neutre par défaut, négation ou hypothèse.

4.4.A. Négation

La modalité prend une valeur "négation" lorsque la relation est explicitement niée dans le document. Cette modalité peut être utilisée lorsqu'une relation de résistance est décrite entre la plante hôte et l'organisme nuisible.

Exemples

-
- (147) négation millet, [...], have not displayed potential as alternative hosts of Foc
- (148) négation Cavendish variety with complete resistance to Panama disease will be available in 2024
- (149) negation The positives have been detected in Mallorca, Menorca and Ibiza; Formentera is the only island, at the moment, free of 'Xylella fastidiosa'
- (150) negation in the framework of the 2022 survey program for the harmful organism xylella fastidiosa in its host plants [...] in order to confirm the absence of the harmful organism (detection survey).
-

4.4.B. Hypothèse

La modalité d'une relation prend la valeur "hypothèse" lorsque la relation est explicitement hypothétique dans le document.

Exemples

-
- (151) Hypothèse Crotalaria [...] identified as potential hosts of Foc
- (152) Hypothèse Gathering direct evidence about whether the bacteria infect pepper tissues may prove difficult, as Hansen's team only had a single sample, and L. capsica cannot be grown in a laboratory.
-

