



El *Thesaurus de la UB* en Linked Open Data (Linked Open Data/Dades obertes vinculades)

Adaptació de: *El Thesaurus de la Universitat de Barcelona: 25 anys en constant evolució*
per Rosa Fabeiro (UB. CRAI Procés Tècnic) i Pep Casals (Nubilum)

Carme Masagué (CRAI Unitat de Procés Tècnic)



Què és el *Thesaurus de la Universitat de Barcelona*?

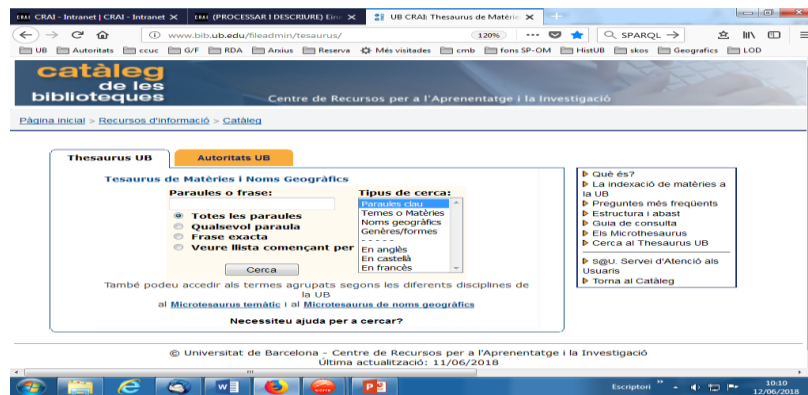
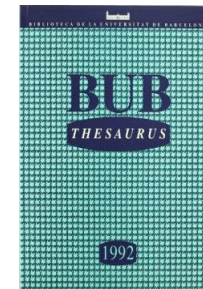
És una eina que facilita la consulta per matèries:

- **normalitza** la forma d'accedir a les matèries
- **identifica** sense ambigüitat el terme buscat
- **agrupa** els documents del catàleg a partir de les formes acceptades
 - és un **vocabulari controlat i estructurat** amb relacions d'equivalència, jeràrquiques i associatives
 - s'utilitza per a la **indexació i recuperació per matèries** de tots els recursos d'informació del CRAI-UB
 - és **multilingüe**, redactat en **català** amb **equivalències en castellà, anglès i francès**
 - és **multidisciplinari**, incorpora **classificació per categories** que són la base dels **Microthesaurus UB**.
 - desenvolupat per la **Unitat de Procés Tècnic del CRAI de la UB**.

Cronologia del *Thesaurus de la UB* (THUB) (I)

1992: Publicació del *Thesaurus de la UB* en paper, 1a ed.
(8.500 termes)

1998: Publicació del *Thesaurus de la UB* en línia



Cronologia del *Thesaurus de la UB* (THUB) (II)

2005: els termes del Thesaurus es classifiquen en 29 Microthesaurus temàtics.

Des del 2006 : Ús del *THUB* com a únic vocabulari controlat per a la indexació dels documents, segons l'esquema de metadades Dublin Core, en tots els repositoris digitals propis i consorciats.



2007: s'incorpora el Microthesaurus de Noms geogràfics.

2009: s'incorpora al catàleg la cerca per gènere i forma.

Cronologia del *Thesaurus de la UB* (THUB) (III)

2014: Iniciativa de publicar el Thesaurus UB en dades obertes vinculades en format SKOS amb el programari de software lliure *TemaTres*.

El projecte es va aturar per manca de termes equivalents del thesaurus multilingüe i per no permetre la gestió des del nostre Sistema Integrat de Gestió de Biblioteca (SIGB) de la UB

- 26.500 descriptors amb 58 % d'equivalències multilingües

2018: (maig) Amb la col·laboració de l'empresa **Nubilum** i amb tecnologia **Coeli** hem transformat les dades del THUB al format SKOS i s'ha iniciat el període de proves per a la publicació del thesaurus en dades obertes vinculades (LOD) i en versió multilingüe. La gestió es farà dins del nostre SIGB.

- 27.506 descriptors i 80% de termes amb equivalències multilingües

El Web semàntic



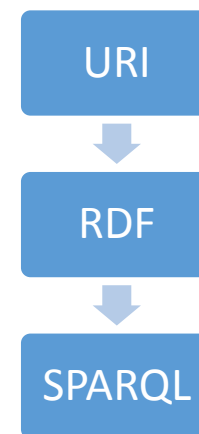
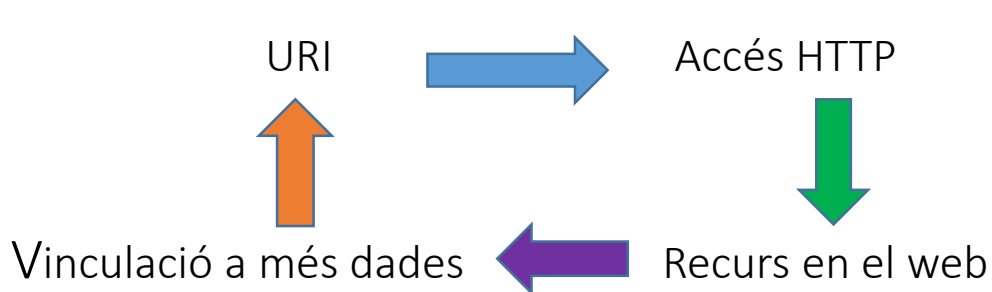
Web semàntic o Web de les dades

- Dades llegibles per màquina: metadades en format RDF, XML, JSON
- Noves tecnologies que permeten obrir les dades, vincular-les i enriquir-les
- Llicència aplicada a les dades

Què són les dades obertes vinculades?

(també anomenades LOD, o dades obertes enllaçades)

- Vincular dades és essencial per connectar el web semàntic.
- Linked Open Data és el projecte que treballa per la evolució del web basat en documents enllaçats cap al web de dades vinculades.
- Publicació de bases de dades en RDF establint vincles RDF entre dades de fonts diferents.
- Les dades obertes poden ser utilitzades, reutilitzades i redistribuïdes lliurement per qualsevol persona i estan disponibles sota una llicència oberta.



* Europeana: [Datos abiertos vinculados. ¿Qué son?](#)

Iniciativa CRAI-UB

El CRAI de la UB amb la col·laboració de l'empresa Nubilum i amb la seva tecnologia Coeli s'ha sumat al projecte de dades vinculades en accés obert amb la publicació del Thesaurus de la UB en LOD .

The screenshot shows a web browser window displaying the CRAI-UB website. The address bar shows the URL crai.ub.edu/ca/recursos-d-informacio/thub. The website header includes the Universitat de Barcelona logo and the CRAI logo (Centre de Recursos per a l'Aprenentatge i la Investigació). Navigation links include 'Coneix el CRAI', 'Recursos d'informació', and 'Serveis que ofereix el CRAI'. A search bar is visible with the text 'Thesaurus de la Universitat de Barcelona (THUB)'. Below the search bar is a large empty input field and a 'CERCA' button. The footer contains copyright information and social media icons.

Thesaurus de la UB com a recurs del CRAI-UB

Possible interfície de cerca
(en construcció)

Com el difonem?

Funcionalitats i possibilitats:

- Integrat com **un recurs més del CRAI**.
- Experiència d'usuari
 - **Estructura multilingüe completa**, no només a nivell de traducció del terme sinó a nivell global;
 - **Cerca predictiva**, que ofereix als usuaris els termes que s'aproximen a la seva construcció de cerca segons els termes usats al thesaurus.
 - Facetes a partir de l'estructura dels microthesaurus.
 - Navegació gràfica a través de constel·lacions de nodes.
- **Interoperabilitat** a través de diferents EndPoints:
 - **API Rest** amb sortida JSON
 - **Linked Open Data** en RDF (SPARQL)
- Vinculació a múltiples eines de descoberta **no només Cercabib**.
- **Simplificar i optimitzar** els processos d'edició i revisió.

UNIVERSITAT DE BARCELONA

Inici CRAI Inici UB Cerca Català English Español

Intranet PAS CRAI SIGU Directori

Biblioteques i horaris Preguntes més freqüents El CRAI en línia

Conèix el CRAI Recursos d'informació Servels que ofereix el CRAI

Inici >> Recursos d'informació >> Thesaurus UB (proves)

Thesaurus UB (proves)

Més informació

CERCA +

< TORNA AL LLISTAT

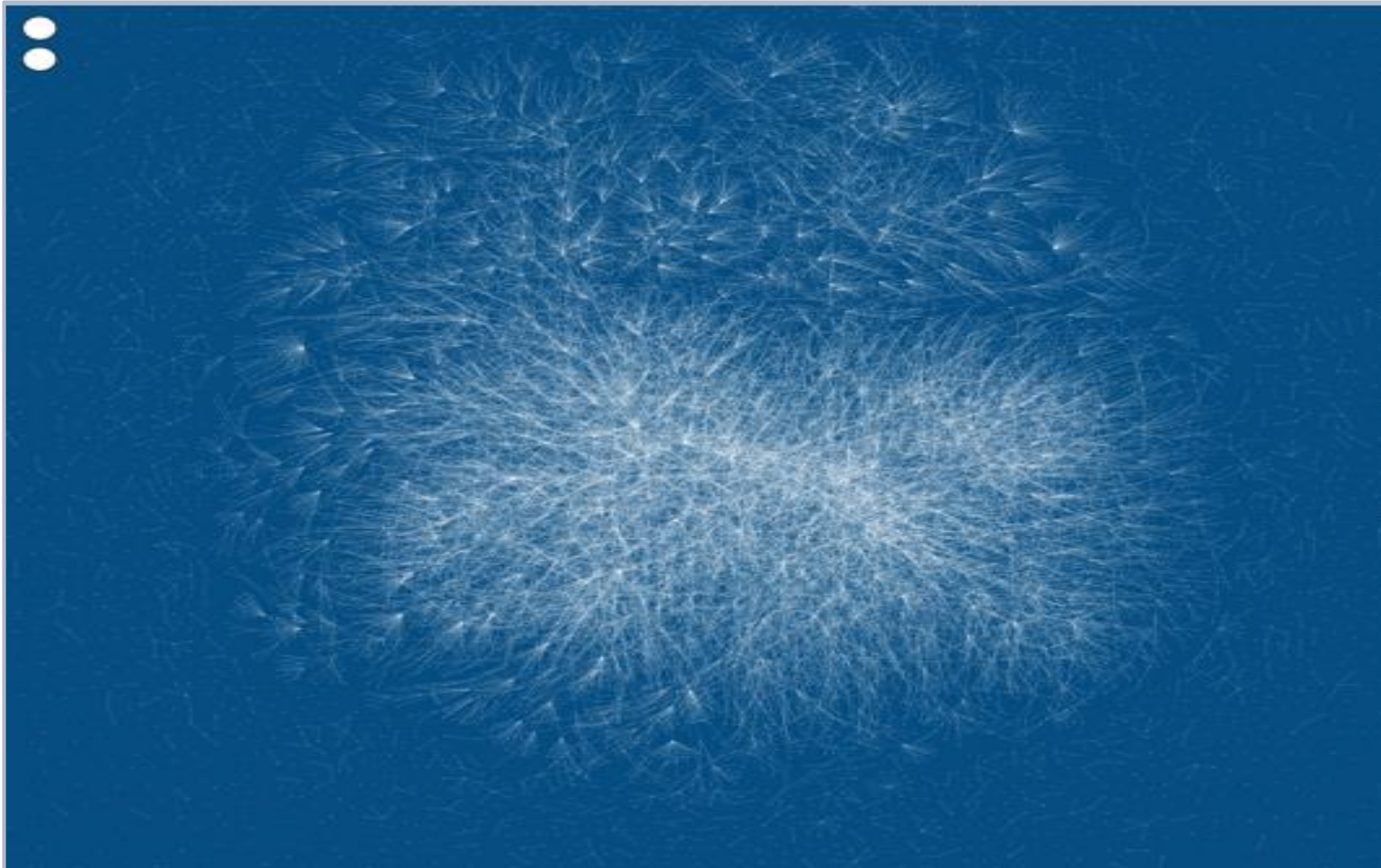
Bacteris nitrificants

DESAR COM: RDF/XML JSON-LD MARC21

TESAURUS	Matèries
MICROTESAURUS	Ciències i tecnologia agroalimentàries, agricultura, silvicultura i zootècnia
TERME PREFERENT	Bacteris nitrificants
TERMES NO PREFERENT	Nitrobacteriàcies
TERMES GENÈRICS	Bacteriologia agrícola · Bacteris · Nitrificació
IDENTIFICADOR	.a1176885x
CERCA DOCUMENTS	CERCABIB

Graph showing relationships between Bacteris, Bacteriologia agrícola, Nitrificació, and Bacteris nitrificants.

Representació gràfica del Thesaurus de la UB



Què ens queda pendent?

- Completar les equivalències en castellà, anglès i francès.
Actualment només el 80% de termes estan traduïts.
- Completar els nodes que manquen en l'estructura jeràrquica
Queda un 12,6% de termes que encara no estan associats a un descriptor genèric.
- Ampliar els horitzons:
 - Les institucions interessades en el thesaurus començaran a usar-lo?
 - Els usuaris interns perceben la millora?...



Més informació:

Fabeiro, Rosa; Casals, Pep. [*El Thesaurus de la Universitat de Barcelona: 25 anys en constant evolució*](#). (Experiència presentada a les 15es Jornades Catalanes d'Informació i Documentació, maig 2018)

Iniciatives de dades obertes vinculades en altres Biblioteques

[Datos enlazados en la BNE](#)

[Dades vinculades en la Bibliothèque nationale de France](#)

[Dades vinculades en la British Library](#)

[Dades vinculades en la Deutsche Nationalbibliothek](#)

[Dades vinculades de la Library of Congress \(Linked Data Service\)](#)

[Dades vinculades dels Getty Vocabularies \(Linked Open Data\)](#)



Moltes gràcies!

