

Base de Datos para a análise da Producción Científica Galega: Necesidades

Cláusula de confidencialidade:

Este documento é propiedade do Consorcio de Bibliotecas Universitarias de Galicia. Tódolos dereitos reservados. Non está permitida a reprodución total ou parcial destes documentos, nin o seu tratamento informático, nin a transmisión de ningunha forma ou por calquera medio, sen o permiso previo e por escrito do Consorcio de Bibliotecas Universitarias de Galicia. Queda expresa e terminantemente prohibida a súa comercialización.

Para o desenvolvemento dunha base de datos que almacene a información sobre a produción científica galega é necesario levar a cabo os seguintes procesos:

- Configuración do servidor SQL Server 2005: Configuración de políticas de seguridade segundo o nivel de acceso permitido aos distintos usuarios:
 - Administrador
 - Perfil de carga de datos
 - Perfil de normalización
 - Creador de informes
 - Distintos niveles de consultor de informes
 - ...
- Estudo dos formatos e modos de extracción dos datos da produción científica de distintas fontes de datos:
 - ISI Web of Science: <http://www.accesowok.fecyt.es/wos/>
 - Scopus: <http://www.scopus.com/>
 - BDs do "Instituto de Estudios Documentales sobre Ciencia y Tecnología", IEDCYT, (do "Consejo Superior de Investigaciones Científicas", CSIC): Consta de tres bases de datos:
 - ICYT: Ciencia e Tecnoloxía.
 - ISOC: Ciencias Sociais e Humanas.
 - IME: Biomedicina.
 - Outras fontes de datos

Un mesmo documento pode estar referenciado en varias bases de datos, o cal debe terse en conta xa que é necesario identificar a que base de datos pertence cada información.

Polo de agora existen dúas APIs dispoñibles: a de ISI e a de SCOPUS pero non se descarta a posibilidade de que poidan aparecer novas APIs para outras fontes ou que as xa existentes sufran modificacións, o cal debe terse en conta á hora da implementación.

Resumo dos formatos segundo a fonte e o modo de extracción:

- ISI Web of Science:
 - A través da plataforma:
Un exemplo sería:

```
FN ISI Export Format
VR 1.0
PT S
AU Daniels, GM
   Amara, SG
AF Daniels, Gwynn M.
   Amara, Susan G.
TI Selective labeling of neurotransmitter transporters at the cell surface
SO NEUROTRANSMITTER TRANSPORTERS
SE METHODS IN ENZYMOLOGY
LA English
DT Review
ID POLARIZED EPITHELIAL-CELLS; EXPRESSION
C1 Oregon Hlth & Sci Univ, Vollum Inst Adv Biomed Res, Dept Cell & Dev Biol,
Portland, OR 97201 USA.
   Oregon Hlth & Sci Univ, Howard Hughes Med Inst, Portland, OR 97201 USA.
```

RP Daniels, GM, Oregon Hlth & Sci Univ, Vollum Inst Adv Biomed Res, Dept
Cell & Dev Biol, Portland, OR 97201 USA.
CR AHN J, 1996, J BIOL CHEM, V271, P6917
GOTTARDI CJ, 1993, SCIENCE, V260, P552
GOTTARDI CJ, 1995, AM J PHYSIOL, V268, P285
GU HH, 1996, J BIOL CHEM, V271, P18100
MATTER K, 1994, CURR OPIN CELL BIOL, V6, P545
NGUYEN TT, 1996, J NEUROCHEM, V67, P645
PIETRINI G, 1994, J BIOL CHEM, V269, P4668
POVLOCK SL, 1998, METHODS ENZYMOL, V296
SARGIACOMO M, 1989, J MEMBRANE BIOL, V107, P277
NR 9
TC 29
PU ELSEVIER ACADEMIC PRESS INC
PI SAN DIEGO
PA 525 B STREET, SUITE 1900, SAN DIEGO, CA 92101-4495 USA
SN 0076-6879
J9 METH ENZYMOLOGY
JI Methods Enzymol.
PY 1998
VL 296
BP 307
EP 318
PG 12
SC Biochemical Research Methods; Biochemistry & Molecular Biology
GA BGJ33
UT ISI:000247400300021
ER

PT J
.....
ER

EF

A explicación de que indica cada campo sería a seguinte:

Etiqueta	Descrición
FN	Tipo de ficheiro.
VR	Número de versión do formato de ficheiro.
PT	Tipo de publicación (p.e.: libro, diario, serie de libro).
AU	Autor/es. Cada autor aparece nunha liña.
TI	Título do artigo.
SO	Título completo da fonte.
LA	Idioma.
DT	Tipo de documento (p.e.: Book Review, Editorial Material, Fiction, Note, Review)
NR	Número de referencias citadas.
SN	ISSN
PU	Editor.
C1	Dirección da institución ou institucións investigadoras que

Etiqueta	Descripción
	participaron na elaboración do documento. De existir máis de unha aparecerán en liñas distintas. Se varios departamentos dunha institución participan no documento, aparecerán en liñas distintas, como institucións diferentes.
DE	Palabras clave do autor.
ID	Palabras clave de ISI.
AB	Resumo.
CR	Referencias citadas. Cada referencia citada aparece nunha liña, co seguinte formato: Autor, Ano, Título abreviado da publicación, Volume, Páxina. Por exemplo: BATTEN M, 2006, NAT IMMUNOL, V7, P929 No caso dalgúns xornais electrónicos emprégase o DOI (Digital Object Identifier) no canto dos números de páxina e números de artigo. Se non hai información de número de páxina ou do número de artigo citado aparecera o DOI.
TC	Número de citas; número de veces que o documento foi citado.
BP	Páxina inicial.
EP	Páxina final.
PG	Número de páxinas.
JI	Abreviatura en formato ISO do título da fonte.
SE	Título da serie de libros.
BS	Subtítulo da serie de libros.
PY	Ano de publicación
PD	Data de publicación.
VL	Volume.
IS	Publicación.
PN	Número de parte.
SU	Suplemento.
SI	Publicación especial.
GA	Número ISI de entrega de documento.
PI	Cidade do editor.
WP	Dirección web do editor.
RP	Dirección do autor.
CP	Patente citada.
J9	Abreviatura do título da fonte (ata 29 caracteres).
PA	Dirección do editor.
UT	Identificador único de ISI para cada artigo.
ER	Fin do rexistro.

- A través da API: neste caso o ficheiro ten formato XML. Un exemplo de ficheiro XML sería o seguinte.

```
<RECORDS>
<REC inst_id="2" recid="155636269" hot="yes" sortkey="3117154304"
timescited="3" sharedrefs="0" inpi="false">

<ut>000202990700007</ut>
<source_title>INTERNATIONAL MICROBIOLOGY</source_title>
<source_abbrev>INT MICROBIOL</source_abbrev>
<item_title>Locus NMB0035 codes for a 47-kDa surface-accessible conserved
antigen in Neisseria</item_title>
<bib_id>9 (4): 273-280 DEC 2006</bib_id>
<editions full="SCI"/>
<languages count="1">
<primarylang code="EN">English</primarylang>
</languages>
<authors count="6">
<primaryauthor>Arenas, J</primaryauthor>
<fullauthorname>
<AuRole>Author</AuRole>
<AuLastName>Arenas</AuLastName>
<AuFirstName>Jesus</AuFirstName>
<AuCollectiveName>Arenas, Jesus</AuCollectiveName>
</fullauthorname>
<author key="12554">Abel, A</author>
<fullauthorname>
<AuRole>Author</AuRole>
<AuLastName>Abel</AuLastName>
<AuFirstName>Ana</AuFirstName>
<AuCollectiveName>Abel, Ana</AuCollectiveName>
</fullauthorname>
<author key="6098088">Sanchez, S</author>
<fullauthorname>
<AuRole>Author</AuRole>
<AuLastName>Sanchez</AuLastName>
<AuFirstName>Sandra</AuFirstName>
<AuCollectiveName>Sanchez, Sandra</AuCollectiveName>
</fullauthorname>
<author key="99232">Alcala, B</author>
<fullauthorname>
<AuRole>Author</AuRole>
<AuLastName>Alcala</AuLastName>
<AuFirstName>Belen</AuFirstName>
<AuCollectiveName>Alcala, Belen</AuCollectiveName>
</fullauthorname>
<author key="1459236">Criado, MT</author>
<fullauthorname>
<AuRole>Author</AuRole>
<AuLastName>Criado</AuLastName>
<AuFirstName>Maria T.</AuFirstName>
<AuCollectiveName>Criado, Maria T.</AuCollectiveName>
</fullauthorname>
<author key="2111067">Ferreiros, CM</author>
<fullauthorname>
<AuRole>Author, Reprint Author</AuRole>
```

```
<AuLastName>Ferreiros</AuLastName>
<AuFirstName>Carlos M.</AuFirstName>
<AuCollectiveName>Ferreiros, Carlos M.</AuCollectiveName>
</fullauthorname>
</authors>
<reprint>
<rp_author>Ferreiros, CM</rp_author>
<rp_address>Univ Santiago de Compostela, Fac Farm, Dept Microbiol &
Parasitol, Campus Sur, Santiago De Compostela 15782, Spain</rp_address>
<rp_organization>Univ Santiago de Compostela</rp_organization>
<rp_suborganizations count="2">
<rp_suborganization>Fac Farm</rp_suborganization>
<rp_suborganization>Dept Microbiol & Parasitol</rp_suborganization>
</rp_suborganizations>
<rp_street>Campus Sur</rp_street>
<rp_city>Santiago De Compostela</rp_city>
<rp_country>Spain</rp_country>
<rp_zips count="1">
<rp_zip location="AC">15782</rp_zip>
</rp_zips>
</reprint>
<research_addrs count="2">
<research>
<rs_address>Univ Santiago de Compostela, Fac Farm, Dept Microbiol &
Parasitol, Santiago De Compostela 15782, Spain</rs_address>
<rs_organization>Univ Santiago de Compostela</rs_organization>
<rs_suborganizations count="2">
<rs_suborganization>Fac Farm</rs_suborganization>
<rs_suborganization>Dept Microbiol & Parasitol</rs_suborganization>
</rs_suborganizations>
<rs_city>Santiago De Compostela</rs_city>
<rs_country>Spain</rs_country>
<rs_zips count="1">
<rs_zip location="AC">15782</rs_zip>
</rs_zips>
</research>
<research>
<rs_address>Natl Ctr Microbiol, Lab Reference Neisseria, Majadahonda,
Spain</rs_address>
<rs_organization>Natl Ctr Microbiol</rs_organization>
<rs_suborganizations count="1">
<rs_suborganization>Lab Reference Neisseria</rs_suborganization>
</rs_suborganizations>
<rs_city>Majadahonda</rs_city>
<rs_country>Spain</rs_country>
</research>
</research_addrs>
<refs count="36">
<ref>128323248</ref>
<ref>126554286</ref>
....
<ref>47450055</ref>
</refs>

<cite>
```

```
<SP>S</SP><CK>ADUB0605030021VJ</CK><AU>ADUBOBIE
J</AU><J2>VACCINE</J2><VL>21</VL><BP>605</BP><PY>2003</PY><AN>
</AN><LOC>128323248</LOC><REF>92835033</REF><RECID>128323248</R
ECID><NR>0</NR>
</cite>

<cite>
<SP>S</SP><CK>BALM0717020051JP</CK><AU>BALMER P</AU><J2>J MED
MICROBIOL</J2><VL>51</VL><BP>717</BP><PY>2002</PY><AN>
</AN><LOC>126554286</LOC><REF>91812694</REF><RECID>126554286</R
ECID><NR>0</NR>
</cite>

.....

<cite>
<SP>S</SP><CK>VAIT0046810073MJ</CK><AU>VAITUKAITIS
JL</AU><J2>METHODS ENZYMOLOGY
B</J2><VL>73</VL><BP>46</BP><PY>1981</PY><AN>
</AN><LOC></LOC><REF>779521</REF><RECID>47450055</RECID><NR>0<
/NR>
</cite>

<categories count="2">
<category>Biotechnology & Applied Microbiology</category>
<category>Microbiology</category>
</categories>
<headings count="1">
<heading>Life Sciences & Biomedicine</heading>
</headings>

<issn>1139-6709</issn>
<ids avail="Y">V44GH</ids>
<bib_vol issue="4" volume="9">9 (4): </bib_vol>
<bib_date date="DEC" year="2006">DEC 2006</bib_date>
<pubtype code="J">Journal</pubtype>
<subjects count="2">
<subject code="DB" edition="SCI">BIOTECHNOLOGY & APPLIED
MICROBIOLOGY</subject>
<subject code="QU" edition="SCI">MICROBIOLOGY</subject>
</subjects>

</REC>

<REC inst_id="2" recid="156148646" hot="yes" sortkey="3117154304"
timescited="2" sharedrefs="0" inpi="false">

<ut>000202946200009</ut>
<source_title>CHEMMEDCHEM</source_title>
<source_abbrev>CHEMMEDCHEM</source_abbrev>
<item_title>Determination of the bound conformation of a competitive nanomolar
inhibitor of Mycobacterium tuberculosis type II dehydroquinase by NMR
spectroscopy</item_title>
<bib_id>1 (9): 990-996 SEP 2006</bib_id>
<article_nos count="1">
<article_no>DOI 10.1002/cmdc.200600100</article_no>
</article_nos>
```

```
<editions full="SCI"/>
<languages count="1">
<primarylang code="EN">English</primarylang>
</languages>
<authors count="9">
<primaryauthor>Prazeres, VFV</primaryauthor>
<fullauthorname>
<AuRole>Author, Reprint Author</AuRole>
<AuLastName>Prazeres</AuLastName>
<AuFirstName>Veronica F. V.</AuFirstName>
<AuCollectiveName>Prazeres, Veronica F. V.</AuCollectiveName>
</fullauthorname>
<author key="6098186">Sanchez-Sixto, C</author>
<fullauthorname>
<AuRole>Author</AuRole>
<AuLastName>Sanchez-Sixto</AuLastName>
<AuFirstName>Cristina</AuFirstName>
<AuCollectiveName>Sanchez-Sixto, Cristina</AuCollectiveName>
</fullauthorname>
<author key="1111437">Castedo, L</author>
<fullauthorname>
<AuRole>Author</AuRole>
<AuLastName>Castedo</AuLastName>
<AuFirstName>Luis</AuFirstName>
<AuCollectiveName>Castedo, Luis</AuCollectiveName>
</fullauthorname>
<author key="1047890">Canales, A</author>
<fullauthorname>
<AuRole>Author</AuRole>
<AuLastName>Canales</AuLastName>
<AuFirstName>Angeles</AuFirstName>
<AuCollectiveName>Canales, Angeles</AuCollectiveName>
</fullauthorname>
<author key="1047628">Canada, FJ</author>
<fullauthorname>
<AuRole>Author</AuRole>
<AuLastName>Canada</AuLastName>
<AuFirstName>Francisco Javier</AuFirstName>
<AuCollectiveName>Canada, Francisco Javier</AuCollectiveName>
</fullauthorname>
<author key="3392841">Jimenez-Barbero, J</author>
<fullauthorname>
<AuRole>Author</AuRole>
<AuLastName>Jimenez-Barbero</AuLastName>
<AuFirstName>Jesus</AuFirstName>
<AuCollectiveName>Jimenez-Barbero, Jesus</AuCollectiveName>
</fullauthorname>
<author key="3960634">Lamb, H</author>
<fullauthorname>
<AuRole>Author</AuRole>
<AuLastName>Lamb</AuLastName>
<AuFirstName>Heather</AuFirstName>
<AuCollectiveName>Lamb, Heather</AuCollectiveName>
</fullauthorname>
<author key="2883419">Hawkins, AR</author>
<fullauthorname>
<AuRole>Author</AuRole>
```

```
<AuLastName>Hawkins</AuLastName>
<AuFirstName>Alastair R.</AuFirstName>
<AuCollectiveName>Hawkins, Alastair R.</AuCollectiveName>
</fullauthorname>
<author key="2540556">Gonzalez-Bello, C</author>
<fullauthorname>
<AuRole>Author</AuRole>
<AuLastName>Gonzalez-Bello</AuLastName>
<AuFirstName>Concepcion</AuFirstName>
<AuCollectiveName>Gonzalez-Bello, Concepcion</AuCollectiveName>
</fullauthorname>
</authors>
<reprint>
<rp_author>Prazeres, VFV</rp_author>
<rp_address>Univ Santiago de Compostela, Lab Quim Organ, CSIC, Ave Ciencias
S-N, Santiago De Compostela 15782, Spain</rp_address>
<rp_organization>Univ Santiago de Compostela</rp_organization>
<rp_suborganizations count="2">
<rp_suborganization>Lab Quim Organ</rp_suborganization>
<rp_suborganization>CSIC</rp_suborganization>
</rp_suborganizations>
<rp_street>Ave Ciencias S-N</rp_street>
<rp_city>Santiago De Compostela</rp_city>
<rp_country>Spain</rp_country>
<rp_zips count="1">
<rp_zip location="AC">15782</rp_zip>
</rp_zips>
</reprint>
<refs count="49">
<ref>114130619</ref>
<ref>101185237</ref>
<ref>156148726</ref>
.....
<ref>142787158</ref>
</refs>

<cite>
<SP>S</SP><CK>ABEL057399 CC</CK><AU>ABELL
C</AU><J2>COMPREHENSIVE NATURA</J2><VL>
</VL><BP>573</BP><PY>1999</PY><AN>
</AN><LOC></LOC><REF>76144101</REF><RECID>114130619</RECID><NR
>0</NR>
</cite>

<cite>
<SP>S</SP><CK>ADAM0249970325BE</CK><AU>ADAMS J
ER</AU><J2>BIOCHEM
1</J2><VL>325</VL><BP>249</BP><PY>1997</PY><AN>
</AN><LOC>101185237</LOC><REF>41019537</REF><RECID>101185237</R
ECID><NR>0</NR>
</cite>

.....
<cite>
```

```
<SP>S</SP><CK>YAO 0661050579FH</CK><AU>YAO HL</AU><J2>FEBS  
LETT</J2><VL>579</VL><BP>661</BP><PY>2005</PY><AN>DOI  
10.1016/j.febslet.2004.11.110</AN><LOC>142787158</LOC><REF>105643016  
</REF><RECID>142787158</RECID><NR>0</NR>  
</cite>  
  
<categories count="2">  
<category>Chemistry, Medicinal</category>  
<category>Pharmacology & Pharmacy</category>  
</categories>  
<headings count="1">  
<heading>Life Sciences & Biomedicine</heading>  
</headings>  
  
<issn>1860-7179</issn>  
<ids avail="Y">V43PE</ids>  
<bib_vol issue="9" volume="1">1 (9): </bib_vol>  
<bib_date date="SEP" year="2006">SEP 2006</bib_date>  
<pubtype code="J">Journal</pubtype>  
<subjects count="2">  
<subject code="DX" edition="SCI">CHEMISTRY, MEDICINAL</subject>  
<subject code="TU" edition="SCI">PHARMACOLOGY & PHARMACY</subject>  
</subjects>  
  
</REC>  
.....  
</RECORDS>
```

- Tódolos campos son importantes pero especialmente destacamos os seguintes polo tratamento que se ten que facer deles:
 - Autores: é necesario un proceso de normalización para non duplicar autores (ver o proceso que se detalla máis adiante).
 - Institucións: pola forma en que aparecen é necesario normalizar as direccións para asignalas a unha institución normalizada (ver o proceso que se detalla máis adiante).
 - Referencias citadas (Bibliografía): é imprescindible para realiza-las análises de cocitacións. Para elo é preciso identificar correctamente os documentos aos que fai referencia un documento concreto. Cada referencia citada aparece na forma: Autor, Ano, Título abreviado da publicación, Volume e Páxina.
Na API ademais do número de documentos (<refs count="49">) aparece a información sobre cada un deles.
 - Citas: nos ficheiros obtidos a través da plataforma só ven o número de citas, nembargante nos ficheiros da API non aparece este número, senón que aparece a lista de citas (<cite>). É importante este campo á hora de traballar coas cocitacións.

- Materias: as materias que veñen asociadas a cada rexistro son materias ISI. As cales téñense que relacionar coas outras clasificacións de materias existentes. Ditas materias poden variar ao longo do tempo é ademais de actualizar os datos cos novos listaxes débese manter un histórico para poder facer comparacións futuras entre estudos sobre materias.
 - Scopus:
 - A través da plataforma:
Un exemplo sería:

TY - JOUR
T1 - Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction
JF - Analytical Biochemistry
JA - ANAL. BIOCHEM.
VL - 162
IS - 1
SP - 156
EP - 159
PY - 1987///
SN - 00032697 (ISSN)
AU - Chomczynski, P.
AU - Sacchi, N.
AD - Laboratory of Biochemistry and Metabolism, NIDDK, National Institutes of Health, Bethesda, MD 20892, United States
AB - A new method of total RNA isolation by a single extraction with an acid guanidinium thiocyanate-phenol-chloroform mixture is described. The method provides a pure preparation of undegraded RNA in high yield and can be completed within 4 h. It is particularly useful for processing large numbers of samples and for isolation of RNA from minute quantities of cells or tissue samples.
KW - messenger rna
KW - rna
KW - gene expression
KW - genetic engineering
KW - human
KW - human cell
KW - priority journal
KW - Animal
KW - Cell Line
KW - Chloroform
KW - Guanidine
KW - Guanidines
KW - Human
KW - Hydrogen-Ion Concentration
KW - Mammae
KW - Phenol
KW - Phenols
KW - Rats
KW - RNA
KW - Solutions
N1 - Cited By (since 1996): 39425
N1 - Export Date: 19 December 2007
N1 - Source: Scopus

N1 - CODEN: ANBCA
 N1 - Language of Original Document: English
 N1 - Chemicals/CAS: RNA, 63231-63-0; Chloroform, 67-66-3; Guanidine, 113-00-8; Guanidines; Phenol, 108-95-2; Phenols; RNA, 63231-63-0; Solutions
 UR - <http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-0023277545&partnerID=40&rel=R7.0.0>
 ER -

A explicación de que indica cada campo sería a seguinte:

Etiqueta	Descrición
Etiquetas de título e referencia	
TY	Tipo de referencia (Tipo de Documento) Debe conter un dos valores definidos na táboa Tipo de Referencia (ve táboa Tipos de Referencia)
ER	Fin de referencia. Debe ser a última etiqueta nunha referencia.
ID	Identificador de referencia. Pode consistir en calquera carácter alfanumérico ata unha lonxitude máxima de 20.
T1 TI CT BT	Título primario. A etiqueta BT emprégase unicamente para libros completos e referencias a traballos non publicados. Este campo pode conter caracteres alfanuméricos; practicamente non existe limitación na lonxitude deste campo.
T2 BT	Título secundario. A etiqueta BT emprégase para tódolos tipos de referencias agás para libros completos e referencias a traballos non publicados; practicamente non existe limitación na lonxitude deste campo.
T3	Título da serie. Este campo pode conter caracteres alfanuméricos; practicamente non existe limitación na lonxitude deste campo.
Etiquetas de autor	
A1 AU	Autor principal. Cada autor debe estar nunha liña separada, precedido desta etiqueta. Cada referencia pode conter un número ilimitado de autores, cunha lonxitude de ata 255 caracteres para cada valor. O nome do autor debe respecta-la seguinte sintaxes: Apelido, Nome, Sufixo. Como nome poden aparecer nomes completos, iniciais ou ámbolos dous. O sufixo é opcional; p.e.: Jr/Sr/II/MD, etc.
A2 ED	Autor secundario. Cada autor debe estar nunha liña separada, precedido desta etiqueta. Terá a mesma sintaxes que o autor principal (A1, AU) e pode ter unha lonxitude de ata 255 caracteres.
A3	Autor da serie. Cada autor debe estar nunha liña separada, precedido desta etiqueta. Terá a mesma sintaxes que o autor principal (A1, AU) e pode ter

Etiqueta	Descripción
	unha lonxitude de ata 255 caracteres.
Etiquetas de ano e texto libre	
Y1 PY	Data principal. Debe estar no seguinte formato: YYYY/MM/DD/outra información. O ano, mes e día son datos numéricos. O campo "outra información" pode conter letras, espazos e guións. Cada unha da información específica da data é opcional, pero non as barras (/) non son opcionais. p.e.: se aparece unicamente ano e outra información: 1998///Spring.
Y2	Data secundaria. Co formato indicado para as etiquetas Y1 e PY.
N1 N2	Notas. Conteñen texto libre con caracteres alfanuméricos; practicamente non hai limitación para a lonxitude deste campo. Detrás desta etiqueta atópase información importante para o sistema. As referencias citadas a outro documento aparecerán coa etiqueta "N1" seguida da palabra "References". Por exemplo: N1 - References: Hall, S.R., Allen, F.H., Brown, I.D., (1991) Acta Cryst., A47, p. 655 As citas feitas a un documento teñen, despois do valor "N1", a palabra "Cited By". Por exemplo: Cited By (since 1996): 39425
AB	Resumo. Texto libre con caracteres alfanuméricos; practicamente non ten limitación na súa lonxitude.
Etiquetas de palabras clave e estado de reimpresión	
KW	Palabras clave. Cada palabra clave ou frase estará nunha liña nova, precedida da correspondente etiqueta. Poden ser palabras ou frases completas, cunha lonxitude de ata 255 caracteres. Practicamente non hai limitación ao número de palabras clave.
RP	Estado de reimpresión. Campo opcional. Pode toma-los valores "IN FILE", "NOT IN FILE", "ON REQUEST (MM/DD/YY) [Non se emprega]
Etiquetas de publicación	
JF JO	Nome completo de publicación periódica. Valor alfanumérico de ata 255 caracteres.
JA	Abreviatura estándar de nome de publicación periódica. Indica a publicación na que o artigo foi (ou vai ser, no caso de referencias en prensa) publicado. Valor alfanumérico de ata 255 caracteres. Si é posible, as abreviaturas seguirán o estilo do <i>Index Medicus</i> , con períodos despois das abreviaturas.
J1	Nome de publicación periódica: abreviatura de usuario 1. Lonxitude de ata 255 caracteres.
J2	Nome de publicación periódica: abreviatura de usuario 2. Lonxitude de ata 255 caracteres.

Etiqueta	Descrición
VL	Número de volume.
IS	Número de publicación.
SP	Páxina inicial.
EP	Páxina final.
CP	Cidade de publicación.
PB	Editor
SN	ISSN/ISBN
AD	Dirección completa da institución asociada ao documento referenciado. Se se trata dunha colaboración e polo tanto hai máis dunha institución, cada unha delas ven codificada nunha nova liña trala etiqueta "AD".
Etiquetas varias	
AV	Dispoñibilidade.
M1 M2 M3	Miscelánea (de 1 a 3 valores).
U1 U2 U3 U4 U5	Valores definidos polo usuario (de 1 a 5).
UR	Web/URL. Practicamente non hai límite á lonxitude deste campo. As direccións URL poden atoparse de xeito individual, unha por etiqueta ou múltiples direccións separadas por punto e coma.
L1	Enlace a PDF.
L2	Enlace a texto completo.
L3	Rexistros relacionados.
L4	Imaxes

- A través da API: descoñécese en que formato se obtén a información pero limitase bastante a información obtida. Só estarían dispoñibles os seguintes campos:

Etiqueta	Descrición
Article Title	Título do artigo
Document Type	Tipo de documento
Cited-by Count	Número de citas
INWARD URL	URL
Source Title	Título da revista
ISSN	ISSN
Vol	Volume
Issue	Número
Page	Páxina
Publication Date	Data de publicación
EID	Identificador único específico de Elsevier para un

	documento en SCOPUS
SCP	Identificador único de SCOPUS para cada documento en SCOPUS
DOI	Digital Object Identifier
First Author	Autor Principal
Complete Author list	Lista de autores
Affiliations	Institucións
Abstracts	Abstracts dos documentos

Esta limitación afecta á hora de calcular as cocitacións, xa que é imposible calculas ao non vir información sobre a bibliografía e as citas nos rexistros descargados de SCOPUS.

Ao igual que con ISI é necesario a normalización tanto de autores coma de institucións, así coma a xestión das materias dos rexistros.

- o BDs do "Instituto de Estudios Documentales sobre Ciencia y Tecnología", IEDCYT, (do "Consejo Superior de Investigaciones Científicas", CSIC): Consta de tres bases de datos:

- ICYT: Ciencia e Tecnoloxía.
- ISOC: Ciencias Sociais e Humanas.
- IME: Biomedicina.

O IEDCYT dispón dun acceso web na Internet que permite descarga-la información de publicacións, aínda que esta opción é pouco operativa debido a limitación do número de rexistros á hora da descarga, só descarga os 100 primeiros rexistros.

Outra posibilidade é solicita-la información en formato electrónico. Neste caso, o CD-ROM facilitado polo IEDCYT contén os datos bibliográficos rexistrados nun formato propio e unha aplicación desenvolvida polo IEDCYT para explota-los.

En calquera caso, o IEDCYT non dispón polo momento da posibilidade de obte-la información vía Web Services, como no caso de Scopus ou ISI.

Cada unha das bases de datos ten unha estrutura similar pero con certas diferenzas.

Certos campos que se poden sinalar como importantes polo traballo engadido que levan son:

- Autores: será necesario tamén aplicar o proceso de normalización.
- Lugar de traballo (Institucións): será necesario tamén aplicar o proceso de normalización.
- Publicacións: Ademais da información sobre as publicacións que ven en cada rexistro, o IEDCYT dispón dun directorio de revistas que pode ser de utilidade para cargar no sistema.
- Clasificación: Para os rexistros da base de datos ICYT a clasificación é por materias UNESCO, polo que é necesario manter as relacións entre as distintas

clasificacións de materias, para á hora de facer estudos ter unha uniformidade.

No caso dos documentos ISOC a clasificación é diferente, polo que será necesario implementar as medidas necesarias para poder facer a relación entre esta clasificación e o resto de clasificacións de materias. Os documentos de IME non teñen unha clasificación de materias asociada.

- Bibliografía e Citas: Nestas bases de datos nos se ten información sobre citas e bibliografía polo que non se van a poder facer estudos de cocitación.

A continuación móstrase un exemplo para cada unha das bases de datos:

- **ICYT: Ciencia e Tecnoloxía.**

Base: ICYT
Documento: 173468
Autor: Rodríguez Bugarín, Miguel; Novales Ordax, Margarita; Orro Arcay, Alfonso
Título: Alta velocidade y territorio. Algunas experiencias internacionais
Inglés-Título:
Otro-Idioma-Título:
Lugar-Trabaja: Univ.A Coruña,A Coruña,España
Obra-Colectiva-Título:
Revista: Ingeniería y territorio
Datos-Fuente: 2005, (70): 4-11, 8 Ref
ISSN: 1695-9647
ISBN:
Idioma: Español
Tipo-Documento: Artículo de revista
Modo-Documento:
Clasificación: 332304 TRANSITO RAPIDO
Congreso:
Descriptor: Ferrocarril de alta velocidade; Transporte ferroviario
Identificadores:
Topónimos: Japón; Francia; Alemania
Resumen:
Ing-Resumen:
Url:

- **ISOC: Ciencias Sociais e Humanas**

Base: ISOC
Documento: 543181
Autor: Acuña Fariña, Juan Carlos
Título: Aspectos de la relación entre teorías de gramática y teorías de procesamiento
Inglés-Título: Aspects of the relationship between theories of grammar and theories of processing
Otro-Idioma-Título:
Lugar-Trabaja: Univ. Santiago de Compostela, España
Revista: Atlantis. Revista de la Asoc. Esp. de Est. Anglo-Norteamericanos
Obra-Colectiva:

Datos-Fuente: 2005, 27(1): 11-27, 83 Ref
ISSN: 0210-6124
ISBN:
Idioma: Inglés
Tipo-Documento: Artículo de revista
Modo-Documento:
Clasificación: 550201 Modelos. Conceptos teóricos y metodológicos
550700 Psicolingüística. Obras generales
550905 Lingüística e informática
Compilador-Director:
Congreso:
Institución-Patrocinadora:
Editor:
Serie:
Signatura:
Notas: Resumen (En)
Descriptores: Teorías gramaticales; Psicolingüística; Investigación científica;
Procesamiento del lenguaje natural; Análisis lingüístico; Interdisciplinariedad
Identificadores:
Topónimos:
Legislación:
Jurisprudencia:
Periodo-Histórico:
Décadas:
Siglo:
Resumen:
Ing-Resumen:
Localización:
Url: http://www.atlantisjournal.org/Papers/27_1/011-027%20Acuña.pdf

▪ **IME: Biomedicina**

Base: IME
Documento: 279384
Autor: Barreiro, M; Núñez, C; Domínguez Muñoz, JE; Lorenzo, A; Barreiro, F; Potel, J; Peña, AS
Título: Asociación entre las mutaciones del gen NOD2/CARD15 con los antecedentes quirúrgicos por enfermedad de Crohn
Otro-Idioma-Título:
Lugar-Trabajo: Hospital Clínico Universitario. Departamento de Gastroenterología. Santiago de Compostela. La Coruña. España; Vrije Universiteit Medisch Centrum. Laboratory of Immunogenetics. Amsterdam. Holanda; Hospital Clínico Universitario. Departamento de Cirugía. Santiago de Compostela. La Coruña. España; Vrije Universiteit Medisch Centrum. Department of Gastroenterology. Amsterdam. Holanda
Revista: Revista Española de Enfermedades Digestivas
Datos-Fuente: 2005, 97 (8): 547-553, 25 ref.
ISSN: 1130-0108
Idioma: Inglés
Tipo-Documento: Artículo de revista
Modo-Documento:
Notas:
Descriptores:
Resumen:
Ing-Resumen:
Url:

- o Outras fontes de datos
 - Información sobre as clasificacións das revistas segundo as materias ISI, ANEP e Ámbitos de coñecemento.
 - Unha materia ISI pode estar relacionada con varias materias ANEP e a súa vez unha materia ANEP pode ter asociados varios ámbitos de coñecemento.
 - Estas tres clasificacións poden variar ao longo do tempo polo que a asignación tería que variar, pero mantendo tamén información sobre as versións anteriores. Para que se poida optar por algunha das asignacións anteriores ou pola nova.
 - A súa vez, tamén existen outras clasificacións como son a UNESCO e a do ISOC que se relacionarán tamén coas materias anteriores, polo que en ambos casos se deberá ter en conta para facilitar aos usuarios a carga/modificación/eliminación de dita información.
 - Pode ser necesario a inclusión de novas clasificacións de materias así coma a relación entre as novas clasificacións e as xa existentes na base de datos.
 - Listados de revistas de cada fonte de datos.

Non se descarta a posibilidade de novas fontes de datos, similares a ISI, SCOPUS, ... que inclúan información sobre a produción de investigadores galegos.

A maiores tamén pode incluírse información adicional do tipo de listados de materias, revistas, información máis detallada dos investigadores, ...

Outra posibilidade é que se amplíen os datos almacenados na base de datos con información máis específica, como pode ser a data de nacemento.

Os casos deben terse en conta á hora do deseño de todo o sistema para facilitar a súa inclusión posterior..

- Desenvolvemento das aplicacións necesarias para a extracción de información das devanditas fontes de datos. Na maioría das fontes será necesario o desenvolvemento de aplicacións para a descarga automática dos rexistros (a través das APIs) das devanditas bases de datos, así coma a carga dos rexistros na base de datos. Ambas APIs teñen certas limitacións co número de rexistros a descargar de cada vez, no caso de ISI o límite é de 100 rexistros de cada vez, e no caso de Scopus de 2.000; polo que é necesario ter en conta estas limitacións á hora de desenvolver uns sistemas que a través de iteracións permitan descargar tódolos rexistros de cada busca. Debido ao tempo que pode levar as descargas, tamén tense que ter en conta a posibilidade de que as APIs non respondan e haxa

que manterse a espera para non perder os datos descargados ata o momento.

No caso de SCOPUS a maiores dos límites de rexistros existen tamén limitacións en canto a información que se pode descargar deste xeito, xa que non aparecen (nestes intres) datos sobre as citas e a bibliografía, entre outros (máis información no apartado sobre estudo das fontes de datos e no informe de consultaría). É dicir, non se pode obter información a través da API de SCOPUS sobre as cocitacións, ao contrario que con ISI.

- Creación dunha base de datos que permite o almacenamento da información obtida das distintas fontes, é dicir, deseñar un modelo de datos que:
 - Facilite a explotación da información almacenada.
 - Sexa flexible na evolución do sistema.
 - Permita a identificación inequívoca do orixe dos datos.

Debe buscarse un deseño que facilite a consulta de información e a integración de datos procedentes de diferentes fontes, para o que se poderían tomar como base as dimensións atopadas no proceso de consultaría aínda que o seu número pode variar a medida que se avance no desenvolvemento do sistema:

- Data: Contén as datas para as que existe información no sistema o que permite, entre outras cousas, facer análise por ano de publicación dun documento.
- Documento: Documentos dispoñibles no Sistema (artigos pertencentes a unha publicación, actas, teses, ...) e a información que os caracteriza: título, volume, serie, ...
Integrar non sistema os documentos procedentes de cada unha das fontes é unha tarefa complexa. No momento do deseño do modelo de datos debe decidirse se é factible traballar cunha lista normalizada de documentos onde se especifique a través dun atributo a súa procedencia, ou é mellor manexar relacións de documentos independentes para cada unha das fontes.
- Publicación: Relación das distintas publicacións para as que se ten recollido información de produción científica. Habería que busca-lo xeito de dispoñer dunha única relación normalizada de publicacións. Neste caso poderían empregarse como valores de referencia as listas de publicacións existentes en Journal Citation Reports e en Scopus.
Nesta dimensión será preciso articula-los mecanismos de carga e actualización de datos necesarios para coñece-la evolución das publicacións ao longo do tempo: cando unha publicación cambia de nome, cando se fusiona con outra, cambia de temática, ...
- Materia: Relación de materias científicas nas que pode clasificarse un documento publicado. Por exemplo: Economía, Química, Enxeñería Civil e Arquitectura, ...
Cada unha das fontes de datos emprega unha clasificación por materias diferente. Tódalas materias deben relacionarse coa clasificación ANEP, e esta a súa vez coa clasificación de

ámbitos de coñecemento. Así mesmo, para tódalas clasificacións de materias empregadas en cada unha das fontes de datos, debe contemplarse a actualización e reorganización das mesmas: por exemplo desglosando unha materia en máis detalle ou incorporando novas categorías.

- o Institución: Conxunto de Institucións Científicas para as que se pode consultar datos de produción científica. Partirase dunha información previa a cal se terá que ir actualizando a medida que se faga a carga de rexistros e a normalización. Debe habilitarse unha ferramenta de fácil manexo para que un usuario cos permisos axeitados poida actualiza-los datos das institucións e decidir a que institución pertence cada dirección de cada documento. Este proceso deberá facerse para cada proceso de carga, pero tamén é necesario que se poida facer a posteriori. Ademais, debe facer un "varrido" sobre as institucións a incorporar ao sistema presentando ao usuario unha lista reducida de datos a revisar.
- o Sector: Sectores nos que se poden encadra-las diferentes institucións rexistradas no Sistema. Por exemplo: Universidade, Sistema Sanitario, Administración, Centros mixtos con dobre adscrición institucional, ...
- o Autor: Relación de autores que teñan publicado algún dos documentos dispoñibles no Sistema de Análise. Non existe unha lista de referencia ou normalizada dende a que obter os autores. Cada unha das fontes de datos que alimenta o sistema identifica aos autores que publicaron un documento empregando unha notación propia. Para cada documento ven rexistrado o autor ou autores mediante o apelido ou apelidos e o nome (ás veces unicamente a inicial do nome). Así, atoparemos nomes distintos que fan referencia ao mesmo autor e nomes idénticos que fan referencia a diferentes autores. Debe habilitarse unha ferramenta de fácil manexo para que un usuario cos permisos axeitados poida actualiza-los datos de autores e decidir se un mesmo nome fai referencia á mesma persoa ou dous nomes diferentes son o mesmo autor, este proceso deberá facerse para cada proceso de carga, pero tamén é necesario que se poida facer a posteriori. Ademais, debe facer un "varrido" sobre os autores a incorporar ao sistema presentando ao usuario unha lista reducida de datos a revisar.
- o Tipo de Documento: Identifica os diferentes tipos de documento nos que clasifica a información de produción científica cada unha das fontes de datos. Por exemplo: Resumo, Artigo, Acta, Tese, ... Cada unha das fontes codifica os tipos de documento dun xeito diferente. En principio non se considera necesaria a definición dunha lista normalizada de Tipos de Documento.
- o Xeografía: Localización xeográfica asociada a cada unha das Institucións Científicas dispoñibles no Sistema de Análise. Habilitar a información de localización de xeito independente

da relación de Institucións facilita a explotación de datos. Debe dar información sobre o país, comunidade autónoma e provincia.

- Tipo de Colaboración: Para aqueles documentos que foron elaborados en colaboración por varias institucións identifica, por ámbito xeográfico, o tipo de colaboración. Posibles valores: Galega, Interrexional e Estranxeira. Esta clasificación faría máis doados os estudos por colaboración.

Na consultaría tamén se destacan certos indicadores, que serían unha propiedade da información que pode ser expresada nun valor cuantitativo e que é preciso medir. Estes indicadores serían unha primeira aproximación que pode ter que ampliarse á hora de levar a cabo o deseño do sistema:

- Indicadores cuantitativos:
 - Documentos
 - Documentos citables
 -
- Indicadores cualitativos:
 - Factor de impacto
 -
- Indicadores de estrutura e relacións
 - Documentos citados
 - Documentos non citados
 - Citas recibidas
 - Citas realizadas
 - Autocitas
 - Documentos en coautoría
 -
- Indicadores de colaboración:
 - Documentos en colaboración
 - ...
- Indicadores derivados cuantitativos:
 - % documentos
 - % documentos citables
 - Tasa de variación interanual
 - Índice de esforzo temático
 -
- Indicadores derivados cualitativos:
 - Factor de impacto normalizado
 - Factor de impacto medio esperado relativo
 - Factor de impacto medio das revistas onde se producen as citas
 - Factor de impacto ponderado
 - Potencial investigador
 - Potencial investigador comparado
 -
- Indicadores derivados de estrutura e relacións:
 - % de documentos citados
 - % de Documentos non citados
 - Índice relativo de citación

- Taxa de documentos en coautoría
- Índice de coautoría
-
- Indicadores derivados de colaboración:
 - Taxa de documentos en colaboración
 - Potencial investigador comparado
 - Índice de internacionalización
 - Índice de coautoría asimétrica
 -

Mais información sobre estas dimensións e indicadores pode verse no informe da consultaría previa.

- Aproveitamento das distintas bases de datos existentes para dar máis información a devandita base de datos. As bases de datos existentes son:
 - BD que almacena a información sobre a produción científica dos investigadores do SUG: nesta base de datos tense a seguinte información:
 - Grupos de investigación das tres universidades galegas.
 - Asociados aos grupos de investigación, poderase engadir información sobre as noticias publicadas na web ou no blog do consorcio que teñan relación co devandito grupo.
 - Investigadores asignados a cada grupo:
 - Nome
 - Nome normalizado segundo as instrucións da FECYT
http://www.accesowok.fecyt.es/recomendaciones_publicaciones.html
 - Cargo
 - Tipo de investigador
 - Email
 - Índice H
 - DNI
 - Sexo
 - Producción de cada investigador
 - Artigos
 - Capítulos de libros
 - Congresos
 - Libros
 - Teses
 - Revisións (Reviews)
 - Contratos con Empresas
 - Outros Proxectos
 - Datos do JCR (Journal Citation Reports)
 - Dentro dos servizos proporcionados polo ISI Web of Knowledge atópase Journal Citation Reports, que ofrece información de publicacións científicas a nivel mundial e datos do seu impacto e influencia na comunicade científica

- Os JCRs actualízanse cada ano pero ten que manterse toda a información anterior.
- Journal Citation Reports ten dúas edicións:
 - JCR Science Edition: Contén os datos de máis de 5.900 revistas en 172 categorías temáticas.
 - JCR Social Sciences Edition: Contén máis de 1.700 revistas en 55 categorías temáticas.
- Journal Citation Reports contén, para un ano, información de publicacións científicas, indicando o seu título, título abreviado, factor de impacto, ... Ademais engádense a estes datos información sobre a clasificación das revistas dentro de cada materia correspondente
- Máis información dispoñible no informe de auditoría.
- Relacións entre as distintas formas de aparición das institucións.
 - Relaciona a dirección que aparece no rexistro coa correspondente institución.
 - Dispónse dunha lista inicial que deberá actualizarse a medida que se vaian cargando novos rexistros e se vaia completando o proceso de normalización.
 - A información de partida obtense dunha base de datos externa a nós, e que ten a seguinte estrutura:

Nome	Descrición
Idins	Identificador
IdInsa	
IdInsNor	Identificador da institución normalizada
rs_organization	Nome da institución procedente de ISI.
organizacion	Nome da institución modificada pola U. Granada.
country	Código do país
rs_city	Nome da cidade procedente de ISI
ciudad	Nome da cidade modificado pola U. Granada
ccaa	Código da Comunidade Autónoma.
sector	Sector a que se dedica a institución
normalizada	
aceptada	
idssinsti	
observaciones	
usuario	
fecha	
hora	
fecham	
fechacarga	
elentic	
aborg	
borrado	
activa	
ncat	

Cando coinciden o campo Idins e IdInsNor estamos ante unha institución normalizada.

Tamén se dispón dunha táboa similar para as institucións non españolas, a cal haberá que introducir no sistema e actualizar á hora de facer a carga.

idinse	idinsnor	rs_organization	rs_city	fechacarga	numcountry	origen
10	10	106 Ippocratous St	Athens	18/04/2007	32	
32	32	Sir Alister Hardy Fdn Ocean Sci Lab	Plymouth	18/04/2007	6	
33	33	Sir Charles Gairdiner Hosp	Nedlands	18/04/2007	36	

- Relacións entre as institucións e os sectores: Partindo dunha táboa inicial que relacione as institucións normalizadas cos sectores nos que se engloban (Universidade, Administración, Servizo Sanitario, ...) debe permitirse no proceso de normalización a asignación entre as institucións e os seus sectores, é dicir, tanto a inserción de novas agrupacións como a modificación das xa existentes.
- Procesos de carga dos rexistros: unha vez realizada a carga dos rexistros ou ben durante dito proceso é necesario realizar procesos de normalización tanto de autores coma de institucións
 - Normalización dos autores: Debe comprobar se os autores que se están a introducir coinciden exactamente con autores xa cargados na base de datos, en dito caso, ampliarase a súa produción coa nova produción que se está a introducir. Se non existe ese autor exactamente, ou existen varios con ese nome, ha de ofrecer ao usuario que autor/es se aproximan má is (tendo en conta a materia do artigo, as institucións que aparezan na dirección) para ofrecer os autores que xa existen na base de datos que se aproximan nun maior grado. A medida que se vaia realizando a normalización terá que gardarse información das correspondencias que se vaian realizando.
 - Normalización das institucións: As institucións que veñen nos rexistros non están normalizadas, é dicir, por exemplo para a Universidade de A Coruña podemos atopar moitas versións diferentes tal como "Univ A Coruna", "Univ La Coruña", "Universidad Coruna",
 Debido a esta variedade de nomes para cada institución, partindo dunha táboa de relación entre direccións e Institución normalizada, trátase de asignarlle os recursos as súas correspondentes institucións. No caso de non atoparse a relación ou ser confusa, debe ofrecer a posibilidade ao usuario de asignala manualmente e actualizar a táboa de correspondencia para novas cargas.
 - Asignación de sectores: asignar a cada institución, se non existe xa, o sector no que traballa.

Para levar a cabo dita normalización será necesario que o encargado/a da carga dos datos, no momento da carga ou nun momento posterior, teña acceso a unha aplicación web que lle permita realizar ditos procesos. Dita aplicación como se dicía anteriormente debe facer un "varrido" para incorporar automaticamente toda a información que sexa posible e dar unha lista reducida de datos a revisar.

- Proporcionar unha serie de informes:
 - Predefinidos: que permitan aos usuarios do sistema realizar de maneira sinxela as consultas de datos máis habituais e obter a información requirida. Os informes han de estar accesibles vía web pero tendo en conta que non todo o mundo pode ver tódolos informes. Unha opción para este tipo de informes sería empregar Reporting Services, sempre que se adapta aos requisitos previos (selección de informes a mostrar segundo o nivel de acceso dos usuarios).
 - Específicos: Serán consultas máis específicas, como poden ser os diagramas de cocitación, ou ben informes puntuais ou informes avanzados.

Tense que ter en conta que debido ao tipo de informe pode ser necesario a realización de cálculos complexos ou representacións gráficas non soportadas por Reporting Services e que haberá que desenrolar. Un exemplo deste tipo de informe serán os relacionados coa cocitación. Con este fin é indispensable que no sistema estea dispoñible toda a información que puidera precisarse para dar resposta ás consultas plantexadas.

Debe terse en conta que aínda que para algúns informes poda empregarse as ferramentas das que dispón SQL Server 2005 Enterprise Edition, será necesario o desenvolvemento dalgunha aplicación que permita a realización de informes máis complexos como será o caso dos informes de coautoría.

Exemplos de informes necesarios, aínda que poden xurdir máis necesidades ao longo do desenvolvemento do sistema ás que habería que facer fronte:

- Informes de cantidade de produción científica: Análise dos aspectos cuantitativos de produción científica, realizados a partires dos indicadores cuantitativos dispoñibles no Sistema. Algúns dos informes sobre a cantidade da produción científica que sería preciso publicar son:
 - Número de documentos por autor
 - Número de documentos por revista
 - Número de documentos por institución galega
 - Número de revistas nas que publican os investigadores/autores galegos
 - Número de documentos por ano
 - Tipos de documentos por ano, institución ou autor
 - Documentos por sectores (administración, sistema sanitario, empresa...)

- Número de documentos nunha área temática (totais e por anos)
- Autores que máis documentos publican nunha área temática
- Institucións que publican nunha área temática, número de documentos e títulos
- Revistas nas que se publican documentos dunha determinada área temática
-
- Informes de calidade da produción científica: Este conxunto de informes pretenden medir a calidade da produción científica, empregando para elo indicadores cualitativos como Factor de impacto ou o Potencial investigador. Como exemplo de informes de calidade da produción científica publicados no Sistema estarían:
 - Revistas con factor de impacto máis alto nunha área temática
 - Autores que publican nas revistas de maior factor de impacto
 - Potencial Investigador por institución
 - Potencial investigador nunha área temática
 -
- Informes de relacións estruturais: O obxectivo destes informes é amosar unha imaxe da estrutura e relacións existentes no conxunto da produción científica rexistrada no Sistema, establecida a través dos contidos temáticos das publicacións. Entre os informes desta categoría estarían as representacións multivariadas, que establecen comparacións entre os datos empregando para elo os indicadores de produción e calidade agrupados por área temática e sector, por institución e sector, por xeografía e area temática, Como exemplos de informes desta categoría estarían:
 - Documentos citables e porcentaxe dos mesmos das institucións que publican nunha área temática
 - Factor de Impacto Relativo con respecto a Galicia e a España por institución nunha área temática
 - Autores por institución
 - Autores máis citados nunha área temática
 - Citas por revista
 - Citas por autor
 - Citas por institución
 - Documentos máis citados da produción científica galega
 - Porcentaxe de documentos citables das institucións galegas
 - Táboa de cocitacións (máis información sobre as cocitacións no informe de consultaría) por área temática.
 -
- Informes de colaboración: Os informes de colaboración tamén reflicten as relacións da produción científica, neste caso en base aos autores dos documentos (máis información sobre as cocitacións no informe de consultaría).

- ...
- Posta en funcionamento do sistema realizando as cargas iniciais: o sistema ten que estar en funcionamento coa produción científica de España ata a actualidade de tódalas fontes tratadas durante a fase de deseño (ISI, SCOPUS,)
- Formación ao persoal técnico do Consorcio de Bibliotecas:
 - Na ferramenta de xeración de informes de xeito que teñan a facultade de crear novos informes cando así o precisen.
 - Na administración do portal de informes e na xestión de accesos ao mesmo.
 - No manexo de tódalas aplicacións desenroladas ao abeiro deste proxecto.

A empresa adxudicataria deberá por a dispor do proxecto unha persoa dedicada a tempo completo e de forma exclusiva ao mesmo.