

# Projet ISTE<sup>X</sup>

## Guide d'annotation des entités nommées (LI-INIST)

1

**ISTEX**  
L'excellence documentaire pour tous



# Table des matières

---

## ***PREMIERE PARTIE : HISTORIQUE ET METHODE***

---

<b>1</b>	<b>INTRODUCTION</b>	<b>5</b>
<b>2</b>	<b>L'ANNOTATION</b>	<b>6</b>
2.1	DEFINITION	
2.2	METHODE D'ANNOTATION UTILISEE	
2.3	LE CAS DES DATES ECRITES EN ABREGE	
<b>3</b>	<b>LES ENTITES NOMMEES</b>	<b>8</b>
3.1	DEFINITION	
3.2	ENTITES NOMMEES CIBLES	
3.3	ENTITES NOMMEES COMPLEMENTAIRES	
<b>4</b>	<b>METHODOLOGIE DE LA DETECTION ET DE L'EVALUATION DES ENTITES NOMMEES</b>	<b>9</b>
4.1	DETECTION : UNITEX-CASSYS	
4.2	OPTIMISATION DE LA DETECTION	
4.3	EVALUATION DE LA DETECTION	

## ***DEUXIEME PARTIE : ANNOTATION DES ENTITES NOMMEES D'ISTEX***

---

<b>1</b>	<b>- PERSONNES : &lt;PERSNAME&gt;</b>	<b>13</b>
1.1	REGLE	
1.2	EXEMPLES	
<b>2</b>	<b>- LIEUX LIES A L'INTERVENTION HUMAINE : &lt;PLACENAME&gt;</b>	<b>14</b>
2.1	REGLE	
2.2	EXEMPLES	
<b>3</b>	<b>- LIEUX LIES A LA GEOGRAPHIE PHYSIQUE : &lt;GEOGNAME&gt;</b>	<b>15</b>
3.1	REGLE	
3.2	EXEMPLES	
<b>4</b>	<b>- DATES : &lt;DATE&gt;</b>	<b>16</b>

**4.1 REGLE****4.2 EXEMPLES****5 - ORGANISATIONS : <ORNAME> 17****5.1 REGLE****5.2 EXEMPLES****6 - ORGANISMES FINANCEURS / PROJETS FINANCES : <ORNAME TYPE="FUNDER"> 19****6.1 REGLE****6.2 EXEMPLES****7 - ORGANISMES HÉBERGEURS DE RESSOURCES : <ORNAME TYPE="PROVIDER"> 20****7.1 REGLE****7.2 EXEMPLES****8 - URL : <REF TYPE ="URL"> 21****8.1 REGLE****8.2 EXEMPLES****9 - POINTEURS VERS LES RÉFÉRENCES BIBLIOGRAPHIQUES <REF TYPE="BIBR"> 22****9.1 RÈGLE****9.2 EXEMPLES****10 - REFERENCES BIBLIOGRAPHIQUES : <BIBL> 23****10.1 REGLE****10.2 EXEMPLES**

## ISTEX : initiative d'excellence en Information scientifique et technique

L'Investissement d'Avenir ISTEX est plus qu'une bibliothèque : il en a certes toutes les caractéristiques en offrant l'archive numérique rétrospective de la Science mais, bien au-delà, le projet ISTEX entend offrir tous les moyens accessibles de consultation et d'analyse aujourd'hui disponibles dans toutes les communautés de la Science. Le Comité Exécutif d'ISTEX réunit toutes les compétences nécessaires à l'innovation dans l'analyse et dans le partage de tous les grands champs de connaissance : l'Agence Bibliographique de l'Enseignement Supérieur (ABES), le Consortium Unifié des Etablissements Universitaires et de Recherche pour l'Accès aux Publications Numériques (Couperin.org), et l'Université de Lorraine agissant pour le compte de la Conférence des Présidents d'Université (CPU) travaillent main dans la main avec le Centre National de la Recherche Scientifique (CNRS) pour assembler les ressources, définir les projets et organiser les services qui donneront au projet ISTEX une valeur ajoutée inédite dans le champ toujours croissant des grandes archives numériques. La recherche est évidemment la clé irremplaçable de ce développement. Autour du LORIA, de l'ATILF sous la haute compétence de Jean-Marie PIERREL, Professeur des Universités, les équipes scientifiques nationales concernées par l'analyse des corpus scientifiques travaillent continument à la définition des services qui permettront de « mieux partager les connaissances » quand ISTEX sera en service. Au nom de tout le Comité Exécutif ISTEX, je souhaite adresser au Commissariat d'Investissement d'Avenir les salutations et les remerciements de toute l'équipe de ce projet innovant qui apportera à toutes les communautés de la recherche les perspectives et les outils d'une grande plateforme multi-usages ouverte sur l'avenir.

Renaud FABRE  
Chef du projet ISTEX  
Directeur de l'Information Scientifique et Technique au CNRS



Financement : ANR-10-IDEX-0004-02

## *Première partie : Historique et Méthode*

---

### 1 Introduction

---

Ce guide a pour objet de préciser les choix d'annotation effectués conjointement par le Laboratoire d'Informatique (LI) de l'Université François Rabelais de Tours et l'INIST-CNRS de Vandœuvre-lès-Nancy lors de l'identification des entités nommées dans des corpus constitués à partir de documents scientifiques puisés dans le fonds ISTEEX.

Des essais de constitution de corpus commencent en juin 2014.

Les premiers corpus sont constitués de formats natifs des éditeurs Elsevier et Springer.

Un corpus exploitable est livré au LI en septembre et permet la construction des graphes (Unitex) pour la détection des entités nommées dans des documents en anglais.

Les graphes pour la détection des entités nommées dans des documents en français existent au LI préalablement au projet ISTEEX où ils sont utilisés moyennant une légère adaptation.

Un corpus général (multi-domaines) anglais de 100 documents et un corpus général français de 100 documents servent à l'optimisation de la détection des entités nommées retenues dans le cadre du projet ISTEEX.

Deux autres corpus servent ensuite à l'évaluation de la méthode. Cette évaluation est effectuée en décembre 2015 et janvier 2016. Il s'agit d'un corpus général anglais de 49 documents et d'un corpus général français de 40 documents.

La cascade de graphes permettant la détection d'entités nommées en anglais est créée au LI pour le projet ISTEEX.

La cascade de graphes permettant la détection d'entités nommées en français est révisée au cours de l'année 2017 de manière à correspondre aux changements requis par l'évolution du projet ISTEEX. Une nouvelle évaluation de cette cascade est réalisée en septembre 2017.

## 2 L'annotation

### 2.1 Définition

Il s'agit de reconnaître, désambiguïser et étiqueter des entités nommées dans des documents textuels scientifiques.

6

### 2.2 Méthode d'annotation utilisée

Les entités nommées sont étiquetées avec des balises XML selon le standard TEI (Text Encoding Initiative), qui permet l'échange de données textuelles stockées sous forme électronique par des utilisateurs utilisant des systèmes informatiques différents.

Voir le *TEI Guidelines* : <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html>

Le projet ISTEEX fait usage d'une TEI simplifiée : les balises de détail ne sont pas incluses dans l'étiquetage.

#### 2.2.1 Exemples

##### *Les noms de personnes*

Texte à étiqueter : *Franklin Delano Roosevelt*.

On a choisi pour ISTEEX un étiquetage qui ne laisse apparaître que la balise principale `<persName>`. Les balises de détail (`<surname>` pour le nom de famille ou patronyme, `<forename>` pour le prénom) n'apparaissent pas.

TEI	<pre>&lt;persName&gt; &lt;forename&gt;Franklin&lt;/forename&gt; &lt;forename&gt;Delano&lt;/forename&gt; &lt;surname&gt;Roosevelt&lt;/surname&gt; &lt;/persName&gt;</pre>
ISTEEX	<pre>&lt;persName&gt;Franklin Delano Roosevelt&lt;/persName&gt;</pre>

##### *Les noms de lieux*

###### ➤ Exemple 1 :

Texte à étiqueter : *à 20 km au nord de Paris*.

Là encore, l'étiquetage ne laisse apparaître que les balises principales.

TEI	<code>&lt;placeName&gt; &lt;measure&gt;20 km&lt;/measure&gt; &lt;offset&gt;au nord de&lt;/offset&gt; &lt;settlement type="city"&gt;Paris&lt;/settlement&gt; &lt;/placeName&gt;</code>
ISTEX	<code>à 20 km au nord de&lt;placeName&gt;Paris&lt;/placeName&gt;</code>

➤ Exemple 2 :

Texte à étiqueter : *la partie sud de Madagascar*

TEI	<code>la &lt;placeName&gt;&lt;offset&gt;partie Sud de&lt;/offset&gt; Madagascar&lt;/placeName&gt;</code>
ISTEX	<code>la partie Sud de &lt;placeName&gt;Madagascar&lt;/placeName&gt;</code>

➤ Exemple 3 :

Texte à étiqueter : *au sud de l'estuaire de la Loire*

Les attributs dans les balises (`type="river"`), ainsi que les balises de détail (`<geogFeat>`) n'apparaissent pas.

TEI	<code>&lt;offset&gt;au sud de&lt;/offset&gt; l'&lt;geogName type="river"&gt;&lt;geogFeat&gt;estuaire &lt;/geogFeat&gt; de la &lt;name&gt;Loire&lt;/name&gt;&lt;/geogName&gt;</code>
ISTEX	<code>au sud de l'&lt;geogName&gt;estuaire de la Loire&lt;/geogName&gt;.</code>

## 2.3 Le cas des dates écrites en abrégé

On rencontre souvent dans les textes des dates abrégées en 2 chiffres : 1990/91.

On ne pourra pas traiter un cas comme : *des épidémies en 1992, 93 et 97.*

Dans cet exemple, seul « 1992 » recevra une balise `<date>`. On ne tient pas compte de « 93 », ni de « 97 » parce que créer un graphe pour reconnaître ces années abrégées risquerait de générer plus d'erreurs que d'en éviter.

En effet dans l'exemple : *En 1998, 27 cas d'infections...*, « 27 » n'est pas une date. Cette formulation pouvant se révéler fréquente dans les textes, on fait le choix de ne pas reconnaître les formes de dates ambiguës.

Mais le non-étiquetage de "93" et "97" sera comptabilisé comme 2 erreurs d'omission lors de l'évaluation puisqu'il s'agit bien de dates.

## 3 Les entités nommées

---

### 3.1 Définition

Les entités nommées sont des mots ou groupes de mots présentant un intérêt particulier sur le plan sémantique et pouvant être catégorisés. Il s'agit de noms propres (personnes, lieux, organisations) ou d'autres unités lexicales (dates, projets financés, URL).

8

### 3.2 Entités nommées cibles

Les entités nommées retenues pour ISTEEX lors de la réunion préparatoire du 30 septembre 2014 sont :

- Les noms de personnes : <persName>
- Les noms de lieux : <placeName> et <geogName>
- Les noms d'organisations : <orgName>
- Les dates : <date>
- Les organismes financeurs et projets financés : <orgName type="funder">
- Les URL : <ref type="url">

### 3.3 Entités nommées complémentaires

D'autres entités nommées sont retenues au fur et à mesure de l'avancement du travail :

- Les pointeurs vers les références bibliographiques : <ref type="bibr"> et les références bibliographiques : <bibl>
- Les organismes hébergeurs de ressources : <orgName type="provider">



## 4 Méthodologie de la détection et de l'évaluation des entités nommées

---

### 4.1 Détection : Unitex-CasSys

9

**Unitex** est un logiciel permettant un traitement automatique de corpus de textes en langue naturelle à l'aide de ressources linguistiques. Il utilise entre autre des dictionnaires compressés (.bin) qui sont appliqués selon une règle de priorité (un mot trouvé dans un dictionnaire n'est plus pris en compte quand les dictionnaires de priorité inférieure sont appliqués) et en respectant les majuscules et les espaces.

Unitex utilise ensuite des grammaires locales (réseaux récurrents de transition) possédant un alphabet de symboles terminaux, un alphabet de symboles non terminaux et un ensemble de règles représentées sous forme de graphes. Ces grammaires locales permettent différents niveaux d'analyse des textes (caractère, mot, groupe de mots).

Unitex est un logiciel gratuit de licence libre.

<http://unitexgramlab.org/>

**CasEN** est une cascade de reconnaissance des entités nommées ; elle est implantée avec le programme de création de cascades **CasSys** présent sur la plate-forme Unitex.

CasSys est un outil d'exploration et d'annotation de corpus dont le prototype a été créé au Laboratoire d'Informatique de Tours.

La Société Ergonotics participe à l'optimisation du package informatique d'Unitex-CasSys dans le cadre du projet ISTEEX.

### 4.2 Optimisation de la détection

#### 4.2.1 Méthode

L'interaction entre le LI de Tours et l'équipe ISTEEX-RD de l'INIST s'est organisée de manière à optimiser les règles de détection des entités nommées :

- a) ISTEEX-RD : Détection d'erreurs / Correction d'erreurs : un type d'erreur est repéré dans un ou plusieurs documents et il est signalé dans son contexte d'apparition.
- b) LI : Modification des graphes ayant pour finalité la disparition, si possible, des erreurs signalées.
- c) Le processus a + b est poursuivi jusqu'à ce que l'outil soit suffisamment amélioré pour procéder à une évaluation.

## 4.2.2 Types d'erreurs

- Erreur de frontière: les balises d'une entité nommée sont mal placées.
- Erreur de typage, ou de catégorie: l'étiquette d'une entité nommée n'est pas celle qui lui correspond.
- Erreur de typage + frontière.
- Erreur d'insertion: un terme neutre est reconnu comme une entité nommée.
- Suppression: une entité nommée n'est pas reconnue.

## 4.2.3 Exemples d'erreurs

### ➤ **Exemple 1 : erreur de frontière**

comme celles qu'envisage le `<orgName>Parlement</orgName>` européen :

L'organisation citée est « Parlement européen » ; la balise fermante doit être placée après « européen ».

### ➤ **Exemple 2 : erreur de typage**

A `<persName>Santa Catarina</persName>` :

Santa Catarina est un état du sud du Brésil, c'est un `<placeName>`.

### ➤ **Exemple 3 : erreur de typage/frontière**

into the Gulf of `<placeName>Mexico</placeName>` :

L'entité nommée est « Gulf of Mexico » qui est un `<geogName>`.

### ➤ **Exemple 4: erreur d'insertion**

`<persName>I. Experimental</persName>` study :

Dans ce titre, le grand "I" suivi de « Experimental » est reconnu comme étant l'initiale d'un prénom suivi du nom de famille. C'est pourquoi il est étiqueté `<persName>`.

### ➤ **Exemple 5 : erreur d'insertion**

the energetical value of `<placeName>London</placeName>` forces

« London » est ici un nom de personne : il s'agit du physicien Fritz London. Mais il est inclus dans un terme (« London forces ») et ne peut pas être reconnu comme une entité nommée. « London forces », « forces de London », « forces de dispersion de London », sont des variantes d'un même terme qui n'appartient pas aux types d'entités nommées retenus dans ISTEEX.

### ➤ **Exemple 6 : erreur de suppression**

financée par la <orgName>Fondation Russell Sage</orgName> entre 2004 et 2007 :  
la période « entre 2004 et 2009 » n'a pas été reconnue comme étant une date.

➤ **Exemple 7 : erreur de suppression**

*Pal (2004) a constaté :*

La référence bibliographique n'a pas été balisée.

11

➤ **Exemple 8 : erreur de suppression**

*± 4,04 % au col) [57].*

La référence bibliographique [57] n'a pas été balisée.

## 4.3 Evaluation de la détection

### 4.3.1 Méthode

La détection des entités nommées ne s'effectue que dans le corps du texte (body). Les erreurs hors du corps du texte seront ignorées. Les erreurs sur une citation dans une autre langue que celle de l'article seront prises en considération.

### 4.3.2 Eléments pris en compte

- D = entités non reconnues
- I = erreur d'insertion
- T = erreur de typage
- E = erreur de frontière
- TE = erreur de typage + frontière
- S = entités détectées
- R = entités réelles

### 4.3.3 Calculs

- Le Slot Error Rate (SER) :  $D + I + TE + 0,5 (T+E) / R$
- Le rappel :  $(S-I)/R$
- La précision :  $(S-I)/S$
- La précision du typage :  $(S-I-T-TE)/S$
- La précision du balisage :  $(S-I-E-TE)/S$

## 4.3.4 Evaluation décembre 2015 – janvier 2016

### 4.3.4.1 Corpus anglais

#### *Le corpus*

- 49 documents TEI issus du fonds ISTEEX.
- Etiquetage par l'équipe ISTEEX-RD pour avoir un corpus de référence.

12

#### *Les résultats*

Types d'erreurs	D	I	T	E	TE	S	R
<b>Résultats</b>	1516	281	64	265	131	3296	5414

Calculs	SER	Rappel	Précision	Précision du typage	Précision du balisage
<b>Résultats</b>	38,6 %	55,7 %	91,5 %	85,6 %	79,5 %

### 4.3.4.2 Corpus français

#### *Le corpus*

- 40 documents TEI issus du fonds ISTEEX.
- Etiquetage par l'équipe ISTEEX-RD pour avoir un corpus de référence.

#### *Les résultats*

Types d'erreurs	D	I	T	E	TE	S	R
<b>Résultats</b>	1017	486	15	202	76	3841	4754

Calculs	SER	Rappel	Précision	Précision du typage	Précision du balisage
<b>Résultats</b>	35,5%	70,6%	87,3%	85%	80,1%

## Deuxième partie : Annotation des entités nommées d'ISTEX

---

### 1 - Personnes : <persName>

---

#### 1.1 Règle

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 », « Université François Rabelais », « Rue du Général Leclerc ».

Tous les noms doivent être étiquetés séparément : *Chun et Seban* est étiqueté : <persName>Chun</persName> et <persName>Seban</persName>.

#### 1.2 Exemples

##### Français

➤ **Exemple 1 :**

<persName>T. Higuchi</persName> propose dans <ref type="bibr">[7, 8]</ref> un estimateur...

➤ **Exemple 2 :**

le professeur <persName>Y. Coppens</persName>

##### Anglais

➤ **Exemple 1 :**

We thank Prof. <persName>Harry Green</persName> and Dr <persName>Larissa Dobrzhinetskaya</persName> for assistance

➤ **Exemple 2 :**

`<persName>`Raymond M. Gilmore`</persName>`

➤ **Exemple 3 :**

(later Sir `<persName>`Kenneth Goadby`</persName>`)

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

---

### 2.1 Règle

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

« Australie », « Paris », « Région parisienne », « Bourgneuf » sont des `<placeName>`.

### 2.2 Exemples

#### Français

➤ **Exemple 1 :**

`<placeName>`Europe orientale`</placeName>`

➤ **Exemple 2 :**

A l'est de l'`<placeName>`Europe`</placeName>`

#### Anglais

➤ **Exemple 1 :**

... in northern `<placeName>`Sweden`</placeName>`

➤ **Exemple 2 :**

The urban structure in `<placeName>`Greater Stockholm`</placeName>` is concentrated...

(Remarque : Dans un autre contexte « Greater Stockholm » pourrait être reconnu comme étant un `<orgName>`.)

## 3 - Lieux liés à la géographie physique : <geogName>

---

### 3.1 Règle

La balise <geogName> annote un nom de lieu géographique, désignant un lieu "naturel" (par opposition au placeName).

« Baie de Bourgneuf », « Bassin parisien », « Estuaire de la Gironde » sont des <geogName>.

15

### 3.2 Exemples

#### Français

➤ **Exemple 1 :**

... au sud de l'<geogName>estuaire de la Loire</geogName>,

➤ **Exemple 2 :**

La fécondité chez *S. m. melanotheron* du <geogName>lac d'Ayamé</geogName> varie...

#### Anglais

➤ **Exemple 1 :**

...located on the coast of the <geogName>Baltic Sea</geogName>, on the inlet to <geogName>Lake Malaren</geogName>.

➤ **Exemple 2 :**

near <geogName>Canary Islands</geogName>

➤ **Exemple 3 :**

dolphins from the <geogName>Pacific Ocean</geogName>

## 4 - Dates : <date>

---

### 4.1 Règle

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

Dans les dates comportant jour/mois/année, seule l'année est prise en compte.

Les décennies, les siècles et les millénaires sont également pris en considération.

16

### 4.2 Exemples

#### Français

➤ **Exemple 1 :**

*le 23 juin <date>1990</date>*

➤ **Exemple 2 :**

*Dans les <date>années quatre-vingt-dix</date> ,*

➤ **Exemple 3 :**

*Dans la première moitié du <date>14ème siècle</date>*

➤ **Exemple 4 :**

*... avant le milieu du <date>XXIe siècle</date> ,*

➤ **Exemple 5 :**

*son lancement en <date>2006</date>-07*

➤ **Exemple 6 :**

*<date>Entre 1995 et 2004</date> ,*

➤ **Exemple 7 :**

*<date>De 1986 à 1997</date> , ce même pays*

➤ **Exemple 8 :**



Sur la période `<date>1983-1999</date>`,

➤ **Exemple 9 :**

Pendant la période `<date>1993-94 à 1996-97</date>`.

## Anglais

➤ **Exemple 1 :**

Friday, 14 May `<date>1993</date>`;

➤ **Exemple 2 :**

`<date>1620</date>`-11-09

➤ **Exemple 3 :**

In `<date>1994</date>`-5 (wave 1),

➤ **Exemple 4 :**

around the period `<date>1985-1986</date>`,

➤ **Exemple 5 :**

as used in late `<date>16th-century</date>`

## 5 - Organisations : `<orgName>`

---

### 5.1 Règle

La balise `<orgName>` annote les organisations, regroupant les institutions, les entreprises, les associations, les agences nationales et internationales.


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

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.

 Attention ! L'étiquetage tient compte du contexte :  
*La <orgName>Turquie</orgName> a déclaré la guerre à ...*  
*Je vais passer mes vacances en <placeName>Turquie</placeName>.*

## 5.2 Exemples

### Français

➤ **Exemple 1 :**

... l'<orgName>Association Hongroise pour le Développement Communautaire</orgName>  
(<orgName>HACD</orgName>)

➤ **Exemple 2 :**

... le vice-président du <orgName>Conseil d'Etat</orgName>.

➤ **Exemple 3 :**

Après que <orgName>Deutsche Telekom</orgName>, une des entreprises allemandes...

➤ **Exemple 4 :**

... contrat de recherche entre la société <orgName>Turboméca</orgName> et le  
<orgName>laboratoire de mécanique des contacts</orgName> de l'<orgName>Insa de  
Lyon</orgName>.

➤ **Exemple 5 :**

Département de Psychologie, <orgName>Université du Québec à Montréal</orgName>,  
<orgName>UQAM</orgName>, <placeName>Québec</placeName>,  
<placeName>Montréal</placeName>

➤ **Exemple 6 :**

Service de médecine interne, <orgName>hôpital Purpan</orgName>, CHU, place du Docteur-  
Baylac, 31059 Toulouse cedex, France

## Anglais

➤ **Exemple 1 :**

Within the `<orgName>American Heart Association</orgName>`

➤ **Exemple 2 :**

published by the `<orgName>National Bank of Belgium</orgName>`

➤ **Exemple 3 :**

at Department of Psychology, `<orgName>Rutgers University </orgName>`, `<placeName>New Brunswick</placeName>`, `<placeName>NJ</placeName>`

## 6 - Organismes financeurs / Projets financés : `<orgName type="funder">`

---

### 6.1 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"`).

Les organismes financeurs sont généralement suivis par un numéro qui désigne le projet financé. Si l'organisme financeur est immédiatement suivi du numéro du projet financé, les deux sont inclus dans la balise. Si l'organisme financeur et le numéro du projet financé sont séparés, ils sont tous les deux étiquetés avec la balise `<orgName type="funder">`.

### 6.2 Exemples

#### Français

➤ **Exemple 1 :**

Cette étude a pu être réalisée grâce au soutien financier du `<orgName type="funder">NEFF </orgName>` (`<orgName type="funder">Schweizer Nationaler EnergieForschungs-Fonds </orgName>`)

➤ **Exemple 2 :**

Cette recherche a été financée par le `<orgName type="funder">`projet A97H01 du programme de coopération franco-argentinEcos-Secyt`</orgName>` et par `<orgName type="funder">`Total Austral`</orgName>`.

## Anglais

### ➤ Exemple 1 :

We thank (...)the `<orgName type="funder">`US National Institutes of Health NIH (GM 23851)`</orgName>` for financial

### ➤ Exemple 2:

This work was supported by `<orgName type="funder">`INCO-DC grant IC18CT96-0116`</orgName>` from the `<orgName>` European Commission`</orgName>`

### ➤ Exemple 3:

This study was funded by the `<orgName type="funder">`Japan Public Health Association`</orgName>`

## 7 - Organismes hébergeurs de ressources : `<orgName type="provider">`

---

### 7.1 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 calculateur, 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"`). Cette balise est utilisée de manière identique pour l'organisme hébergeur et pour la ressource.

### 7.2 Exemples

## Français

### ➤ Exemple 1 :

effectuée sur les calculateurs mis à disposition par l'`<orgName type="provider">`Institut de Développement et de Recherches en Informatique Scientifique`</orgName>`,

## Anglais

### ➤ Exemple 1 :

... numerical simulations were run on the `<orgName type="provider">PARADOX-III cluster</orgName>` hosted by the `<orgName type="provider">Scientific Computing Laboratory of the Institute of Physics Belgrade</orgName>`.

### ➤ Exemple 2 :

The parallel-beam x-ray tomography experiments were carried out at the beamline for `<orgName type="provider">TOMCAT at the SLS in Villigen, Switzerland</orgName>`

21

## 8 - URL : `<ref type="url">`

---

### 8.1 Règle

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

### 8.2 Exemples

#### Français

### ➤ Exemple 1 :

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

#### Anglais

### ➤ Exemple 1 :

PENEASY is freely available from `<ref type="url">http://www.upc.es/inte/english/download_engALL.htm</ref>`.

### ➤ Exemple 2 :

a community web site `<ref type="url">www.scruminresearch.org</ref>` is being built.

## 9 - Pointeurs vers les références bibliographiques <ref type="bibr">

---

### 9.1 Règle

La balise <ref type="bibr"> annote dans le corps du texte les pointeurs vers les références bibliographiques se trouvant en bas de page ou en fin de document.

22

### 9.2 Exemples

#### Français

➤ **Exemple 1 :**

(<ref type="bibr">Yan [10]</ref>)

➤ **Exemple 2 :**

(voir <ref type="bibr">Salditt, Whiteford et Adema, 2007</ref>)

➤ **Exemple 3 :**

<ref type="bibr">Kaoua et al. [7]</ref>

➤ **Exemple 4 :**

(<ref type="bibr">OMS, 2001</ref>).

#### Anglais

➤ **Exemple 1 :**

proposed by <ref type="bibr">[Broyden12]</ref>

➤ **Exemple 2 :**

... are fewer in number<ref type="bibr">[2-4]</ref> and ...

➤ **Exemple 3 :**

<ref type="bibr">Baumann et Thiele [3, 4]</ref> négligent pour le film liquide

➤ **Exemple 5 :**

([<ref type="bibr">Gage and Margolin 2000</ref>](#))

➤ **Exemple 6 :**

... as described in [<ref type="bibr">Grimnes and Martinsen \(2008\) </ref>](#)

23

➤ **Exemple 7 :**

...droplet model of [<ref type="bibr">Myers and Swiatecki \(Ann. Phys. 84, 186; 1974\)</ref>](#)

➤ **Exemple 8 :**

[<ref type="bibr">Beezhold et al. \(1994\)</ref>](#)

➤ **Exemple 9 :**

[<ref type="bibr"> Smith et al., 1999a</ref>](#)

➤ **Exemple 10 :**

[<ref type="bibr">Staaav \(1990, unpublished\)</ref>](#)

➤ **Exemple 11 :**

[<ref type="bibr">Municipality of Stockholm, 1991; unpublished</ref>](#))

➤ **Exemple 12 :**

([<ref type="bibr">M. J. Schmitt and D. J. Tipper, unpublished results</ref>](#))

## 10 - Références bibliographiques : <bibl>

---

### 10.1 Règle

La balise <bibl> annote la référence bibliographique qui apparaît en totalité dans le corps du texte (Nom + titre + date) et qui ne renvoie pas vers les références bibliographiques de fin de texte ou de bas de page.

## 10.2 Exemples

### Français

➤ **Exemple :**

(Voir `<bib>`GIRODET, Jean (2007). *Dictionnaire des pièges et difficultés de la langue française*, Paris : Bordas`</bib>`).

24

### Anglais

➤ **Exemple :**

«... *the scene of our finitude, the place where we encounter the limits of our subjectivity*».  
`<bib>`Diane Michelfelder et Richard Palmer, *Dialogue and Déconstruction. The Gadamer-Derrida Encounter*, Albany, Suny Press, 1989, p. XI`</bib>`. (Trad.fr., François Doyon).