

# Corpus EN-ISTEX

## Guide d'annotation des entités nommées

(INIST-CNRS / LIFAT)

1



## Table des matières

---

<b>1</b>	<b>CONSTRUCTION D'UN CORPUS IStEX</b>	<b>3</b>
1.1	ORIGINE DE LA DEMANDE	3
1.2	CHOIX DES DOCUMENTS	3
1.3	CAMPAGNE D'ANNOTATION	4
1.4	ACCORD INTERANNOTATEUR	7
1.5	CORPUS ANNOTE	9
1.6	PUBLICATION DU CORPUS	9
<b>2</b>	<b>REGLES D'ANNOTATION AVEC LES ENTITES NOMMEES D'ISTEX</b>	<b>10</b>
2.1	PERSONNES : <PERSNAME>	10
2.2	LIEUX LIES A L'INTERVENTION HUMAINE : <PLACENAME>	11
2.3	LIEUX LIES A LA GEOGRAPHIE PHYSIQUE : <GEOGNAME>	13
2.4	DATES ET PERIODES : <DATE>	15
2.5	ORGANISATIONS : <ORGNAME>	18
2.6	ORGANISMES FINANCEURS : <ORGNAME TYPE="FUNDER">	20
2.7	ORGANISMES HEBERGEURS DE RESSOURCES : <ORGNAME TYPE="PROVIDER">	22
2.8	URL : <REF TYPE ="URL">	23
<b>3</b>	<b>ANNEXE: LEXIQUE DES PÉRIODES</b>	<b>24</b>
3.1	PERIODES GEOLOGIQUES	24
3.2	PERIODES ARCHEOLOGIQUES	25
3.3	PERIODES HISTORIQUES, PROTOHISTORIQUES ET PREHISTORIQUES	25
3.4	PERIODES HISTORIQUES CHINE, JAPON, COREE	25

# 1 Construction d'un corpus ISTEEX

---

## 1.1 Origine de la demande

En décembre 2019, le Laboratoire d'Informatique Fondamentale et Appliquée de Tours (LIFAT) demande au Service *Analyser et Fouiller l'Information Scientifique* de l'Inist-CNRS d'étudier la possibilité de constituer un corpus gold représentatif du fonds ISTEEX, afin de disposer d'un corpus qui permettrait d'évaluer les performances d'outils de reconnaissance d'entités nommées.

ISTEX étant un réservoir renfermant plus de vingt-trois millions de documents, il s'avérait impossible d'annoter manuellement un corpus représentatif de la totalité d'ISTEX. Il a été alors décidé de prendre en considération les deux plus importants éditeurs du fonds :

- Elsevier : six millions de documents
- Wiley : cinq millions et demi de documents

Cent documents de langue anglaise ont été choisis pour chacun des deux éditeurs.

La constitution du corpus, son annotation et sa publication sur <https://data.istex.fr/>, ainsi que la rédaction de ce guide, se sont déroulées au long de l'année 2020.

## 1.2 Choix des documents

### 1er critère : les catégories scientifiques Science-Metrix

Science-Metrix attribue aux revues de sa couverture une catégorie scientifique à trois niveaux. Exemples :

- 1 - applied sciences ; 2 - engineering ; 3 - environmental engineering
- 1 - health sciences ; 2 - public health & health services ; 3 - public health
- 1 - natural sciences ; 2 - earth & environmental sciences ; 3 - environmental sciences

Dans ISTEEX, l'outil Multicat attribue par appariement aux documents des revues appartenant à cette couverture la catégorie correspondante.

Voir <http://data.istex.fr/> : Enrichissement Multicat/Science-Metrix :

<https://enrichment-process.data.istex.fr/ark:/67375/R0H-Q29BJZWS-N>

Le premier niveau de la classification Science-Metrix est retenu pour constituer le corpus. La proportion des catégories dans les corpus-éditeurs est conservée pour constituer le corpus à annoter.

Catégories Science-Metrix	Nbre de docs Wiley	%	Corpus gold Nbre de docs	Nbre de docs Elsevier	%	Corpus gold Nbre de docs
Applied sciences	631290	16,1	16	874572	16,89	17
Arts & Humanities	141103	3,6	4	20279	0,39	1
Economic & Social Sciences	347294	8,9	9	173273	3,35	3
General	74735	1,9	2	5657	0,11	1
Health Sciences	1470224	37,6	37	2611993	50,45	50
Natural Sciences	1248538	31,9	32	1491368	28,81	28
	3913184	100	100	5177142	100	100

Proportions des catégories scientifiques Science-Metrix dans les corpus-éditeurs Elsevier et Wiley et dans les 200 documents retenus pour le corpus à annoter.

## 2ème critère : les types d'entités nommées

Pour être retenu, un document doit contenir un minimum de dix entités nommées réparties dans plusieurs types d'entités nommées.

Les types d'entités nommées sont :

- Les noms de personnes <persName>
- Les lieux administratifs <placeName>
- Les lieux géographiques <geogName>
- Les dates et les périodes <date>
- Les organisations <orgName>
- Les organismes financeurs <orgName type="funder">
- Les organismes hébergeurs de ressources <orgName type="provider">
- Les url <ref type="url">

### 1.3 Campagne d'annotation

Un sous-corpus de vingt documents (10 %) est extrait du corpus pour faire l'objet d'une campagne d'annotation. L'annotation de ces vingt documents par un groupe d'annotateurs qui permettra le calcul d'un accord interannotateur, critère de fiabilité de l'annotation du corpus.

#### Recrutement des annotateurs

Le réseau Ingénierie Documentaire de l'Inist-CNRS, dont le rôle est de favoriser le partage des compétences au sein de l'Institut, sert de tremplin pour diffuser

l'information dans les différents services de l'Institut : on recherche des volontaires pour annoter manuellement un corpus.

Huit personnes travaillant dans six services de l'Inist-CNRS vont constituer l'équipe d'annotateurs.

## Formation à l'usage de l'outil Oxygen

5

Le format des documents annotés est un fichier TEI (TEI = ensemble de balises et d'attributs XML) obtenu par transformation du XML fourni par l'éditeur via une feuille de style. L'annotation s'opère uniquement dans le corps du texte (<body>) et dans les paragraphes (<p>). On ne pose pas d'annotation dans les titres <head>, les tableaux et encadrés <figure> qui se trouvent dans le corps du texte.

L'annotation est facilitée par la structuration des documents. Cependant le fonds ISTEEX contenant des documents non structurés, il paraissait important d'en inclure quelques-uns dans le corpus.

Le format XML fourni par l'éditeur est structuré ou non structuré. Il en est de même pour le fichier TEI qui en découle. La proportion de documents structurés et non structurés est indiquée dans [1.5 Corpus annoté](#).

```

<?xml version="1.0" encoding="UTF-8" ?>
<text>
  <body>
    <div>
      <p>Unemployment in Spain: Its Causes and the Policy Response Lluís Fina Absrruct. O
      loss of jobs in Spain is basically the reverse side of the specific characteristics
      the Franco years. The first section is dedicated t o a general description of this
      which began in the mid-Seventies, identifying the main sectors of the economy and e
      evolution of employment in each of them. The two following sections discuss the fac
      the permanent or 'structural' factors which may have had a negative effect on emplo
      be considered as given by the new situation and whether they can or should be alter
      temporary factors, which have also had a negative effect on employment. The fifth s
      Although these factors have had a relatively minor influence o n unemployment, they
      complete picture of the situation and particularly of the prospects of the Spanish
      how Spain's economic policy is responding t o the problems examined in the earlier
      economic policy can be examined in detail, and so the analysis is limited to a ment
      policy and then proceeds to examine the principal labour market policies, and those
      problems identified earlier in the paper. Finally, a concluding section reviews and
      policy implications. Introduction During 1985 the number of unemployed in Spain ave
      population and is the highest unemployment rate registered in any OECD country. Fif
      Barcelona. Paper presented at the International Conference o n the Employment Probl
      the Associazione Italiana degli Economizti del Lavoro, Rome, 17-18 October, 1986. R
      Editorial Board on June 30, 1987. Labour 1 (2): 29-69 (1987) 29 Lluís Fina 30 ago t
      official sources, the Spanish unemployment rate was then about 1.5%, a figure compa
      countries. In other words, unemployment in Spain is now not only higher than in any
      faster. The Spanish unemployment picture differs from that of most other European c
      is basically due to a loss of jobs rather than to stagnation or insufficient growth
      increase in the active population. Between 1977 and 1985, a period for which a fair
      number of unemployed increased by 2,278,600 and 1,884,000 jobs were lost, i.e. the 1
      84.5% of the increase in unemployment. In the ten EEC countries, these proportions
      period. &apos; Figure 1. Total Labour Force and Total Employment. Spain, 19771985
      i..T .&apos; .?:0. 11000 1981 10000 1977 1978 1979 1980 + LabourForce *0- Employmen
      1 1983 I.?o ~ 1984 1985 Unemployment in Spain: Its Causes and the Policy Response
      modest growth even though the Spanish population is among the fastest growing in Eu
      'potentially active population', i.e. the total population 16 years of age and olde
      whom have entered the labour market. The rate of participation in the Spanish labou
    
```

Document non structuré : un seul paragraphe englobe résumé, titres de sections, texte, pieds de pages, tableaux, etc., sans séparation. Note : les titres de sections sont indiqués en rouge pour permettre leur lecture.

```

▼<text>
  ▼<body>
    ▼<div xml:id="ijpp12032-sec-0005" type="section">
      <head>Introduction</head>
      ▼<p>
        Studies have shown that medication errors, particularly those
        are often a cause for potentially avoidable morbidity and mo
        <ref type="bibr" target="#ijpp12032-bib-0001 #ijpp12032-bib-
        and secondary care.
        <ref type="bibr" target="#ijpp12032-bib-0003"/>
        With the majority of prescribing taking place in primary car
        and address these types of medication error, although the ev
        <ref type="bibr" target="#ijpp12032-bib-0004 #ijpp12032-bib-
      </p>
      ▼<p>
        However, more recently the results of the PINCER Trial, a la
        pharmacist-led information technology-based intervention res
        medication errors within the primary care setting, compared
        <ref type="bibr" target="#ijpp12032-bib-0007"/>
        This was a parallel-group, pragmatic, cluster trial in which
        generated feedback ('simple feedback') in which practices we
        (2) the pharmacist-led intervention comprising computer-gene
        Quest Browser software was used to search general practice c
        specific pharmaceutical indicators (outcomes) (Table
        <ref type="bibr" target="#ijpp12032-tbl-0001"/>
        ). These pharmaceutical indicators (outcomes) were chosen be
        errors resulting in serious morbidity,
    
```

Document structuré : Identification aisée du texte à annoter (balises <p>).

6

Un temps de formation est consacré à la maîtrise du logiciel Oxygen qui permet l'annotation des documents, la vérification de la validité des balises et l'extraction des entités nommées via une feuille de style.

Le document *A comparison of temperate and boreal peats from Ontario, Canada: possible modern analogues for Permian coals* sert à l'entraînement des annotateurs.

Ce document comprend deux cent quatre-vingt-sept entités nommées réparties dans cinq types d'entités nommées : <persName>, <placeName>, <geogName>, <date> et <orgName type="funder">.

Une première annotation de ce document est faite individuellement, puis plusieurs réunions entre annotateurs permettent de discuter sur les frontières et les types des entités nommées, ainsi que sur les règles d'annotations. Les différentes règles établies par les annotateurs sont regroupées dans la deuxième partie de ce guide.

## Annotation

Après cette période de formation, chacun des huit annotateurs traite individuellement les vingt documents du sous-corpus. Des réunions sont

organisées ponctuellement tout au long du processus d'annotation pour discuter de points problématiques soulevés par le travail d'annotation.

Quelques exemples des problèmes soulevés :

- Quelles limites au balisage des termes? Par exemple inclure ou non l'article (*Netherlands* ou *The Netherlands*).
- Faut-il tenir compte de toutes les façons d'écrire une date ou bien filtrer? Par exemple: *17 December 1982 - 17 December, 1982 - 1982, 17 December - 17 December in 1982 - etc.*
- Comment gérer le changement de type d'entité nommées en fonction du contexte? Par exemple dans *les décisions prises par Washington*, *Washington* ne désigne pas un lieu mais une instance du pouvoir; ainsi si ce terme est généralement balisé comme un `<placeName>`, il devient un `<orgName>` dans cet exemple.

7

## 1.4 Accord interannotateur

### Méthode

Les entités nommées sont extraites des vingt documents annotés à l'aide d'une feuille de style et les résultats de l'annotation sont regroupés dans des classeurs Excel avec une ligne par annotateur, une colonne par entité nommée et le type d'entité nommée à l'intersection. Les types sont convertis en chiffres.

	A	B	C	D	E	F	G
1	Annotateur	Canada	Baffin Island	Greenland	Jan Mayen	Svalbard	Iceland
2	A	2	3	3	3	3	3
3	B	2	3	2	2	2	2
4	C	2	3	3	3	3	3
5	D	2	3	2	3	3	2
6	E	2	3	2	2	2	2
7	F	2	3	3	3	3	3
8	G	2	3	3	3	3	3
9	H	2	3	3	3	3	3

Extrait d'un fichier Excel regroupant les annotations des huit annotateurs : on note une divergence d'annotation entre `<placeName>` (2) et `<geogName>` (3).

Les divergences donnent lieu à des discussions entre annotateurs pour que chacun exprime les arguments qui sous-tendent les choix effectués.

Les discussions ne résolvent pas toutes les divergences, mais elles permettent de faire progresser chaque annotateur dans sa connaissance du processus d'annotation. Les réunions/discussions entre annotateurs se répètent jusqu'à ce

que chacun soit sûr des choix qu'il a faits dans les types d'entités nommées et les entités nommées elles-mêmes.

Les tableaux sont ensuite convertis en fichiers CSV pour le calcul de l'accord interannoteur. Ces fichiers ne contiennent que des chiffres et indiquent le nombre d'annotateurs, le nombre de types d'entités nommées du document (classes) et le nombre d'entités nommées du document (observables).

Nbannotateurs=8				
Nbclasses=6				
Nbobservables=135				
4	4	2	2	2
4	4	2	2	2
4	4	2	2	2
4	4	2	2	2
4	4	2	2	2
4	4	2	2	2
4	4	2	2	2
4	4	2	2	2

Extrait d'un fichier CSV destiné au calcul de l'accord interannoteur

Les fichiers CSV sont envoyés au LIFAT où des chercheurs expérimentés travaillent sur les mesures d'accord interannoteur.

## Résultats

L'estimation porte sur la valeur du coefficient kappa de Cohen sur le sous-corpus de vingt documents annoté par huit annotateurs (2 153 observations et 17 224 annotations). La version multi-annotateurs du Kappa (Davies & Fleiss, 1982) est utilisée. La mesure globale obtenue est de 0,91 de Kappa (mesuré sur chaque document), ce qui traduit un fort accord entre annotateurs.

```

*****en moyennant sur les 20 fichiers*****
nbtotale d'observations=2153
moyenne des accords entre les 20 fichiers=0.8458
moyenne pond'eree par le nombre d'observations sur chaque fichier=0.869003

*****resultats globaux*****
classes  0      1      2      3      4      5      6      7      8
0      72.86  28.29  115.14  89.14  115.00  112.57  10.43  1.43  7.14
1      28.29  2896.57  1.14  0.29  0.00  1.71  0.00  0.00  0.00
2      115.14  1.14  3749.16  47.43  1.71  56.43  0.00  0.00  0.00
3      89.14  0.29  47.43  2167.15  2.00  0.00  0.00  0.00  0.00
4      115.00  0.00  1.71  2.00  2688.29  2.00  0.00  0.00  0.00
5      112.57  1.71  56.43  0.00  2.00  3960.30  23.57  11.43  0.00
6      10.43  0.00  0.00  0.00  0.00  23.57  350.00  0.00  0.00
7      1.43  0.00  0.00  0.00  0.00  11.43  0.00  5.14  0.00
8      7.14  0.00  0.00  0.00  0.00  0.00  0.00  0.00  80.86
S=17224
alpha calculé sur l'ensemble des observations =0.910509

```

Accord interannoteur et matrice de coïncidences des classes



## 1.5 Corpus annoté

Le premier tableau présente quelques données bibliographiques du corpus, dont on peut trouver une exposition plus détaillée dans l'instance *Corpus EN-ISTEX* publiée dans [data.istex.fr](http://data.istex.fr) (voir [1.6 Publication du Corpus](#)).

Le deuxième tableau présente la répartition par type des entités nommées du corpus.

9

Editeur	Nbre de titres de revues	Années de publication	Nbre de documents structurés
Elsevier	66	1967 à 2002	79 /100
Wiley	68	1980 à 2016	84 /100

Indications bibliographiques du corpus

Editeur	persName	placeName	geogName	date	orgName	orgName type funder	orgName type provider	ref type url
Elsevier	773	2796	1761	2059	1843	151	8	0
Wiley	1897	3631	1209	1840	1522	172	12	46
Total	2670	6427	2970	3899	3365	323	20	46
%	13,53	32,59	16,06	19,77	17,06	1,63	0,1	0,23

Nombre d'entités nommées par type dans le corpus et pourcentages par rapport au nombre total d'entités nommées du corpus (19720)

## 1.6 Publication du corpus

L'instance Corpus EN-ISTEX donne accès aux données bibliographiques des documents du corpus, à travers des graphiques offrant une exploration de ces données. Cette exploration se poursuit sur les entités nommées du corpus annoté.

En page d'accueil, on trouve dans le champ Application la possibilité de télécharger le corpus en fichiers TEI et les offsets des entités nommées.

Le corpus peut également être téléchargé via <https://dl.istex.fr/> (onglet Voir Plus / Télécharger)

L'onglet Ressources permet d'accéder aux données particulières de chacun des documents constitutifs du corpus.

## Règles d'annotation avec les entités nommées d'ISTEX

### 1.7 Personnes : <persName>

#### 1.7.1 Règles

La balise <persName> annote les noms de personnes.

- Dans les textes anonymisés (arrêtés de justice, rapports médicaux, etc.), les initiales remplaçant les noms de personnes (X.X.) ne sont pas reconnues comme entités nommées.
- Les noms de personnes inclus dans une expression ne sont pas étiquetés comme <persName>.

Exemples : *forces de dispersion de London, London dispersion forces, Université François Rabelais, Rue du Général Leclerc, maladie d'Alzheimer, Alzheimer's disease, Ohm's law, Bouguer gravity, Ringer's solution, Knudsen diffusion.*

- Les titres (Mr, Mme, Sir, Dr, etc.) ne sont pas inclus dans la balise, mais on inclut *St* ou *Ste* dans la balise, les saints n'étant désignés que par leur prénom.
- On n'annote pas les fonctions comme: Dauphin of France ou Président de la République.
- Si le nom de personne dans une citation n'est pas inclus dans la balise <ref>, on l'annote s'il fait partie du texte (*Darwin*), mais on ne l'annote pas s'il fait partie de la référence (*Meyer*) : exemple : *an early visit by <persName>Darwin</persName> (<ref type="bibr" target="#jbi12723-bib-0028"/>), and extensive endemism in their biota (Meyer, <ref type="bibr" target="#jbi12723-bib-0073"/>), they have received...*
- On n'annote pas un nom d'auteur suivi de *et al.*

#### 1.7.2 Exemples

##### Français

<i>T. Higuchi propose dans [7, 8] un estimateur...</i>	<i>&lt;persName&gt;T. Higuchi&lt;/persName&gt; propose dans [7, 8] un estimateur...</i>
<i>le professeur Y. Coppens</i>	<i>le professeur &lt;persName&gt;Y. Coppens&lt;/persName&gt;</i>

##### Anglais

<i>We thank Prof. Harry Green and Dr Larissa Dobrzhinetskaya for assistance...</i>	<i>We thank Prof. &lt;persName&gt;Harry Green&lt;/persName&gt; and Dr &lt;persName&gt;Larissa</i>
--	---

	<i>Dobrzhinetskaya</i> <code>&lt;/persName&gt;</code> for assistance...
<i>Raymond M. Gilmore</i>	<code>&lt;persName&gt;</code> <i>Raymond M. Gilmore</i> <code>&lt;/persName&gt;</code>
<i>(later Sir Kenneth Goadby)</i>	<i>(later Sir</i> <code>&lt;persName&gt;</code> <i>Kenneth Goadby</i> <code>&lt;/persName&gt;</code> <i>)</i>
<i>...the birthplace of St Bridget</i>	<i>...the birthplace of</i> <code>&lt;persName&gt;</code> <i>St Bridget</i> <code>&lt;/persName&gt;</code>
<i>...the crucified Christ...</i>	<i>...the crucified</i> <code>&lt;persName&gt;</code> <i>Christ</i> <code>&lt;/persName&gt;</code> <i>...</i>
<i>Indeed, Lucas</i> ( <i>&lt;ref type="bibr" target="#bij12294-bib-0041"/&gt;</i> : 22) states that	<i>Indeed,</i> <code>&lt;persName&gt;</code> <i>Lucas</i> <code>&lt;/persName&gt;</code> ( <i>&lt;ref type="bibr" target="#bij12294-bib-0041"/&gt;</i> : 22) states that
<i>According to Yalow, 76 a distinguished pioneer in medical...</i>	<i>According to Yalow, 76 a distinguished pioneer in medical...</i>
<i>Louis XIV claimed descent from the Gauls,...</i>	<code>&lt;persName&gt;</code> <i>Louis XIV</i> <code>&lt;/persName&gt;</code> claimed descent from the Gauls,...
<i>...in the second half of the Franco years.</i>	<i>...in the second half of the</i> <code>&lt;persName&gt;</code> <i>Franco</i> <code>&lt;/persName&gt;</code> years.

**Note** : Dans *Franco years*, on ne tient pas compte de years parce que cela n'a pas une signification de `<date>` ; Mais un certain nombre de termes qui pourraient suivre *Franco* (*government, dictatorship, regime, ...*) seraient pris en considération dans la balise qui serait alors `<orgName>`.

## 1.8 Lieux liés à l'intervention humaine : `<placeName>`

### 1.8.1 Règle

La balise `<placeName>` annote un nom de lieu administratif (ville, région, pays, etc.).

- Le vocabulaire de l'orientation est inclus dans la balise pour les continents quand il est adjectif qualificatif du nom (*oriental, méridional, septentrional, occidental, eastern, southern, northern, western*), mais ne l'est pas pour les pays, excepté pour *Northern Ireland*.
- Les termes *East, West, North, South* sont inclus dans la balise pour les pays ou états: *East Germany, South Korea, North Carolina*, etc. quand ils ont une majuscule.
- Quand deux entités nommées sont liées par *ET*, la deuxième étant abrégée, on inclut les deux dans la balise.
- Quand deux entités nommées sont liées par un tiret, on les annote séparément.
- Les continents sont toujours des `<placeName>`.
- Pas d'annotation pour les bâtiments: *Drottningholm palace, Palais de l'Élysée, Arc de Triomphe* ne sont pas des `<placeName>`.

- Pas d'annotation pour des termes généraux comme Pays européens / European countries / Western countries / Western societies / Pays arabes / Arab countries / Arab Nations.
- *EEC countries* s'annote : `<orgName>EEC</orgName> countries`

*Australie, Paris, Région parisienne, Bourgneuf, Silicon Valley, European space*, sont des `<placeName>`, ainsi que *Peace River (Rivière-la-Paix)* : ville du Canada, *Clyde River* : village de l'île de Baffin, *Rhode Island* : État des États-Unis.

## 1.8.2 Exemples

### Français

<i>Europe orientale</i>	<code>&lt;placeName&gt;Europe orientale&lt;/placeName&gt;</code>
<i>A l'est de l' Europe</i>	<i>A l'est de l'</i> <code>&lt;placeName&gt;Europe&lt;/placeName&gt;</code>
<i>En France méridionale</i>	<i>En </i> <code>&lt;placeName&gt;France&lt;/placeName&gt;</code> <i> méridionale</i>
<i>Dans l'hémisphère nord</i>	<i>Dans l'</i> <code>&lt;placeName&gt;hémisphère nord&lt;/placeName&gt;</code>
<i>La région Bretagne-Vendée</i>	<i>La région </i> <code>&lt;placeName&gt;Bretagne&lt;/placeName&gt;</code> <i>-</i> <code>&lt;placeName&gt;Vendée&lt;/placeName&gt;</code>
<i>...dans la ville de Paris</i>	<i>...dans la ville de </i> <code>&lt;placeName&gt;Paris&lt;/placeName&gt;</code>
<i>La Guadeloupe</i>	<i>La </i> <code>&lt;placeName&gt;Guadeloupe&lt;/placeName&gt;</code>

Note : L'île de la Guadeloupe est un `<geogName>`, il convient de tenir compte du contexte.

### Anglais

<i>... in northern Sweden</i>	<i>... in northern </i> <code>&lt;placeName&gt;Sweden&lt;/placeName&gt;</code>
<i>... in East Germany</i>	<i>... in </i> <code>&lt;placeName&gt;East Germany&lt;/placeName&gt;</code>
<i>... in northern Europe</i>	<i>... in </i> <code>&lt;placeName&gt;northern Europe&lt;/placeName&gt;</code>
<code>&lt;hi rend="bold"&gt;Western Europe&lt;/hi&gt;</code> <i>(Ghent, Belgium</i> <u>Note</u> : <i>Western Europe</i> est en gras dans le texte	<code>&lt;hi rend="bold"&gt;&lt;placeName&gt;Western Europe&lt;/placeName&gt;&lt;/hi&gt;</code> <code>&lt;placeName&gt;Ghent&lt;/placeName&gt;</code> , <code>&lt;placeName&gt;Belgium&lt;/placeName&gt;</code>
<i>...south-eastern Polynesia, Samoa or Hawaii...</i>	<i>...south-eastern </i> <code>&lt;placeName&gt;Polynesia&lt;/placeName&gt;</code> <i>,</i> <code>&lt;placeName&gt;Samoa&lt;/placeName&gt;</code> <i> or</i> <code>&lt;placeName&gt;Hawaii&lt;/placeName&gt;</code> <i>...</i>
<i>...the California coast,</i>	<i>...the </i> <code>&lt;placeName&gt;California&lt;/placeName&gt;</code> <i> coast,</i>
<i>...the whole southern California borderland,</i>	<i>...the whole southern</i> <code>&lt;placeName&gt;California&lt;/placeName&gt;</code> <i> borderland,</i>
<i>...in Lincoln and Pipestone counties in southwestern Minnesota.</i>	<i>...in </i> <code>&lt;placeName&gt;Lincoln and Pipestone counties&lt;/placeName&gt;</code> <i> in southwestern</i>

	<placeName>Minnesota</placeName>.
... the Kibale National Park (1,500–1,700 m) and parts of Mahale Mountains National Park	... the <placeName>Kibale National Park</placeName> (1,500–1,700 m) and parts of <placeName>Mahale Mountains National Park</placeName>
...six views of Drottningholm palace near Stockholm	...six views of Drottningholm palace near <placeName>Stockholm</placeName>
...the Gothic tribes that conquered Rome.	...the Gothic tribes that conquered <placeName>Rome</placeName>.
... in New York City.	... in <placeName>New York City</placeName>.
... in the City of New York	... in the City of <placeName>New York</placeName>
The urban structure in Greater Stockholm is concentrated...	The urban structure in <placeName>Greater Stockholm</placeName> is concentrated...

Note: Dans un autre contexte *Greater Stockholm* pourrait être reconnu comme étant un <orgName> ;  
Idem pour *Kibale National Park*, exemple : *The direction of Kibale National Park has decided to...*

## 1.9 Lieux liés à la géographie physique : <geogName>

### 1.9.1 Règle

La balise <geogName> annote un nom de lieu géographique, un élément géographique tel qu'un mont, une île, une rivière, un bassin, un lac, etc..

- Quand deux entités nommées sont liées par ET, la deuxième étant abrégée, on inclut les deux dans la balise.
- Quand deux entités nommées sont liées par un tiret, on les annote séparément.
- Dans une expression générale comme *les îles du Pacifique*, aucune île n'étant précisément désignée, on annote seulement *Pacifique*.
- Les termes *East*, *West*, *North*, *South* sont inclus dans la balise quand ils ont une majuscule, comme dans *South Pacific*.
- Bois et forêts sont des <geogName>, *Bois de Boulogne* est un <geogName>.
- Des éléments géographiques comme *river*, *ocean*, *island*, *craton*, *desert*, *valley*, *sea*, suivant ou précédant un nom propre ou une entité qu'ils qualifient, entraînent la balise <geogName>; si le terme *island* n'est pas présent dans l'entité, on annote avec <placeName>. Ces éléments géographiques seuls ne sont pas annotés.

- Attention : *Peace River (Rivière-la-Paix)* est une ville du nord de l'Alberta au Canada, *Clyde River* est un village inuit situé sur le bord de l'île de Baffin, *Rhode Island* est un état des États-Unis, ce sont des `<placeName>`.
- En géologie, on annote les « Formations » qui renvoient bien à des entités géographiques, mais pas les « Groups » qui sont plutôt de l'ordre de la classification :

Exemple: from the `<geogName>Yongancun Formation</geogName>` of the Jiayin Group.

*Baie de Bourgneuf, Bassin parisien, Estuaire de la Gironde* sont des `<geogName>`.

## 1.9.2 Exemples

### Français

<i>... au sud de l'estuaire de la Loire</i>	<i>... au sud de l'<code>&lt;geogName&gt;estuaire de la Loire&lt;/geogName&gt;</code></i>
<i>La fécondité chez <i>S. m. melanotheron</i> du lac d'Ayamé varie...</i>	<i>La fécondité chez <i>S. m. melanotheron</i> du <code>&lt;geogName&gt;lac d'Ayamé&lt;/geogName&gt;</code> varie...</i>
<i>Les océans atlantique et indien</i>	<i>Les <code>&lt;geogName&gt;océans atlantique et indien&lt;/geogName&gt;</code></i>
<i>les îles du Pacifique</i>	<i>les îles du <code>&lt;geogName&gt;Pacifique&lt;/geogName&gt;</code></i>

### Anglais

<i>...located on the coast of the Baltic Sea, on the inlet to Lake Malaren.</i>	<i>...located on the coast of the <code>&lt;geogName&gt;Baltic Sea&lt;/geogName&gt;</code>, on the inlet to <code>&lt;geogName&gt;Lake Malaren&lt;/geogName&gt;</code>.</i>
<i>near Canary Islands</i>	<i>near <code>&lt;geogName&gt;Canary Islands&lt;/geogName&gt;</code></i>
<i>(the southern Cook, Austral, Tuamotu, Marquesas, Gambier or Pitcairn Islands)</i>	<i>(the southern <code>&lt;geogName&gt;Cook, Austral, Tuamotu, Marquesas, Gambier or Pitcairn Islands&lt;/geogName&gt;</code>)</i>
<i>...proximity to the southern Cook, northern Austral and Tuamotu archipelagos,...</i>	<i>...proximity to the southern <code>&lt;placeName&gt;Cook&lt;/placeName&gt;</code>, northern <code>&lt;geogName&gt;Austral and Tuamotu archipelagos&lt;/geogName&gt;</code>,...</i>
<i>dolphins from the Pacific Ocean</i>	<i>dolphins from the <code>&lt;geogName&gt;Pacific Ocean&lt;/geogName&gt;</code></i>
<i>from Lake Ontario to Lake Erie</i>	<i>from <code>&lt;geogName&gt;Lake Ontario&lt;/geogName&gt;</code> to <code>&lt;geogName&gt;Lake Erie&lt;/geogName&gt;</code></i>
<i>Lakes Erie and Ontario</i>	<i><code>&lt;geogName&gt; Lakes Erie and Ontario&lt;/geogName&gt;</code></i>
<i>Great Lakes–St. Lawrence Lowland</i>	<i><code>&lt;geogName&gt;Great Lakes&lt;/geogName&gt;</code>–</i>

	<geogName>St. Lawrence Lowland</geogName>
California Current	<geogName>California Current</geogName>
...are located in the central south Pacific at roughly 16–18° S,	...are located in the central south <geogName>Pacific</geogName> at roughly 16–18° S,
However, they are younger than the Austral and Northwest Hawaiian islands	However, they are younger than the <geogName>Austral and Northwest Hawaiian islands</geogName>
Hawaiian, Marquesas, Cook-Austral and Samoan islands.	<geogName>Hawaiian, Marquesas, Cook-Austral and Samoan islands</geogName>.
...the pelitic sediments of the Serra da Saudade Formation,	...the pelitic sediments of the <geogName>Serra da Saudade Formation</geogName>,
...for the upper Lapa Formation. Samples from the Serra do Poço Verde and Serra do Garrote formations	...for the upper <geogName>Lapa Formation</geogName>. Samples from the <geogName>Serra do Poço Verde</geogName> and <geogName>Serra do Garrote</geogName> formations

## 1.10 Dates et périodes : <date>

### 1.10.1 Règle

La balise <date> annote les dates et les périodes.

- L'élément principal d'une <date> est l'année, elle doit être présente pour qu'il y ait annotation, par exemple *le 25 mai* ou *entre janvier et avril* ne sont pas annotés.
- Les décennies, les siècles et les millénaires sont annotés.
- On annote toute la date si elle est écrite normalement (*le <>24 mai 1978</>*), mais on ne prend que l'année si un mot s'interpose : (*le 24 mai en <>1978</>*) ;
- Pour les périodes, on n'annote que l'année si la période se situe entre deux mois d'une même année et qu'il y a une ellipse (*du 30 juin au 5 septembre <>2001</>*), mais s'il n'y a pas d'ellipse: (*<>du 30 juin 2001 au 5 septembre 2001</>*) on annote toute la période.
- L'expression d'une période doit comporter *from-to*, *from-through*, *for-through*, *between-and*, *de-à*, *entre-et*, sans mots intermédiaires.
- Si une période regroupe deux périodes, chacune est annotée séparément.

- Les périodes géologiques et les périodes historiques sont annotées (voir tableaux) ; on n'annote que le nom de la période et pas les adjectifs qui précèdent (*early, upper, etc.*) ; on n'inclut pas le mot *age* dans la balise, sauf pour *age du fer, age du bronze, age de la pierre*.
- On annote *Late Glacial Maximum, LGM, Late Glacial Cycle, LGC, Late Glacial Interstadial, LGI, Lateglacial*, chaque fois que l'on rencontre ces termes avec des majuscules.
- On n'annote pas les dates
  - qui ne correspondent pas à une année précise, comme les dates précédées de *CA* ("circa" qui veut dire "environ")
  - comportant *BP* ("before present") ou *AP* ("avant le présent") correspondant à des dates obtenues par le carbone 14
- On n'annote pas les événements historiques (*première guerre mondiale, guerre froide...*)

### 1.10.2 Exemples

#### Dates - Français

<i>le 23 juin 1990</i>	<i>le &lt;date&gt;23 juin 1990&lt;/date&gt;</i>
<i>le 23 juin en 1990</i>	<i>le 23 juin en &lt;date&gt;1990&lt;/date&gt;</i>
<i>Dans les années quatre-vingt-dix,</i>	<i>Dans les &lt;date&gt;années quatre-vingt-dix&lt;/date&gt;,</i>
<i>Dans les années 1990,</i>	<i>Dans les &lt;date&gt;années 1990&lt;/date&gt;,</i>
<i>Dans le milieu des années 60</i>	<i>Dans le milieu des &lt;date&gt;années 60&lt;/date&gt;</i>
<i>Dans la première moitié du 14ème siècle</i>	<i>Dans la première moitié du &lt;date&gt;14ème siècle&lt;/date&gt;</i>
<i>... avant le milieu du XXIe siècle,</i>	<i>... avant le milieu du &lt;date&gt;XXIe siècle&lt;/date&gt;,</i>
<i>Au Moyen Age et à la Renaissance</i>	<i>Au &lt;date&gt;Moyen Age&lt;/date&gt; et à la &lt;date&gt;Renaissance&lt;/date&gt;</i>
<i>Avant septembre 1980</i>	<i>Avant &lt;date&gt;septembre 1980&lt;/date&gt;</i>

#### Dates - Anglais

<i>Friday, 14 May 1993 ;</i>	<i>Friday, &lt;date&gt;14 May 1993&lt;/date&gt; ;</i>
<i>on 26 May, 1977</i>	<i>on &lt;date&gt;26 May, 1977&lt;/date&gt;</i>
<i>1620-11-09</i>	<i>&lt;date&gt;1620-11-09&lt;/date&gt;</i>
<i>By the year &lt;date&gt;2012&lt;/date&gt;</i>	<i>By the year &lt;date&gt;2012&lt;/date&gt;</i>
<i>In the 1990s</i>	<i>In the &lt;date&gt;1990s&lt;/date&gt;</i>
<i>In 1994–5 (wave 1),</i>	<i>In &lt;date&gt;1994–5&lt;/date&gt; (wave 1),</i>



<i>as used in late 16th-century</i>	<i>as used in late &lt;date&gt;16th-century&lt;/date&gt;</i>
<i>in May and June 2011</i>	<i>in May and &lt;date&gt;June 2011&lt;/date&gt;</i>
<i>In mid1945</i>	<i>In mid&lt;date&gt;1945&lt;/date&gt;</i>
<i>in November of 1523</i>	<i>in November of &lt;date&gt;1523&lt;/date&gt;</i>
<i>In the mid-fifteenth century</i>	<i>In the mid&lt;date&gt;fifteenth century&lt;/date&gt;</i>
<i>sixteenth- and seventeenth-century</i>	<i>sixteenth- and &lt;date&gt;seventeenth-century&lt;/date&gt; (singulier)</i>
<i>in the fifteenth and sixteenth centuries</i>	<i>in the &lt;date&gt;fifteenth and sixteenth centuries&lt;/date&gt; (pluriel)</i>

## Périodes - Français

<i>Entre 1995 et 2004,</i>	<i>&lt;date&gt;Entre 1995 et 2004&lt;/date&gt;,</i>
<i>de 1990 à 2001</i>	<i>&lt;date&gt;de 1990 à 2001&lt;/date&gt;</i>
<i>De 1986 à 1997, ce même pays</i>	<i>&lt;date&gt;De 1986 à 1997&lt;/date&gt;, ce même pays</i>
<i>de juin 1990 à mai 2001</i>	<i>&lt;date&gt; de juin 1990 à mai 2001&lt;/date&gt;</i>
<i>Pendant la période 1993-94 à 1996-97.</i>	<i>Pendant la période &lt;date&gt;1993-94 à 1996-97&lt;/date&gt;.</i>
<i>d'octobre 2010 à mai 2012</i>	<i>&lt;date&gt;d'octobre 2010 à mai 2012&lt;/date&gt;</i>
<i>du 12 au 31 octobre en 1978</i>	<i>du 12 au 31 octobre en &lt;date&gt;1978&lt;/date&gt;</i>
<i>du 30 juin au 5 septembre 2001</i>	<i>du 30 juin au 5 septembre &lt;date&gt;2001&lt;/date&gt;</i>
<i>du 30 juin 2001 au 5 septembre 2001</i>	<i>&lt;date&gt;du 30 juin 2001 au 5 septembre 2001&lt;/date&gt;</i>
<i>Sur la période 1983-1999,</i>	<i>Sur la période &lt;date&gt;1983-1999&lt;/date&gt;,</i>
<i>du Moyen Age à la Renaissance</i>	<i>&lt;date&gt;du Moyen Age à la Renaissance&lt;/date&gt;</i>
<i>En été 1942</i>	<i>En été &lt;date&gt;1942&lt;/date&gt;</i>

## Périodes - Anglais

<i>around the period 1985-1986,</i>	<i>around the period &lt;date&gt;1985-1986&lt;/date&gt;,</i>
<i>from December 2010 to December 2011</i>	<i>&lt;date&gt;from December 2010 to December 2011&lt;/date&gt;</i>
<i>between 24 November and 13 December 2010</i>	<i>between 24 November and 13 December &lt;date&gt;2010&lt;/date&gt;</i>
<i>the summer of 1977</i>	<i>the summer of &lt;date&gt;1977&lt;/date&gt;</i>
<i>from 10 May 2003 to 30 August 2003</i>	<i>&lt;date&gt;from 10 May 2003 to 30 August 2003&lt;/date&gt;,</i>
<i>the first event (mid-Oligocene to early Miocene)</i>	<i>the first event (mid-&lt;date&gt;Oligocene&lt;/date&gt; to early &lt;date&gt;Miocene&lt;/date&gt;)</i>
<i>from post-Eocene onwards</i>	<i>from post-&lt;date&gt;Eocene&lt;/date&gt; onwards</i>

<i>event (mid-Miocene&lt; to Quaternary)</i>	<i>event (mid-<a href="#">&lt;date&gt;Miocene&lt;/date&gt;</a> to <a href="#">&lt;date&gt;Quaternary&lt;/date&gt;</a>)</i>
<i>from the second quarter of the 17th century to the 19th century</i>	<i>from the second quarter of the <a href="#">&lt;date&gt;17th century&lt;/date&gt;</a> to the <a href="#">&lt;date&gt;19th century&lt;/date&gt;</a></i>
<i>from 1958–1977 to 1978–2004</i>	<i>from <a href="#">&lt;date&gt;1958–1977&lt;/date&gt;</a> to <a href="#">&lt;date&gt;1978–2004&lt;/date&gt;</a></i>
<i>between 1978–2004 and 1958–1977</i>	<i>between <a href="#">&lt;date&gt;1978–2004&lt;/date&gt;</a> and <a href="#">&lt;date&gt;1958–1977&lt;/date&gt;</a></i>
<i>From the year 1799 to 2006...</i>	<i>From the year <a href="#">&lt;date&gt;1799&lt;/date&gt;</a> to <a href="#">&lt;date&gt;2006&lt;/date&gt;</a>.</i>
<i>...from the 1970s to the 1980s.</i>	<i>...from the <a href="#">&lt;date&gt;1970s&lt;/date&gt;</a> to the <a href="#">&lt;date&gt;1980s&lt;/date&gt;</a>.</i>
<i>...data sets covering January 2000 to June 2004...</i>	<i>data sets covering <a href="#">&lt;date&gt;January 2000&lt;/date&gt;</a> to <a href="#">&lt;date&gt;June 2004&lt;/date&gt;</a></i>
<i>we used values for July 2003 to June 2004</i>	<i>...we used values for <a href="#">&lt;date&gt;July 2003&lt;/date&gt;</a> to <a href="#">&lt;date&gt;June 2004&lt;/date&gt;</a>...</i>
<i>...15 May–15 June from 2002 through 2004...</i>	<i>...15 May–15 June <a href="#">&lt;date&gt;</a>from 2002 through 2004<a href="#">&lt;/date&gt;</a>...</i>
<i>...for 1991 through 1998...</i>	<i><a href="#">&lt;date&gt;</a>for 1991 through 1998<a href="#">&lt;/date&gt;</a></i>
<i>...from 1991 through at least 1996...</i>	<i>...from <a href="#">&lt;date&gt;1991&lt;/date&gt;</a> through at least <a href="#">&lt;date&gt;1996&lt;/date&gt;</a>...</i>
<i>...from January 1, 2010 through November 30, 2010</i>	<i>...from January 1, <a href="#">&lt;date&gt;2010&lt;/date&gt;</a> through November 30, <a href="#">&lt;date&gt;2010&lt;/date&gt;</a></i>
<i>...upon a mid- to late-Cinquecento,</i>	<i>...upon a mid- to late-<a href="#">&lt;date&gt;Cinquecento&lt;/date&gt;</a>,</i>
<i>...from the Cambrian to the Lower Silurian.</i>	<i>...from the <a href="#">&lt;date&gt;Cambrian&lt;/date&gt;</a> to the Lower <a href="#">&lt;date&gt;Silurian&lt;/date&gt;</a>.</i>
<i>...sand of early Holocene age.</i>	<i>...sand of early <a href="#">&lt;date&gt;Holocene&lt;/date&gt;</a> age.</i>

## 1.11 Organisations : [<orgName>](#)

### 1.11.1 Règle

La balise [<orgName>](#) annote les organisations, institutions, entreprises, associations, agences nationales et internationales.

- Un terme désignant une organisation peut comporter plusieurs variantes : exemple pour le *Haut-Commissariat des Nations unies pour les réfugiés*, il existe également :
  - en français: HCR, Haut-Commissariat aux Réfugiés, Agence des Nations Unies pour les Réfugiés,
  - en anglais: UNHCR, United Nations High Commissioner for Refugees, UN Refugee Agency.
- Cependant, le lien entre ces variantes ne sera pas établi.
- On annote un gouvernement
    - s'il est spécifié de quel gouvernement il s'agit : *Gouvernement des Etats Unis, US Government, le Gouvernement de Mitterand* (dans la mesure où l'on parle du groupe de personnes qui a gouverné pendant la présidence de mitterand), *Obama's Government*, sont des <orgName>,
    - ou si un adjectif qualifie le gouvernement sans laisser de doute sur le fait qu'il s'agit d'une organisation: *Federal government* est un <orgName>.

Mais on n'annote pas le mot *gouvernement* seul, même avec une majuscule.

- On n'annote pas les bases de données: *PubMed, Embase, BIOSIS, British Nursing Index, Cinahl, Pascal, PsychInfo, Social Science Citation Index and Science Citation Index*, etc.
- On n'annote pas les études, enquêtes, programmes, projets etc.: *Luxembourg Income Study, Agriculture Department Survey, Treatment and Intervention of Psychosis project*, etc., sauf si cela apparaît dans le paragraphe "Acknowledgements".
- Attention: L'étiquetage tient compte du contexte:  
Exemple: La <orgName>Turquie</orgName> déclare la guerre à... ; Je vais passer mes vacances en <placeName>Turquie</placeName>.

*Union Européenne, Amnesty International, CNRS, Ministère de la justice*, sont des <orgName>.

### 1.11.2 Exemples

#### Français

... l'Association Hongroise pour le	... l'<orgName>Association Hongroise pour le
-------------------------------------	--

<i>Développement Communautaire (HACD)</i>	<i>Développement Communautaire&lt;/orgName&gt; (&lt;orgName&gt;HACD&lt;/orgName&gt;)</i>
<i>... le vice-président du Conseil d'Etat.</i>	<i>... le vice-président du &lt;orgName&gt;Conseil d'Etat&lt;/orgName&gt;.</i>
<i>Après que Deutsche Telekom, une des entreprises allemandes...</i>	<i>Après que &lt;orgName&gt;Deutsche Telekom&lt;/orgName&gt;, une des entreprises allemandes...</i>
<i>... contrat de recherche entre la société Turboméca et le laboratoire de mécanique des contacts de l'Insa de Lyon.</i>	<i>... contrat de recherche entre la société &lt;orgName&gt;Turboméca&lt;/orgName&gt; et le &lt;orgName&gt;laboratoire de mécanique des contacts&lt;/orgName&gt; de l'&lt;orgName&gt;Insa&lt;/orgName&gt; de &lt;placeName&gt;Lyon&lt;/placeName&gt;.</i>
<i>Le Département de Psychologie de l'Hôpital de la Pitié-Salpêtrière</i>	<i>Le &lt;orgName&gt;Département de Psychologie de l'Hôpital de la Pitié-Salpêtrière&lt;/orgName&gt;</i>
<i>Département de Psychologie, Hôpital de la Pitié-Salpêtrière</i>	<i>Département de Psychologie, &lt;orgName&gt;Hôpital de la Pitié-Salpêtrière&lt;/orgName&gt;</i>
<i>Service de médecine interne, hôpital Purpan, CHU, place du Docteur-Baylac, 31059 Toulouse cedex, France</i>	<i>Service de médecine interne, &lt;orgName&gt;hôpital Purpan&lt;/orgName&gt;, CHU, place du Docteur-Baylac, 31059 Toulouse cedex, France</i>

## Anglais

<i>Within the American Heart Association</i>	<i>Within the &lt;orgName&gt;American Heart Association&lt;/orgName&gt;</i>
<i>... published by the National Bank of Belgium</i>	<i>... published by the &lt;orgName&gt;National Bank of Belgium&lt;/orgName&gt;</i>
<i>at Department of Psychology, Rutgers University, New Brunswick, NJ</i>	<i>at Department of Psychology, &lt;orgName&gt;Rutgers University &lt;/orgName&gt;, &lt;placeName&gt;New Brunswick&lt;/placeName&gt;, &lt;placeName&gt;NJ&lt;/placeName&gt;</i>

## 1.12 Organismes financeurs : <orgName type="funder">

### 1.12.1 Règle

La balise <orgName type="funder"> annote les organismes qui financent des projets officiels.

L'étiquetage de cette entité nommée faisant usage de la balise <orgName>, on utilise dans ce cas un attribut (type="funder").

On trouve cette entité dans le paragraphe "Remerciements", "Acknowledgements".

### 1.12.2 Exemples

#### Français

21

<i>Cette étude a pu être réalisée grâce au soutien financier du NEFF (Schweizer Nationaler EnergieForschungs-Fonds)</i>	<i>Cette étude a pu être réalisée grâce au soutien financier du &lt;orgName type="funder"&gt;NEFF &lt;/orgName&gt; (&lt;orgName type="funder"&gt;Schweizer Nationaler EnergieForschungs-Fonds &lt;/orgName&gt;)</i>
<i>Cette recherche a été financée par le projet A97H01 du programme de coopération franco-argentinEcos-Secyt et par Total Austral.</i>	<i>Cette recherche a été financée par le &lt;orgName type="funder"&gt;projet A97H01 du programme de coopération franco-argentinEcos-Secyt&lt;/orgName&gt; et par &lt;orgName type="funder"&gt;Total Austral&lt;/orgName&gt;.</i>

#### Anglais

<i>We thank (...)the US National Institutes of Health NIH (GM 23851) for financial...</i>	<i>We thank (...)the &lt;orgName type="funder"&gt;US National Institutes of Health NIH&lt;/orgName&gt; (GM 23851) for financial...</i>
<i>This work was supported by INCO-DC grant IC18CT96-0116 from the European Commission</i>	<i>This work was supported by &lt;orgName type="funder"&gt;INCO-DC&lt;/orgName&gt; grant IC18CT96-0116 from the &lt;orgName&gt;European Commission&lt;/orgName&gt;</i>
<i>This study was funded by the Japan Public Health Association</i>	<i>This study was funded by the &lt;orgName type="funder"&gt;Japan Public Health Association&lt;/orgName&gt;</i>
<i>This work was supported by the Patient Safety Research Portfolio U.K. (grant number: PS024; current controlled trials number: ISRCTN21785299).</i>	<i>This work was supported by the &lt;orgName type="funder"&gt;Patient Safety Research Portfolio U.K.&lt;/orgName&gt; (grant number: PS024; current controlled trials number: ISRCTN21785299).</i>

## 1.13 Organismes hébergeurs de ressources : <orgName type="provider">

### 1.13.1 Règle

La balise <orgName type="provider"> annote les organismes qui hébergent des ressources et les mettent à disposition pour des expériences (par exemple un gros ordinateur, un télescope, mais aussi les produits nécessaires à une expérience...).

L'étiquetage de cette entité nommée faisant usage de la balise <orgName>, on utilise dans ce cas un attribut (type="provider").

On trouve cette entité dans le paragraphe "Remerciements", "Acknowledgements".

### 1.13.2 Exemples

#### Français

<i>...effectuée sur les calculateurs mis à disposition par l'Institut de Développement et de Recherches en Informatique Scientifique,</i>	<i>effectuée sur les calculateurs mis à disposition par l'&lt;orgName type="provider"&gt;Institut de Développement et de Recherches en Informatique Scientifique&lt;/orgName&gt;,</i>
---	---

#### Anglais

<i>... numerical simulations were run on the PARADOX-III cluster hosted by the Scientific Computing Laboratory of the Institute of Physics Belgrade.</i>	<i>... numerical simulations were run on the PARADOX-III cluster hosted by the &lt;orgName type="provider"&gt;Scientific Computing Laboratory of the Institute of Physics Belgrade&lt;/orgName&gt;.</i>
<i>The parallel-beam x-ray tomography experiments were carried out at the beamline for TOMCAT at the SLS in Villigen Switzerland</i>	<i>The parallel-beam x-ray tomography experiments were carried out at the beamline for TOMCAT at the &lt;orgName type="provider"&gt;SLS&lt;/orgName&gt; in &lt;placeName&gt;Villigen&lt;/placeName&gt;, &lt;placeName&gt;Switzerland&lt;/placeName&gt;</i>

## 1.14 URL : <ref type = "url" >

### 1.14.1 Règle

La balise <ref type="url"> annote des identifiants de pages et de sites web.

### 1.14.2 Exemples

#### Français

<i>L'indice mensuel des prix à la consommation publié par l'Institut statistique national de l'Espagne (Instituto Nacional de Estadística de España – <a href="http://www.ine.es">http://www.ine.es</a>)</i>	<i>L'indice mensuel des prix à la consommation publié par l'Institut statistique national de l'Espagne (Instituto Nacional de Estadística de España – &lt;ref type="url"&gt;<a href="http://www.ine.es">http://www.ine.es</a>&lt;/ref&gt;)</i>
--	--

#### Anglais

<i>PENEASY is freely available from <a href="http://www.upc.es/inte/english/download_engALL.htm">http://www.upc.es/inte/english/download_engALL.htm</a>.</i>	<i>PENEASY is freely available from &lt;ref type="url"&gt;<a href="http://www.upc.es/inte/english/download_engALL.htm">http://www.upc.es/inte/english/download_engALL.htm</a>&lt;/ref&gt;.</i>
<i>...a community web site <a href="http://www.scruminresearch.org">www.scruminresearch.org</a> is being built.</i>	<i>...a community web site &lt;ref type="url"&gt;<a href="http://www.scruminresearch.org">www.scruminresearch.org</a>&lt;/ref&gt; is being built.</i>

## 2 Annexe: lexique des périodes

### 2.1 Périodes géologiques

Éon	Ère	Période	Début et fin (en millions d'années)	Durée (en millions d'années)
Phanérozoïque	Cénozoïque	Quaternaire (Pléistocène/ Holocène (de -11 700 ans au présent))	2.588–0	2.588+
		Néogène (Miocène/Pliocène)	23.03–2.588	20.4
		Paléogène (Paléocène/Éocène/ Oligocène)	66.0–23.03	42.9
	Mésozoïque	Crétacé	145.5–66.0	79.5
		Jurassique	201.3–145.0	56.3
		Triassique ou Trias	252.17–201.3	50.9
	Paléozoïque	Permien	298.9–252.17	46.7
		Carbonifère (Mississippien/Pennsylvanien)	358.9–298.9	60
		Dévonien	419.2–358.9	60.3
		Silurien	443.4–419.2	24.2
		Ordovicien	485.4–443.4	42
		Cambrien	541.0–485.4	55.6
	Protérozoïque	Néoprotérozoïque	Édiacarien	635.0–541.0
Cryogénien			850–635	215
Tonien			1000–850	150
Mésoprotérozoïque		Stenien	1200–1000	200
		Ectasien	1400–1200	200
		Calymmien	1600–1400	200
Paléoprotérozoïque		Stathérien	1800–1600	200
		Orosirien	2050–1800	250
		Rhyacien	2300–2050	250
		Sidérien	2500–2300	200

[https://fr.wikipedia.org/wiki/P%C3%A9riode\\_g%C3%A9ologique](https://fr.wikipedia.org/wiki/P%C3%A9riode_g%C3%A9ologique)



## 2.2 Périodes archéologiques

Période	En français	Autre nom
Neolithic	Néolithique	Age de la pierre polie
Mesolithic	Mésolithique	
Paleolithic	Paléolithique	

## 2.3 Périodes historiques, protohistoriques et préhistoriques

Période / anglais	Période / français	Début	Fin
Age of Enlightenment (Europe)	Siècle des Lumières	1715	1789
Renaissance	Renaissance	14 <sup>ème</sup> siècle	16 <sup>ème</sup> siècle
Middle Ages	Moyen Age	5 <sup>ème</sup> siècle	15 <sup>ème</sup> siècle
Antiquity	Antiquité	-3300	5 <sup>ème</sup> siècle
Iron Age	Age du fer	1100 av. J.-C.	800 à 700 av. J.-C.
Bronze Age	Age du Bronze	3000 av. J.-C.	900 av. J.-C.
Stone Age	Age de la pierre	-3,3 Ma	3000 av. J.-C.
Prehistory	Préhistoire	-2,8 Ma	-3300

## 2.4 Périodes historiques Chine, Japon, Corée

L'histoire du Japon et de la Chine traditionnelle se divise en périodes (ou époques) elles-mêmes découpées en ères (ex : période Edo, Ere Ansei), une ère correspond souvent à un règne au Japon, une période à une dynastie en Chine. Quelques exemples.

Période / anglais	Période / français	Début	Fin
	Ere Reiwa	01 mai 2019	En cours
	Ere Heisei	8 janvier 1989	30 avril 2019
	Ere Taishō	30 juillet 1912	25 décembre 1926
	Ere Meiji	1868	1912
Genbun	Ere Genbun (Edo)	Avril 1736	Février 1741
Edo period	Période Edo	1603	1868

Ming dynasty	Période Ming	1368	1644
	Période Sengoku	1467	1615
	Période de Muromachi	1336	1573
Song dynasty	Période Song	960	1297
	Période Heian	794	1185
Nara period	Période Nara	710	794
	Période d'Asuka	592	710