



## NooJ pour l'Intelligence économique

Philippe Lambert

### ► To cite this version:

Philippe Lambert. NooJ pour l'Intelligence économique. École thématique. Semaine de formation NooJ à l'INALCO, INALCO, 2012, pp.14. cel-00677782

**HAL Id: cel-00677782**

**<https://cel.hal.science/cel-00677782>**

Submitted on 9 Mar 2012

**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.

# NooJ pour l'intelligence économique

Philippe Lambert  
Institut Jean Lamour  
[Philippe.lambert@ijl.nancy-universite.fr](mailto:Philippe.lambert@ijl.nancy-universite.fr)



# PLAN



**I. Contexte : l'intelligence économique**

**II. Cas d'études :**

**II. A.: Identification d'experts vietnamiens**

**II. B. : Traitement de questions ouvertes  
d'un questionnaire**

**III. Perspectives de recherche**

# I. : L'intelligence économique



**1/ L'ensemble des activités coordonnées de collecte, de traitement et de diffusion de l'information utile aux acteurs économiques en vue de son exploitation**

**(Rapport Martre 1994)**

**2/ Réduction du niveau de l'incertitude dans la prise de décision**

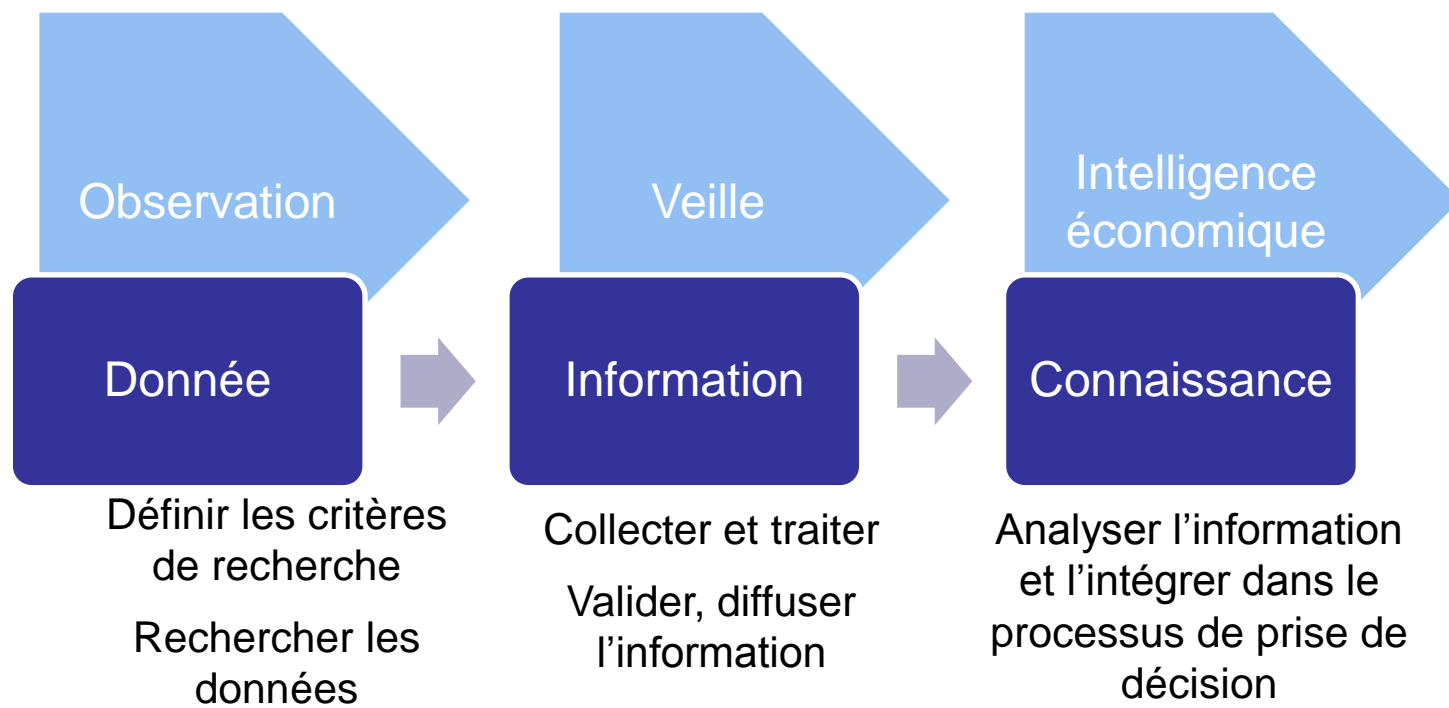
**(Nombreux auteurs)**

**3/ Donner la bonne information à la bonne personne au bon moment**

# I. : L'intelligence économique



## Le cycle de l'IE



# II. A.: Identification d'experts



Scénario : Implantation d'une entreprise au Vietnam



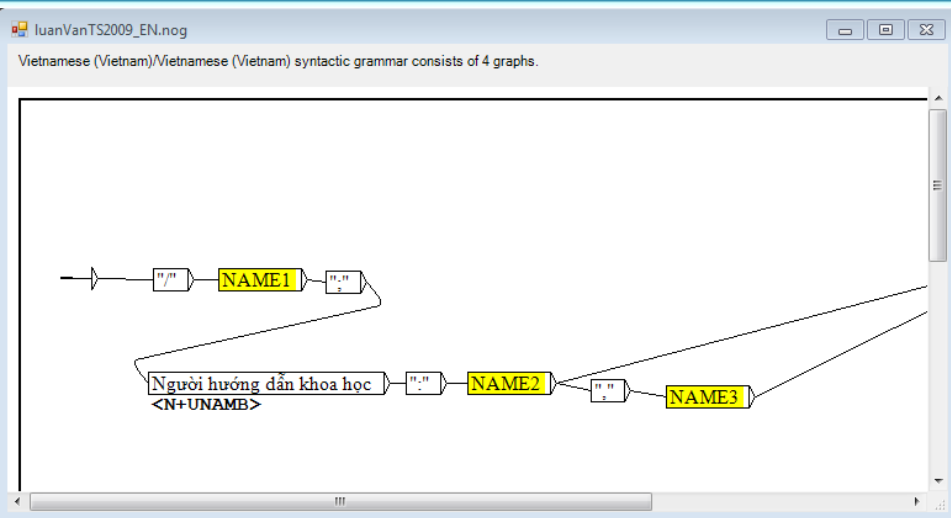
Trouver un point de contact local

The screenshot shows the homepage of the Ho Chi Minh City General Library website. The header features the library's name in Vietnamese and English, along with a search bar. The main content area is divided into several sections:

- Các trang chính** (Main pages): Includes links to search, services, guides, online activities, library events, and contact information.
- Tin tức** (News): A list of recent news items, including announcements about library events and services.
- Giới thiệu thư viện** (Library introduction): A section providing an overview of the library's history, mission, and services.
- Hoạt động thư viện** (Library activities): A section highlighting various programs and events organized by the library.
- Thư viện Lưu động** (Mobile library): A section dedicated to the library's mobile services and outreach programs.

The website also features logos for various partners and sponsors, including PERii, ProQuest, EBSCO, and FORCE. The footer contains contact information and a copyright notice.

## II. A.: Identification d'experts

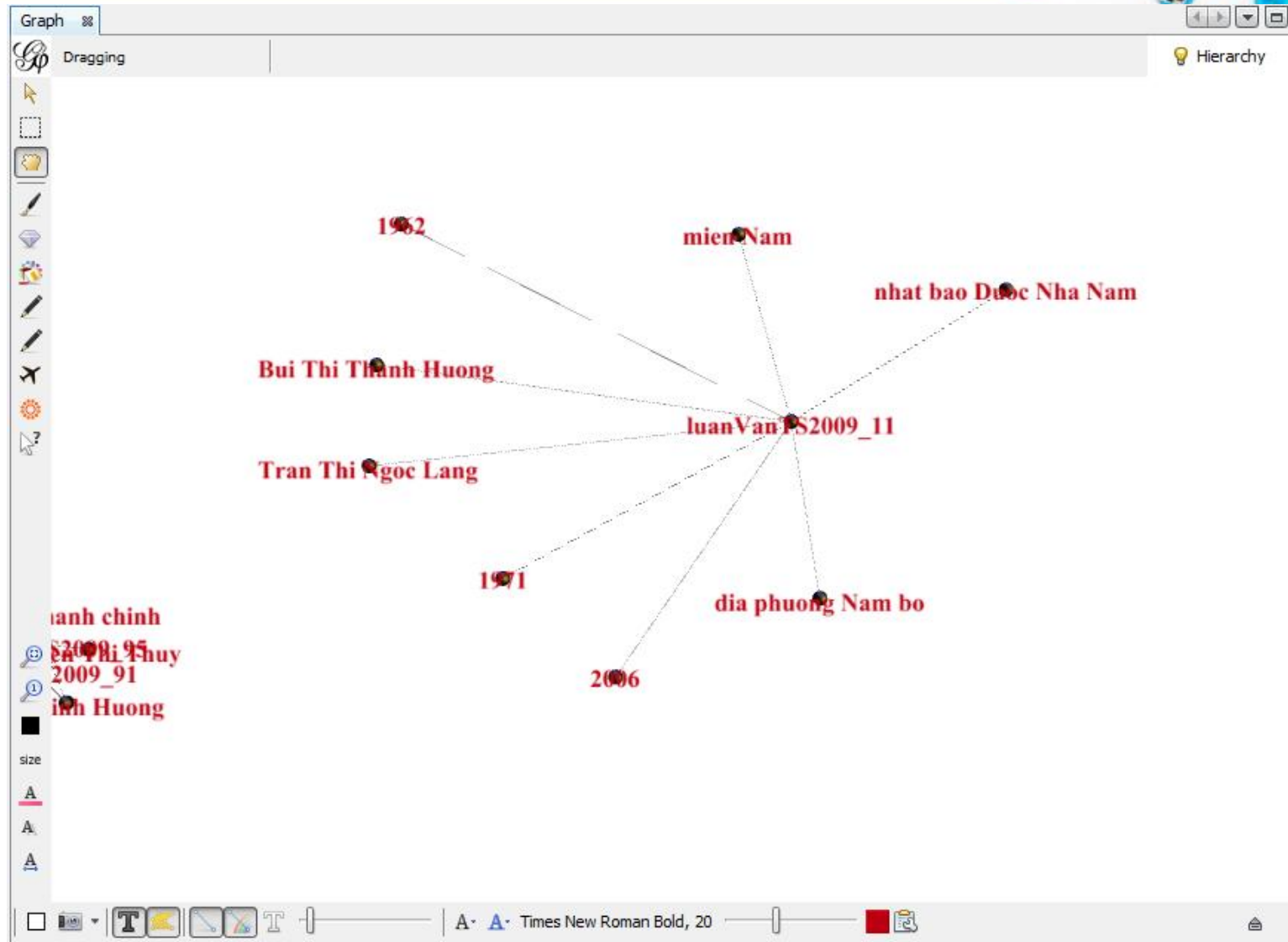


**Restructuration des données**

**Entrepôt de données  
(analyse multidimensionnelle)**

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Source	syntagm												
2	IuanVanTS2009_Titre_01	vong mac dai thao duong				0,4194340737	1,293							
3	IuanVanTS2009_Titre_01	ung dung quang trong toan vong mac					1,20E+13							
4	IuanVanTS2009_Titre_02	Pham tru noi dung												
5	IuanVanTS2009_Titre_03	thanh pho Ho Chi Minh				192,65	63,67							
6	IuanVanTS2009_Titre_03	Quan he bien chuong				0,3040373862	8168							
7	IuanVanTS2009_Titre_05	Bieu hien tu tuong ton giao tin nguoi				0,883156217	294405							
8	IuanVanTS2009_Titre_05	kien truc ton giao Nam bo				0,5661257803	16926							
9	IuanVanTS2009_Titre_06	van ban tin tieng Anh				0,697466963	350453							
10	IuanVanTS2009_Titre_06	tieng Viet				0,2701345931	80383							
11	IuanVanTS2009_Titre_06	van ban tin tieng Anh					1,33E+14							
12	IuanVanTS2009_Titre_07	nguyen vien to phi chinh phu nuoc ngoai					1,73E+14							
13	IuanVanTS2009_Titre_07	thanh pho Ho Chi Minh				0,3567232826	8312							
14	IuanVanTS2009_Titre_08	Phuong phap dich Anh-Viet					1,84E+14							
15	IuanVanTS2009_Titre_10	dong bang ven bien Nam Trung				0,5064426895	7843							
16	IuanVanTS2009_Titre_101	Dan chu Nhan dan Lao				0,307566841	843823							
17	IuanVanTS2009_Titre_101	tin mien Trung nuoc				0,307566841	843823							
18	IuanVanTS2009_Titre_101	tin mien Trung nuoc				0,3229451839	96014							
19	IuanVanTS2009_Titre_102	hanh vi trai phap luat				0,4236948830	63359							
20	IuanVanTS2009_Titre_103	huong hoan thien				0,4281969538	39712							
21	IuanVanTS2009_Titre_105	Tieng Viet				0,4083429896	91277							
22	IuanVanTS2009_Titre_105	To hop cu phap dang lap				0,9391886762	89937							
23	IuanVanTS2009_Titre_106	tieng Viet				0,1416028109	4133							
24	IuanVanTS2009_Titre_106	cau don tieng Anh				0,5650300530	000184							
25	IuanVanTS2009_Titre_106	cau don tieng Anh				0,2991335574	70686							

## II. A.: Identification d'experts





## II. B. : Questionnaire



### **Traitement des questions ouvertes**

Questionnaire en ligne sur les nanosciences

Détermination des compétences des répondants

+

Leurs principales motivations



Comité de pilotage stratégique

### **Répondants :**

- Une centaine de chercheurs des laboratoires en recherche fondamentale (CNRS et universités)  
dans le champs des nanosciences (spintronique, photovoltaïques, optronique, plasmas chauds, etc.

## II. B. : Questionnaire



Fichier CSV



Mappage XSD



### **Préparation des données :**

- Sélection des données
- Nettoyage des données
- Construction de nouvelles données
- Intégration des données
- Formatage des données

## II. B. : Questionnaire



Corpus ClubLNE\_NooJ.noc [Modified]

Corpus Language is "French (fr)";  
Original Text File format is "Default".  
Corpus consists of 71 text files  
XML Text Nodes are: <Mots-cles>  
Corpus contains 213 text units delimited by: <P>

File Name | Size | Last Modif.  
clubLNE01 | 1901 | 01/02/2012  
clubLNE02 | 1658 | 01/02/2012  
clubLNE03 | 2113 | 01/02/2012  
clubLNE04 | 2584 | 01/02/2012  
clubLNE05 | 1856 | 01/02/2012  
clubLNE06 | 2019 | 01/02/2012  
clubLNE07 | 3616 | 01/02/2012  
clubLNE08 | 1880 | 01/02/2012  
clubLNE09 | 1810 | 01/02/2012  
clubLNE10 | 2023 | 01/02/2012  
clubLNE11 | 1561 | 01/02/2012  
clubLNE12 | 2222 | 01/02/2012  
clubLNE13 | 1591 | 01/02/2012  
clubLNE14 | 1734 | 01/02/2012

Remove Add

Locate a pattern in ClubLNE\_NooJ.noc

Pattern is:

- ☐ a string of characters:
- ☐ a PERL regular expression:
- ☒ a NooJ regular expression:

<ACTIVITE>

☐ a NooJ grammar:

☐ Syntactic Analysis

Index

- ☐ Shortest matches
- ☒ Longest matches
- ☐ All matches

Limitation

- ☒ Only: 100 matches
- ☐ All matches
- ☐ 1 example per match

☒ Reset Concordance

clubLNE\_Activate.nog

French /French syntactic grammar.

<WF> <ACTIVITE> <P>

Concordance for Corpus Corpus ClubLNE\_NooJ.noc

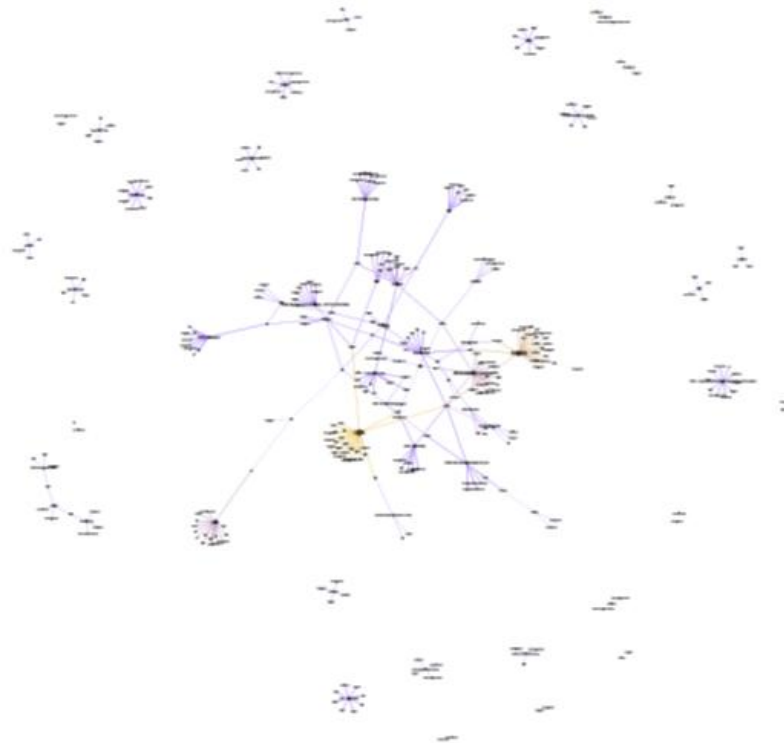
Reset Display: 5 characters word forms before, and 5 after. Display: ☒ Matches ☒ Outputs

Before	Seq.
es minces,	Surface amplificatrice de contraste optique,
ste optique,	Logiciel de traitement et conversion d'
<Mots-cles>	nanoélectronique,
lectronique,	nanophotonique,
photonique,	micro et nanosystèmes,
osystèmes,	nanotechnologies pour la santé,
our la santé,	nanomatériaux,
matériaux,	composants,
omposants,	intégration sur silicium,
sur silicium,	matériaux fonctionnels,
onctionnels,	techniques de dépôts de couches minces.
<Mots-cles>	Optique,
s>Optique,	nanophotonique,
photonique,	champ proche optique,

Query 100/100

Création de dictionnaires

## II. B. : Questionnaire



## II. B. : Questionnaire



Réponse ouverte : « les raisons de votre participation »



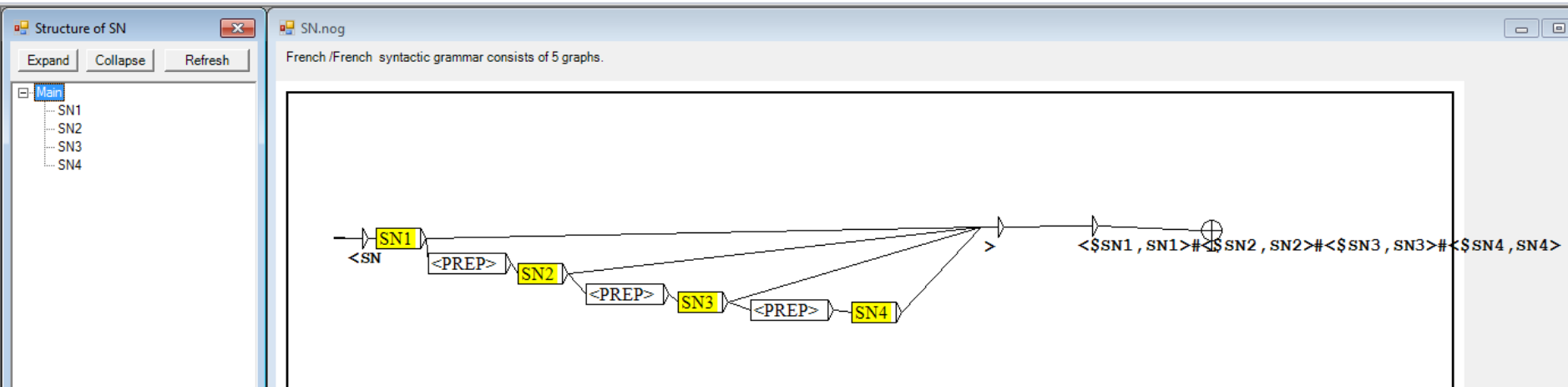
Traitement par NooJ d'un noeud XML spécifique



Extraction des syntagmes nominaux les plus pertinents



## Extraction des SN avec l'outil statistique de NooJ



établir des contacts avec la communauté	0.07811217				
réaliser des caractérisations de couches minces	0.0778794				
DEVELOPPER DE NOUVELLES ACTIVITÉS	0.07638687				
Développement de compétences complémentaires	0.07566451				
une activité de recherche	0.07534191				Débouchés
Structure française de métrologie pour définition	0.07405033				
optimiser la qualité des consortiums pour mes futurs projets	0.0733888				Recherche et technique
autres projets collaboratifs a impact industriel	0.07256933				
Rencontrer des acteurs de ce secteur	0.07195017				Réseau
Avancées vers la métrologie	0.06616665				
pour définition de protocoles	0.06457878				
Faire partie du réseau	0.06431864				
connaître les acteurs français du monde	0.06356252				
REALISER DES ETUDES	0.06222717				
Avoir des contacts	0.0618805				
apporter ensemble les réponses les plus adéquates	0.06041966				
Recherche de nouveaux clients	0.05958986				
mise en place de collaboration	0.05934748				
analyse des nanoparticules	0.0581209				
un réseau industriels	0.05809941				
Etre informé des nouvelles méthodes	0.05720627				
Évaluer les besoins instrumentaux pour mesures	0.05673448				
les autres techniques	0.05672379				
analyse des risques	0.05592926				
développement de mesures	0.05495035				
réfèrent en métrologie dimensionnelle	0.05461787				
établir des collaborations	0.05402814				
Discussion avec divers partenaires	0.05341636				
opérateur dans le traitement des déchets	0.05322692				
enseignement des questions scientifiques	0.05296692				
des projets collaboratifs en cours de montage	0.05237262				
développe une technologie laser	0.0519196				
sur un sujet en plein développement	0.05174325				
utilisation des outils	0.05062894				
les questionnements dans ce domaine	0.05022253				
Acquérir des conseils	0.05015725				



**Merci de votre attention !**