

12^{ème} Colloque international ISKO-France 2019

Université Paul-Valéry Montpellier 3

9-11 octobre 2019

Une classification interdisciplinaire pour l'échange et la médiation des données ouvertes de la recherche

Marcin Trzmielewski – doctorant

Equipe Lerass-Ceric, Université Paul-Valéry Montpellier 3 (France)

Claudio Gnoli – bibliothécaire

Bibliothèque des Sciences et Techniques, Université de Pavie (Italie)



Integrative Levels Classification (ILC) 2004-

Classification
interdisciplinaire à facettes,
reposée sur les phénomènes

Inspirée par le projet du
Classification Research
Group

10 000 concepts en 2019

main classes: (select + to expand)

a	forms	+
b	spacetime	+
c	branes	+
d	energy	+
e	atoms	+
f	molecules	+
g	continuum bodies	+
h	celestial bodies	+
i	rocks	+
j	land	+
k	genes	+
l	bacteria	+
m	organisms	+
n	populations	+
o	instincts	+
p	consciousness	+
q	language	+
r	rituals	+
s	communities	+
t	polities	+
u	enterprises	+
v	technologies	+
w	artifacts	+
x	artworks	+
y	knowledge	+

Facets key

0	as for aspect	+
1	at time	+
2	in place	+
3	by agent	+
4	affected by disorder	+
5	with transformation	+
6	having property	+
7	with part	+
8	as form	+
9	of kind	+

browse special classes:

A	those	+
B	the 1st ones	+
C	the 2nd ones	+
D-T	(etc.)	
U	the typical	+
V	the entirety	+
W	together with	+
X	something	+
Y	the actual	+
Z	the mentioned	+

Intérêt croissant pour les données en SHS

Nombre d'entrepôts (multi)disciplinaires contenant les données en SHS

Augmentation de 96,7% en 3 ans

Entrepôts en 2019

781

NOTRE ÉTUDE

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

Humanities and Social Sciences Search

Entrepôts en 2016

397

GÓMEZ ND, MÉNDEZ E, HERNÁNDEZ-PÉREZ T (2016). Social Sciences and Humanities Research Data and Metadata: A Perspective from Thematic Data Repositories. *El profesional de la información*, 2016, n° 25, 545-555.

0 100 200 300 400 500 600 700 800 900

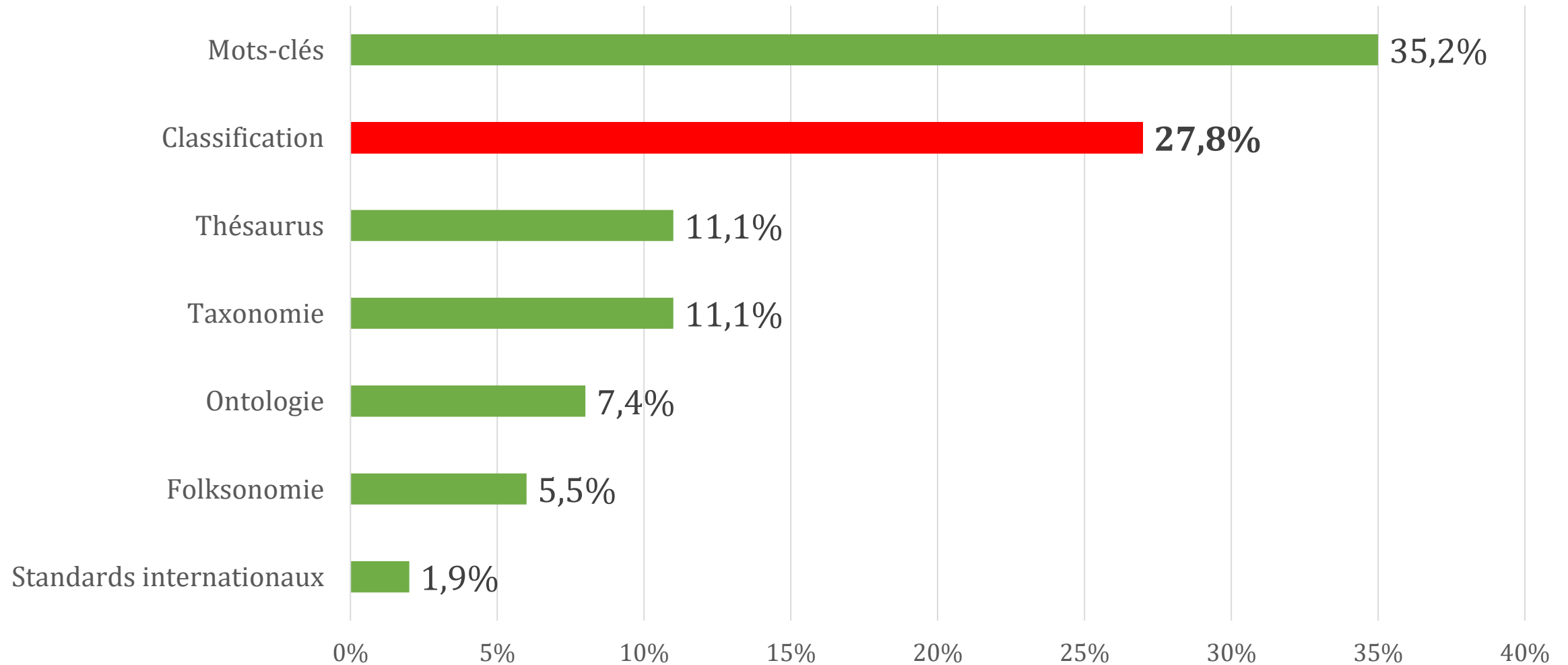
Entrepôts de données en SHS

Indexation/classification sujet pas toujours présente

Données organisées par des mots-clés

Données organisées avec des classifications
disciplinaires

Systemes d'organisation des connaissances (SOC) integres dans les entrepôts contenant des donnees ouvertes en SHS (N=22 ; *in-house data*)



300 000 €

150 000 €

100 000 €

72 000 €

48 000 €

24 000 €

12 000 €

6 000 €

3 000 €

1 500 €

800 €

500 €

300 €

➤ 200 €

ILC pourrait-elle être envisagée pour échanger et médiatiser les données de la recherche en SHS ?

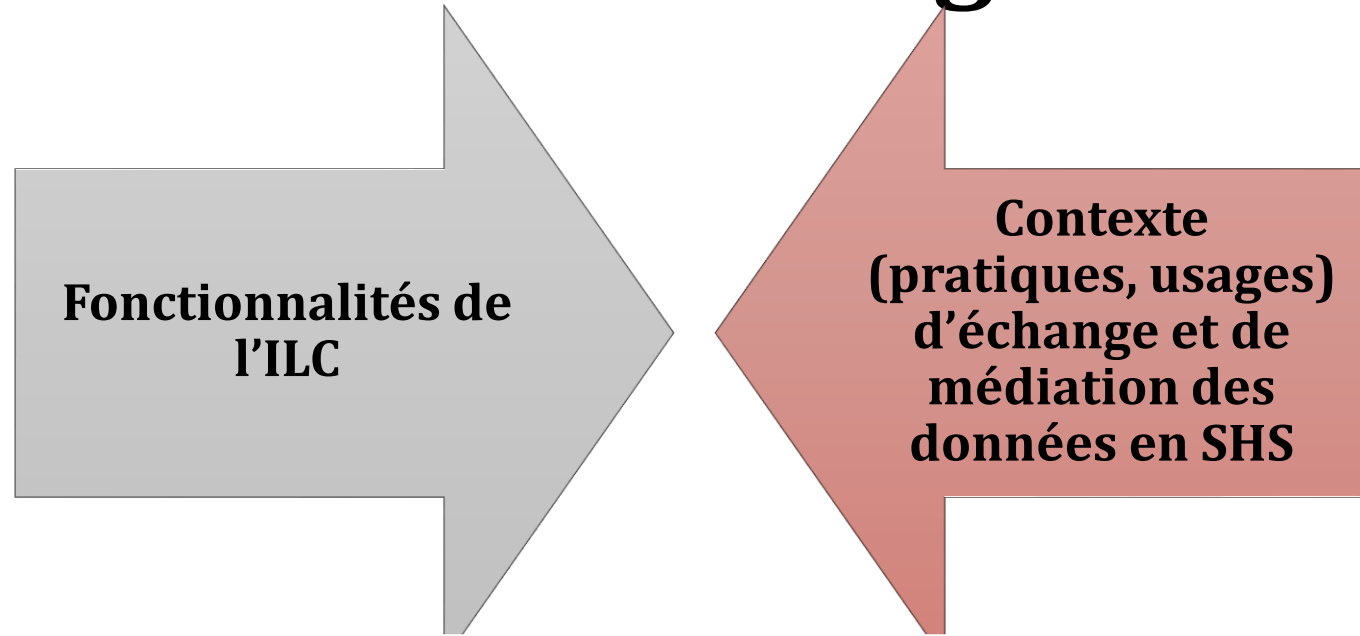
♦A: Oui

♦B: Non

♦C: Peut-être

♦D: Demandez à quelqu'un d'autre

Méthodologie



Analyse des travaux en sciences de l'information, en organisation des connaissances et en philosophie des sciences

Interrogation du registre d'entrepôts



Simulation d'organisation des données

Bénéfices

Chercheurs

- ✓ Avancement de la carrière
- ✓ Moyen de trouver le matériel scientifique
- ✓ Création de nouvelles collaborations

Recherche et innovation

- ✓ Augmentation de la reproductibilité de la science
- ✓ Restriction des coûts liés à la génération des données
- ✓ Encouragement des collaborations pour affronter de grands problèmes sociétaux

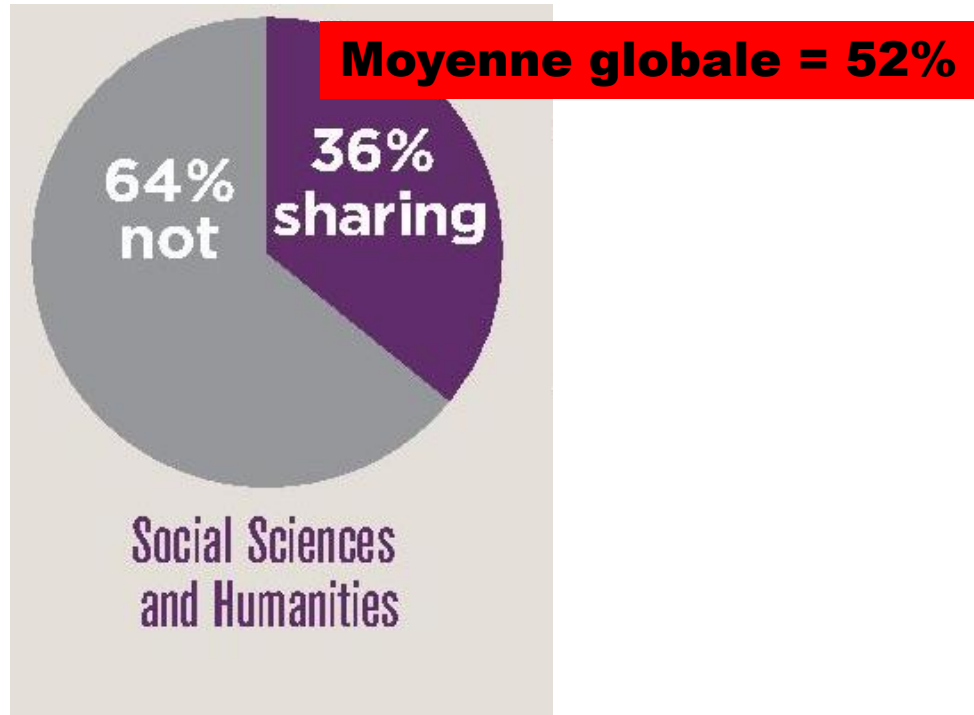
DÉLÉMONTEZ R, BOUKACEM-ZEGHMOURI C (2015). Données de la recherche : entre discours, réalités et valeur. *I2D – Information, Données & Documents*, 2015, n° 53, 56-57.

LYON L (2016). Transparency: the emerging third dimension of open science and open data. *Liber quarterly*, 2016, n° 25, 153-171.

DILLAERTS H (2014). Le libre accès et le financement de projets de recherche transversaux : des vecteurs d'interdisciplinarité dans l'économie de la connaissance ?. *Revue européenne des sciences sociales*, 2014, no° 52, 111-135.

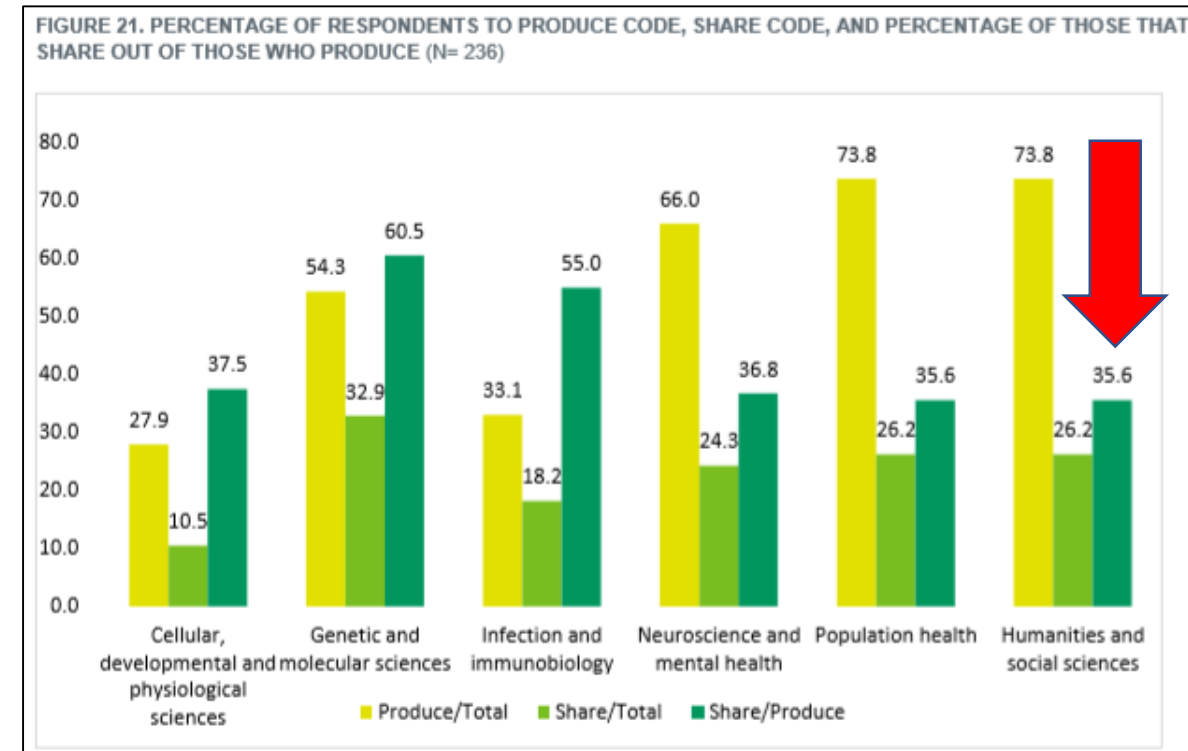
Usages

2014



FERGUSON L (2014). How and Why Researchers Share Data (And Why They don't). Wiley, 3 November 2014

2016



VAN DEN EYNDEN V (2016). *Towards Open Research: practices, experiences, barriers and opportunities*. London : Wellcome Trust, 2016.



Données échangées

DFG Classification of Scientific Disciplines, Research Areas, Review Boards and Subject Areas (2016-2019)

Subject Area	Review Board	Research Area	Scientific Discipline
101-01 Prehistory	101 Ancient Cultures		
101-02 Classical Philology			
101-03 Ancient History			
101-04 Classical Archaeology			
101-05 Egyptology and Ancient Near Eastern Studies			
102-01 Medieval History	102 Medieval Studies		
102-02 Early Modern History			
102-03 Modern History			
102-04 History of Art			
102-05 Cultural Studies			
100-02 Theoretical Philosophy	Philosophy		
108-03 Practical Philosophy			
109-01 General Education and History of Education	109 Educational Research		
109-02 General and Domain-Specific Teaching and Learning			
109-03 Education Systems and Educational Institutions			
109-04 Educational Research on Socialization, Welfare and Organisations			
110-01 General, Biological and Mathematical Psychology	110 Psychology		
110-02 Developmental and Educational Psychology			
110-03 Social Psychology, Industrial and Organisational Psychology			
110-04 Differential Psychology, Clinical Psychology, Medical Psychology, Methodology			
111-01 Sociological Theory	111 Social Sciences	12 Social and Behavioural Sciences	
111-02 Empirical Social Research			
111-03 Communication Sciences			
111-04 Political Science			
112-01 Economic Theory	112 Economics		
112-02 Economic Policy and Public Finance			
112-03 Business Administration			
112-04 Statistics and Econometrics			
112-05 Economic and Social History			
113-01 Principles of Law and Jurisprudence	113 Jurisprudence		
113-02 Private Law			
113-03 Public Law			
113-04 Criminal Law and Law of Criminal Procedure			
113-05 Criminology			

**52 champs en SHS répertoriés dans
la Classification of Scientific
Disciplines
=
52 communautés de pratiques**



Content Types

- Archived data (191)
- Audiovisual data (266)
- Configuration data (24)
- Databases (168)
- Images (349)
- Networkbased data (42)
- Plain text (386)
- Raw data (293)
- Scientific and statistical data formats (409)
- Software applications (119)
- Source code (63)
- Standard office documents (615)
- Structured graphics (191)
- Structured text (244)
- other (288)

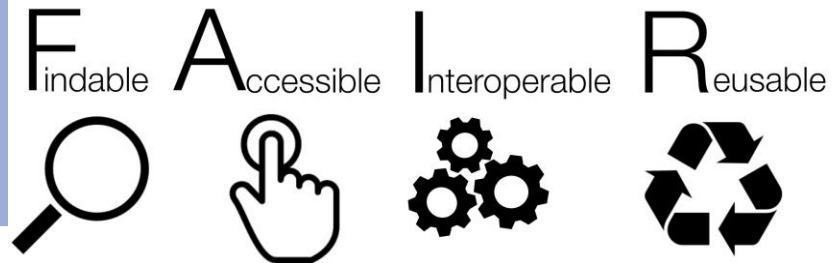
Pratiques de médiation

Partager librement et réutiliser les données pour créer une



open science

Rendre les données de recherche



Sungya Pundir, Wikimedia Commons, CC By-SA 4.0

COMMISSION EUROPEENNE, DIRECTION GENERALE DE LA RECHERCHE ET DE L'INNOVATION (2016).
Lignes directrices pour la gestion des données FAIR dans Horizon 2020 Version 3.0 26 Juillet 2016.



Pratiques de médiation

Recommandations FAIR pour l'organisation des connaissances :

Choisir des SOC validés par les communautés scientifiques

→ *scientific warrant* (Bliss, 1929)

Représenter les SOC en formats du Web sémantique pour décrire et lier les données entre elles d'une façon interdisciplinaire

→ *décontextualisation* (Leonelli, 2019)

Fournir une riche description des données via les métadonnées

→ *recontextualisation* (Leonelli, 2019)

3 défis pour l'ILC



Contexte en SHS	Défi pour l'ILC
➤ Pluralité de champs disciplinaires	(1) Décrire un vaste spectre des connaissances
➤ Des données hétérogènes ➤ Recontextualisation des données pour les bien interpréter	(2) Décrire les données hétérogènes d'une manière riche et nuancée
➤ Taux bas d'usages ➤ Besoin de décontextualiser et lier les données entre elles	(3) Valoriser et faire découvrir l'information au-delà des frontières disciplinaires

(1) Décrire un vaste spectre des connaissances

Systeme général ayant pour vocation de classer tout le savoir humain

Systeme basée sur les *phénomènes*, exprimés par les caractères alphabétiques

Possibilité d'exprimer des relations entre n'importe quelle paire de classes

main classes: (select + to expand)

a	forms	+
b	spacetime	+
c	branes	+
d	energy	+
e	atoms	+
f	molecules	+
g	continuum bodies	+
h	celestial bodies	+
i	rocks	+
j	land	+
k	genes	+
l	bacteria	+
m	organisms	+
n	populations	+
o	instincts	+
p	consciousness	+
q	language	+
r	rituals	+
s	communities	+
t	polities	+
u	enterprises	+
v	technologies	+
w	artifacts	+
x	artworks	+
y	knowledge	+

Théorie des niveaux d'intégration : lorsque les éléments d'un niveau inférieur sont combinés, ils peuvent donner lieu à une nouvelle intégration.

NEEDHAM J (1937). *Integrative Levels: A Revaluation of the Idea of Progress*. Oxford : Clarendon Press, 1937, 59.

FEIBLEMAN JK (1954). Theory of Integrative Levels. *The British Journal for the Philosophy of Science*, 1954, no° 5, 59-66.

q	langue
qv	langues
qve	langues indo-européennes
qves	langues slaves

x	art
xl	littérature
xlb	proverbes

Relations entre les niveaux :

- **Dépendance** – critères phylogénétiques
- **Inclusion** – augmentation de spécificité morphologique

GNOLI C (2017). Classifying phenomena, part 2: Types and Levels. *Knowledge organization*, 2017, n° 44, 37-54.

(2) Décrire les données hétérogènes d'une manière riche et nuancée

Approche synthétique : la possibilité de combiner infiniment les phénomènes, relations et propriétés.

SZOSTAK R (2011). Complex concepts into basic concepts. *Journal of the American Society for Information Society and Technology*, 2011, n° 62, 2247-65.

	<i>Facets key</i>	
0	as for aspect	+
1	at time	+
2	in place	+
3	by agent	+
4	affected by disorder	+
5	with transformation	+
6	having property	+
7	with part	+
8	as form	+
9	of kind	+

Une combinaison entre des phénomènes est possible grâce à une **syntaxe à facettes**, permettant d'établir des relations libres.

GNOLI C (2017). Classifying phenomena, part 2: Types and Levels. *Knowledge organization*, 2017, n° 44, 37-54.

Description du sujet de données

x1b73qves « proverbes, en, langues slaves »

<i>x1b</i> « proverbes »	73 « en »	<i>qves</i> « langues slaves »
phénomène	facette	phénomène

Description du sujet de données avec leurs propriétaires documentaires

r2ttbzw00ytm « traditions, en Pologne, enregistré[es] en vidéo »

<i>r2ttbzw</i> « traditions en Pologne »	00 « enregistré en »	<i>ytm</i> « video »
phénomène	facette dimensionnelle	perspective

Facettes (filtres)



Sujet

France (1)
Pologne (1)
Russie (1)
traditions (2)

Méthode d'étude

enquête (1)
tournage vidéo (1)

traditions

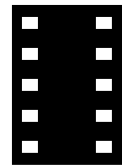


Traditions en Occitanie

Auteur : Jean Bon

Sujet : **traditions** en France étudié par enquête

Méthode d'étude : enquête



La Russie et son influence sur les traditions du peuple polonais

Auteur : Haute Ecole d'Etude Sociales à Varsovie

Sujet : Russie, **traditions** en Pologne,
enregistré en vidéo

Méthode d'étude : vidéo

(3) Valoriser et faire découvrir l'information au-delà des frontières disciplinaires

Phénomènes sont les termes pour lesquels il existe le plus grand accord sur leur sens entre groupes, individus et disciplines.

Szostak R (2014). The Importance of Knowledge Organization, *Bulletin of the Association for Information Science and Technology*, 2014, n° 40, 37-42.

proverbes

ordinateur

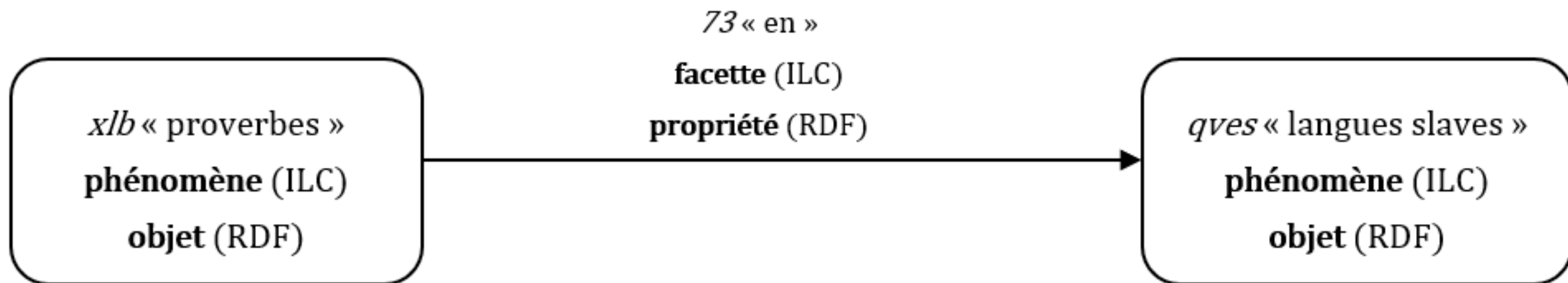
Internet

traditions

Pologne

Phénomènes de l'ILC peuvent donc devenir des unités interdisciplinaires permettant de lier les données entre elles !

Grâce à la combinaison des phénomènes, nous pourrions coder l'information sur les données en syntaxe RDF :



... ce qui permettrait de lier les données décrites via l'ILC avec d'autres données sur le Web.

**Publication
imminente de
l'ILC en SKOS
(Ntriples RDF)**

**Etablissement de passerelles
terminologiques avec d'autres
SOC**

Collaboration avec :

University of
South Wales
Prifysgol
De Cymru

Hypermedia Research Group

CDD	ILC	Libellé
<i>170</i>	<i>p938</i>	valeurs
<i>180</i>	<i>yy99</i>	écoles philosophiques
<i>190, 140</i>	<i>yy99m</i>	philosophie moderne occidentale

Résultats et discussion

Défi	Commentaire	Fonctionnalités de l'ILC	Approche
Décrire des connaissances vastes	✓ Description possible	✓ Représentation des concepts par les phénomènes ✓ Possibilité d'exprimer des relations entre n'importe quelle paire de classes ✓ Libre combinaison de phénomènes et facettes	Ontologique Synthétique
Décrire les données hétérogènes d'une manière riche et nuancée	✓ Description possible		
Valoriser et faire découvrir l'information au-delà des frontières disciplinaires	✓ Connecteurs interdisciplinaires peuvent être générées		

Résultats et discussion

Pas de retour d'expérience de la part des utilisateurs des bases de connaissance organisées via l'ILC



the BioAcoustic Reference Database

by CIBRA: Centro interdisciplinare di Bioacustica e ricerche ambientali
at the University of Pavia, in collaboration with [ISKO Italia](#)

Freely faceted search interface



choose facets: [only part of the records are classified!]

taxon: (anyone)

insects other invertebrates

ray-finned fish amphibians reptiles birds

mammals

cetaceans marine mammals

rodents bats elephants

primates humans

function: (anyone)

alarm threat localization

courtship individual recognition social contact

signals: (anyone) structure analyzed

vibrational infrasounds ultrasounds

related topics: (anyone)

behaviour echolocation

population abundance and distribution

risk mitigation conservation human disturbance

pollution sound, noise fishing ships tourism

Where the Apennine begins

Books and articles

expand a class:

- A Polcevera valley
- B Bisagno valley
- C Scrivia valley
- D Spinti valley
- E Borbera valley
- F Ossona valley
- G Grue valley
- H Curone valley
- J Staffora valley and Oltrepò
- K Tidone valley
- O Trebbia valley
- S Aveto valley
- T Fontanabuona valley
-
- h rocks
- j territory
- m organisms
- n environment
- p personalities
- r language
- s society, community
- t institutions
- u economy
- v technologies
- w artifacts
- x arts, leisure
- y knowledge

or search the classes corresponding to the subject

or search directly by name of an author

or search directly by words in the title

(up to 5 words separated by blanks)

*The bibliographic database currently consists of 823 records
The scheme is arranged according to
the [integrative level classification](#)*

<http://www.iskoi.org/ilc/bard/>

<http://www.iskoi.org/ilc/4p/index.en.php>

Résultats et discussion



● Evaluer les usages
de l'ILC

● Organiser un corpus
de données en SHS

Conclusion

Jokers

1/15

1 MILLION €

300 000 €

150 000 €

100 000 €

72 000 €

48 000 €

24 000 €

12 000 €

6 000 €

3 000 €

1 500 €

800 €

500 €

300 €

➤ 200 €


ILC pourrait-elle être envisagée pour échanger et médiatiser les données de la recherche en SHS ?

➤ **A:** Oui

➤ **B:** Non

➤ **C:** Peut-être

➤ **D:** Demandez à quelqu'un d'autre



Contact



marcin.trzmielewski@etu.univ-montp3.fr



[@MTrzmielewski](https://twitter.com/MTrzmielewski)



claudio.gnoli@unipv.it



[@scritur](https://twitter.com/scritur)

