

Els llenguatges d'interrogació

Ana María Checa Rubio
Pere Masip Masip

P06/09046/00265

Índex

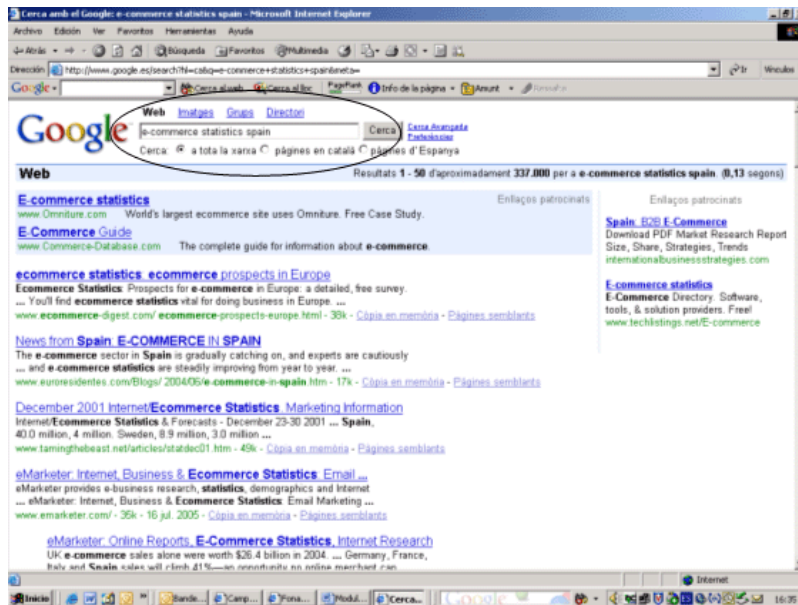
1. Definició de llenguatges d'interrogació	5
2. Operadors	10
2.1. Operadors booleans o lògics	10
2.1.1. L'ús de diferents operadors	11
2.1.2. Operador XOR	12
2.2. Operadors posicionals	13
2.2.1. Operadors posicionals absoluts	13
2.2.2. Operadors posicionals relatius	16
2.3. Cerca per frase	19
2.4. Operadors de truncament	19
2.5. Operadors de rang	20
2.6. Altres funcions	21
2.6.1. Discriminació de majúscules	21
2.6.2. Discriminació dels accents	21
3. Les estratègies de cerca	22
Bibliografia	24

1. Definició de llenguatges d'interrogació

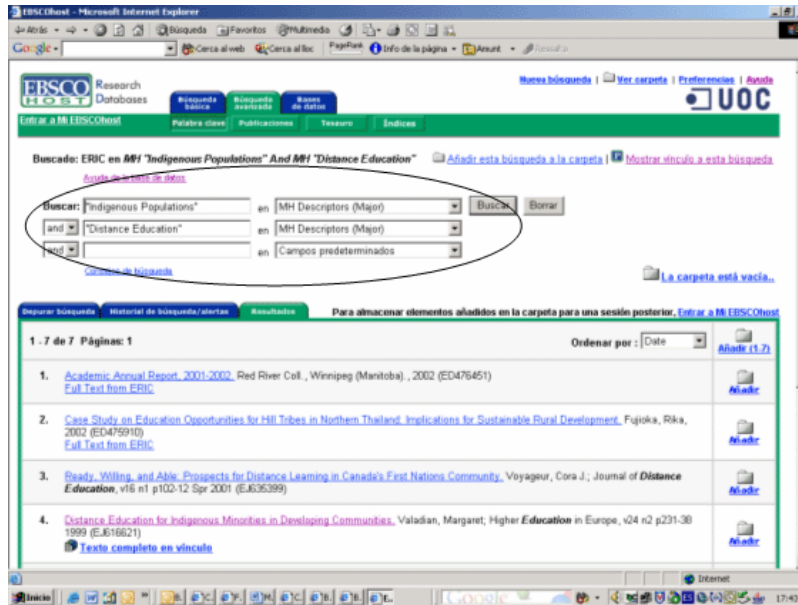
Entenem com a **llenguatge d'interrogació** el conjunt d'ordres, operadors i estructures que, organitzats segons unes normes lògiques, permeten la consulta de fons i recursos d'informació electrònica. El resultat de la combinació d'aquests elements és una expressió coneguda amb el nom d'**equació de cerca** (Tramullas, 1997, pàg. 199).

Exemples d'equació de cerca

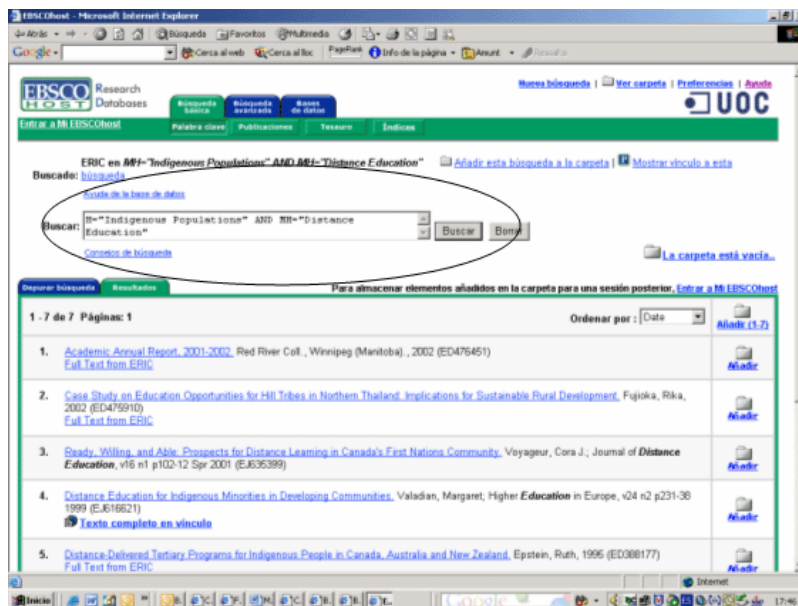
Exemple d'equació de cerca de Google



Exemple d'equació de cerca de la base de dades ERIC



La mateixa equació de cerca a la base de dades ERIC, però usant una única línia de comandaments



Al llarg dels darrers vint anys les interfícies dels llenguatges d'interrogació han evolucionat de forma considerable. Els primers llenguatges d'interrogació obligaven a la formulació escrita de complexes expressions que incloïen tots els components de l'equació de cerca: les ordres, els operadors i els termes que representaven la necessitat d'informació. En l'actualitat, en canvi, les interfícies són molt més intuïtives i fàcils d'usar.

Exemple de cerques per mitjà de cerca amb interfície escrita – distribuïdor DIALOG

Exemple d'una cerca de la base de dades HealthSTAR per mitjà del distribuïdor DIALOG:

```

?begin 151
File 151:HealthSTAR 1975-1998/Jan
(c) format only 1998 The Dialog Corporation
Set  Items  Description
---  -
?select health maintenance organizations(l)trends
      9054  HEALTH MAINTENANCE ORGANIZATIONS/DE
      90441 TRENDS/DE
S1    565   HEALTH MAINTENANCE ORGANIZATIONS(L)TRENDS
?s s1/1997:1998
      565   S1
      223501 PY=1997 : PY=1998
S2    77   S1/1997:1998
?add 149
07jun98 10:38:28 User306002 Session D672.5
$2.75   1.000 DialUnits File151
$2.75   Estimated cost File151
$0.19   TYMNET
$2.94   Estimated cost this search
$9.64   Estimated total session cost  3.016 DialUnits

```

Equació de cerca que indica al sistema els termes que s'han de cercar en el camp descriptors (DE).

Exemple d'una cerca de la base de dades NTIS per mitjà del distribuïdor DIALOG:

```

File 6:NTIS 64-1997/Nov W4
Comp& distr 1997 NTIS, Intl Copyright All Rights Re
Set  Items  Description
---  -
?ss ozone or chlorofluorcarbon? or cfc?
S1   8656  OZONE
S2   8     CHLOROFLUORCARBON?
S3   1117  CFC?
S4   9446  OZONE OR CHLOROFLUORCARBON? OR CFC?
?ss warming or greenhouse (w) effect
S5   1755  WARMING
S6   3893  GREENHOUSE
S7   140800 EFFECT
S8   1938  GREENHOUSE (W) EFFECT
S9   3133  WARMING OR GREENHOUSE (W) EFFECT
?s s4 and s9
      9446  S4
      3133  S9
S10  469   S4 AND S9

```

Font: <http://support.dialog.com/searchaids/success/success.pdf>

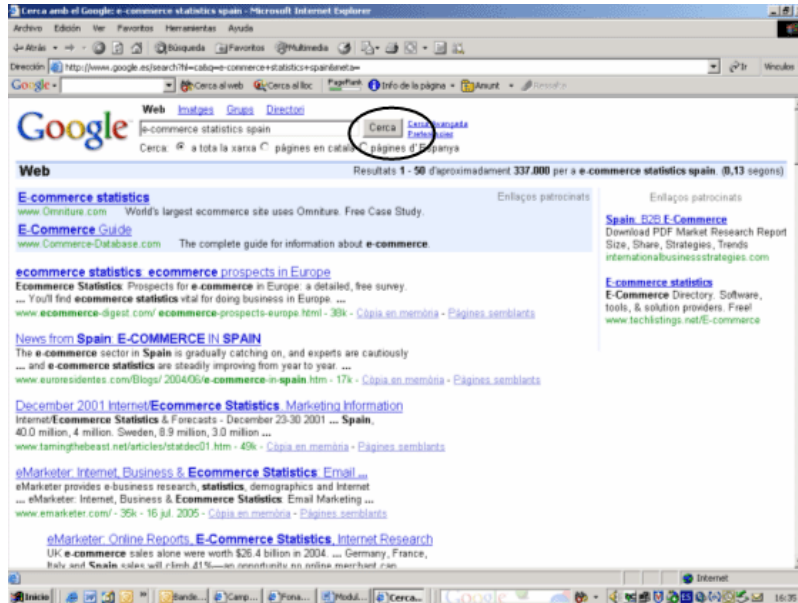
El desenvolupament de les interfícies gràfiques va permetre la irrupció d'una nova generació de llenguatges d'interrogació molt més intuitius que incorporaven sistemes de botons per a indicar les ordres, menús desplegable, caselles de verificació, botons d'opció...

Cada cercador, cada base de dades, en definitiva, cada sistema de recuperació de la informació, pot tenir el seu propi llenguatge d'interrogació. Tanmateix, els llenguatges d'interrogació es fonamenten en unes normes lògiques més o menys comunes, tot i que adaptades a les particularitats de cada sistema, que els regeixen.

Desgranant la definició inicial de Tramullas, observem que l'autor fa referència a les ordres com el primer ingredient del llenguatge d'interrogació. Les **ordres** són aquelles paraules, abreviatures, botons en les interfícies gràfiques, que indiquen al sistema les accions que ha d'executar: buscar, ordenar, imprimir, filtrar... No tots els llenguatges d'interrogació, però, tenen les mateixes ordres, depèn de la riquesa de cada llenguatge, ni són expressades de la mateixa manera.

Ordres

En els exemples anteriors, l'ordre de buscar a Google es dona prement sobre el botó "Cerca",



mentre que a Dialog cal escriure el comandament SELECT.

```
File 6:NIIS 64-1997/Nov W4
Comp&dist 1997 NIIS, Intl Copyright All Rights Re

Set Items Description
---
?ss ozone or chlorofluorcarbon? or cfc?
S1 8656 OZONE
S2 8 CHLOROFLUORCARBON?
S3 1117 CFC?
S4 9446 OZONE OR CHLOROFLUORCARBON? OR CFC?
?ss warming or greenhouse(w) effect
S5 1755 WARMING
S6 3893 GREENHOUSE
S7 140800 EFFECT
S8 1938 GREENHOUSE (W) EFFECT
S9 3133 WARMING OR GREENHOUSE (W) EFFECT
?s s4 and s9
9446 S4
3133 S9
S10 469 S4 AND S9
```

S4 indica buscar els continguts de la línia de cerca 4, on S correspon a l'abreviació de l'ordre SELECT.

El segon element que apareix en la definició de Tramullas són els operadors. Els **operadors** determinen les relacions entre els termes inclosos a les equacions de cerca. La majoria de sistemes reconeixen, per exemple, els operadors booleans, i els més avançats, també els posicionals, de rang, truncament...

Les normes lògiques estableixen els paràmetres bàsics de funcionament, determinant, per exemple, l'ordre dels elements en les equacions de cerca, les possibilitats de combinació dels termes de la cerca, etc.

Operadors

A continuació es mostren com influeixen les normes lògiques internes dels llenguatges d'interrogació en els resultats que s'obtenen. Els exemples estan presos de la base de dades ISTA de la plataforma EBSCOHOST.

En la primera imatge, s'han introduït els parèntesis per a indicar al llenguatge documental com ha de realitzar la combinació de termes. En concret, l'ús dels parèntesis estableix que l'operador OR ha d'unir els termes *semantic web* i *ontolog**. Si no s'haguessin posat els parèntesis, els sistema hauria hagut de decidir per ell mateix si *ontolog** s'havia de combinar amb el terme *semantic web* amb l'operador OR, o bé si ho feia amb l'AND amb el terme *librar**. Per a prendre aquesta decisió, el llenguatge d'interrogació té predefinit un ordre lògic per a l'execució dels operadors.

The screenshot shows the EBSCOhost search interface. The search query is entered as: `(semantic web OR ontolog*) AND librar*`. The search is performed in the Information Science & Technology Abstracts (ISTA) database. The results table is as follows:

Añadir a la búsqueda	Búsqueda n°	Términos de la búsqueda	Limitadores y Ampliadores	Último acceso realizado a través de	Resultados	Modificar búsqueda	Eliminar
<input type="checkbox"/>	S1	((semantic web OR ontolog*) AND librar*)		Interfaz - EBSCOhost Pantalla de búsqueda - Advanced Base de datos - Information Science & Technology Abstracts (ISTA)	56	Modificar	X

En la imatge següent, es mostra el resultat que s'obté si no s'usen els parèntesis (els que apareixen els posa per defecte el sistema). El nombre de documents recuperats és considerablement més elevat, fet que indica que el sistema ha donat prioritat a l'operador OR. És a dir, el sistema primer ha buscat els documents en què apareixien simultàniament els termes *ontolog** i *librar**, als quals s'hi ha sumat els documents que contenien el terme *semantic web*.

The screenshot shows the EBSCOhost search interface. The search query is entered as: `semantic web OR ontolog* AND librar*`. The search is performed in the Information Science & Technology Abstracts (ISTA) database. The results table is as follows:

Añadir a la búsqueda	Búsqueda n°	Términos de la búsqueda	Limitadores y Ampliadores	Último acceso realizado a través de	Resultados	Modificar búsqueda	Eliminar
<input type="checkbox"/>	S2	(semantic web OR ontolog* AND librar*)		Interfaz - EBSCOhost Pantalla de búsqueda - Advanced Base de datos - Information Science & Technology Abstracts (ISTA)	111	Modificar	X
<input type="checkbox"/>	S1	((semantic web OR ontolog*) AND librar*)		Interfaz - EBSCOhost Pantalla de búsqueda - Advanced Base de datos - Information Science & Technology Abstracts (ISTA)	56	Modificar	X

2. Operadors

Els llenguatges d'interrogació més potents reconeixen un ampli ventall d'operadors, que, a grans trets, es poden agrupar en aquestes quatre tipologies:

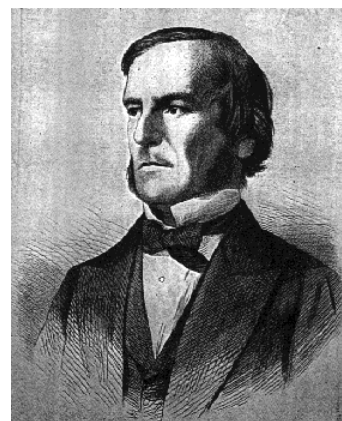
- Operadors booleans o lògics.
- Operadors posicionals.
- Operadors de truncament.
- Operadors de rang.

2.1. Operadors booleans o lògics

Els operadors booleans és la tipologia més estesa i reconeguda per la majoria de llenguatges d'interrogació.

Aquests operadors es fonamenten en el principi que les relacions entre conceptes es poden expressar com a relacions entre conjunts. Les relacions entre dos conjunts de conceptes (que permeten identificar documents) donen com a resultat un altre conjunt (de documents).

Els tres operadors booleans són: l'operador **AND** (operador d'intersecció), l'operador **OR** (operador de suma) i l'operador **NOT** (operador de negació).

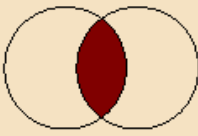
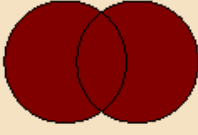
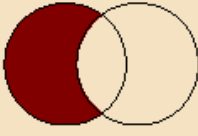


Operadors booleans
Els operadors booleans reben el nom en honor al matemàtic del segle XIX G. Boole, iniciador de l'àlgebra de conjunts.

Operadors booleans o lògics			
	Expressions	Funció	Exemples
AND	AND, Y, I, *, +, un espai en blanc...	Permet recuperar els documents que tenen en comú, almenys, els termes a partir dels quals se cerca. Es tracta d'un operador de caràcter restrictiu que elimina els documents que no contenen tots els termes units per l'operador	Tsunami AND Indonèsia Marsupials AND Austràlia Publicitat AND Creativitat AND Catalunya
OR	OR, O, +...	Permet recuperar els documents que tenen en comú, almenys, un dels termes a partir dels quals se cerca. Aquest és un operador d'ampliació. Molt útil per a la cerca de termes sinònims i quasi sinònims	Televisió OR Ràdio Paludisme OR Malària Compact disc OR CD OR Disc compacte

Operadors booleans o lògics			
NOT	NOT, NO, AND NOT, -...	Permet recuperar els documents que contenen el primer terme de l'equació de cerca, però no el segon. És un operador de caràcter restrictiu	Cinema NOT Hollywood Java NOT cafè Televisió NOT Càmera

La representació gràfica, mitjançant diagrames de Venn, del funcionament dels tres operadors es pot observar en el quadre següent:

Diagrama de Venn	Exemples	Funció
	Tsunami AND Indonèsia	Recupera documents que continguin el dos termes
	Paludisme OR Malària	Recupera documents que continguin, almenys, un dels dos termes
	Cinema NOT Hollywood	Recupera documents que continguin el terme <i>cinema</i> , però no <i>Hollywood</i>

2.1.1. L'ús de diferents operadors

En la construcció de les equacions de cerca es pot usar més d'un operador, aquesta mesura és molt habitual en cerques mínimament complexes. En aquestes circumstàncies, cal tenir en compte les prioritats del sistema a l'hora d'executar els diferents operadors, ja que el fet de processar primer un operador o un altre pot condicionar els resultats que s'obtinguin.

La seqüència a partir de la qual cada operador és processat depèn de cada sistema. Habitualment, però, els sistemes solen donar prioritat a l'operador AND, per damunt dels operadors OR i NOT. Enfront d'equacions que comparteixen operadors amb la mateixa prioritat, el sistema executa el primer operador de l'esquerra.

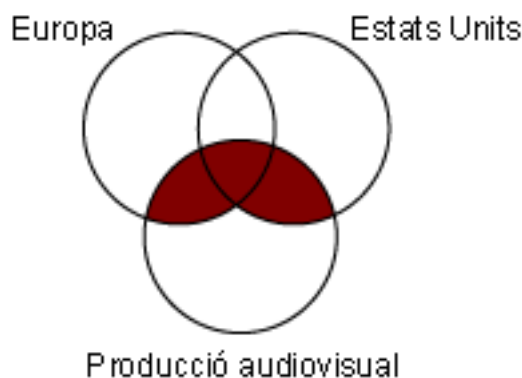
Producció audiovisual AND Europa OR Estats Units

Els sistema executarà primer l'operador AND i a continuació l'operador OR. Així, recuperarà documents que tractin de la producció audiovisual a Europa i tots aquells documents sobre els Estats Units, des de qualsevol punt de vista (història, cultura, cinema, literatura...).

La major part de sistemes, però, permeten modificar l'ordre d'execució dels operadors. Habitualment el símbol que s'utilitza és el parèntesi. Davant un parèntesi, el sistema prioritza la combinació dels termes inclosos en el parèntesi.

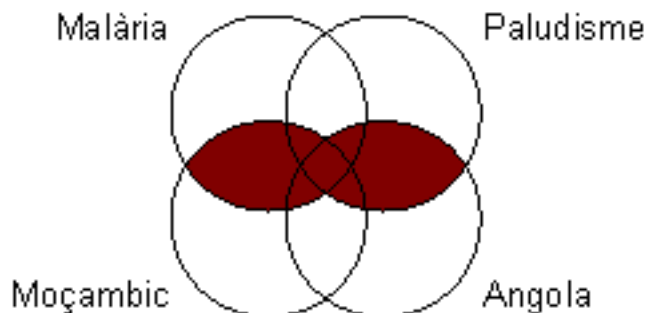
Producció audiovisual AND (Europa OR Estats Units)

En aquest cas, els parèntesis donen prioritat a l'execució del seu contingut, que posteriorment es combinarà amb el terme producció audiovisual mitjançant l'operador AND. Els resultats que s'obtidran són els documents que tracten de la producció audiovisual a Europa i els que ho fan sobre la producció audiovisual als Estats Units. També recuperarà aquells documents que facin referència a la producció audiovisual a les dues àrees simultàniament.



(Paludisme OR malària) AND (Moçambic OR Angola)

Recuperarà els documents que parlin del paludisme, tant si la malaltia apareix identificada amb aquest nom com si ho fa pel seu sinònim *malària*, a Moçambic o a Angola, o a tots dos alhora.

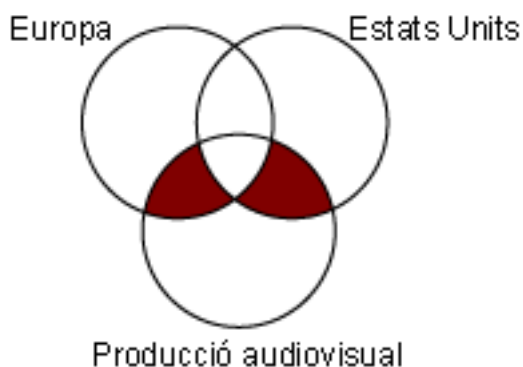


2.1.2. Operador XOR

S'ha vist com l'operador OR recupera els documents que contenen almenys un dels dos termes de la cerca. Alguns sistemes ofereixen un operador similar a l'OR, però amb un caràcter més restrictiu. Habitualment es representa per l'expressió XOR i permet recuperar els documents que contenen qualsevol dels dos termes de la cerca, però no aquells que contenen els dos termes simultàniament.



Gòtic XOR Romànic



Producció audiovisual AND (Europa XOR Estats Units)

2.2. Operadors posicionals

Els operadors posicionals permeten buscar un terme basant-se en el context en què es troba, és a dir, segons la seva posició en el registre o respecte a altres termes del mateix registre.

Els operadors posicionals es poden dividir en dos blocs:

- Operadors posicionals absoluts.
- Operadors posicionals relatius.

2.2.1. Operadors posicionals absoluts

Els operadors posicionals absoluts permeten especificar el camp on el sistema ha de buscar el terme o els termes buscats.

Permeten limitar la cerca en un àmbit concret i és especialment útil en cerques a bases de dades bibliogràfiques. Així, si es busca un autor determinat es pot delimitar la cerca al camp autor, descartant tots aquells documents que continguin el terme en un camp diferent. Per exemple, Autor = Mann, Thomas; Títol = No logo; Matèria: Bioinformàtica.

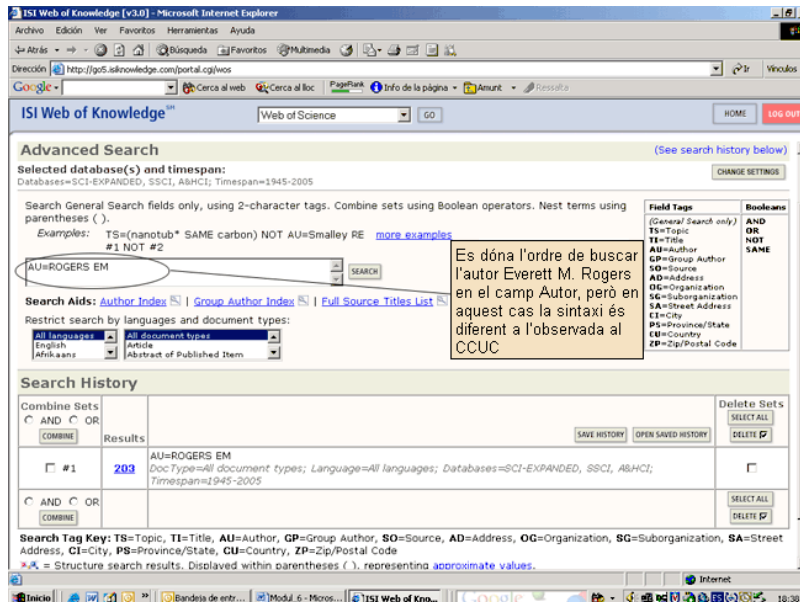
Exemples d'operadors posicionals absoluts

La sintaxi que s'ha d'usar depèn de cada sistema.

Cerca per camp autor al Catàleg Col·lectiu de les Universitats de Catalunya:



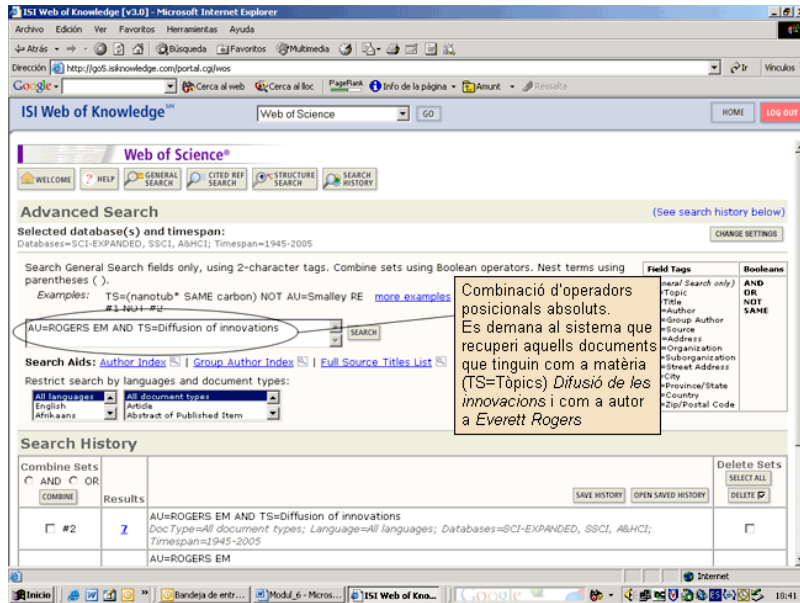
Cerca per camp autor a la base de dades Web of Science:



Els operadors posicionals absoluts, com els altres, es poden combinar amb els operadors booleans. Per exemple: Autor = McQuaiL AND Títol = Teoria de la comunicació.

Combinació d'operadors posicionals absoluts amb operadors booleans

Ús de diversos operadors absoluts en una cerca a la base de dades Web of Science:



Com s'ha dit més amunt, l'ús d'operadors posicionals absoluts permet delimitar els resultats a un camp concret, reduint la possibilitat d'obtenir soroll documental. A continuació es mostren les conseqüències derivades de no usar l'operador posicional absolut en la cerca d'articles d'un autor amb el cognom Floss a la base de dades Medline.

Soroll documental

En no indicar en quins camps es vol realitzar la consulta, el sistema busca en tots els camps de la base de dades. Per aquest motiu ens retorna documents que tenen, efectivament, un autor amb el cognom Floss –registres 1 i 17–, però també recupera documents que en realitat no tenen res a veure amb el que es buscava, ja que *floss* apareix en el camp títol (registre 18) i en el camp resum (registre 16). En aquest cas, el problema és degut al fet que en anglès *dental floss* significa *fil dental*, que és la matèria a la qual fan referència els registres 16 i 18.



2.2.2. Operadors posicionals relatius

La tipologia d'operadors posicionals relatius és una tipologia àmplia que acull diversos operadors com els de proximitat, adjacència o de context. Malgrat existir diferències entre cada tipologia, sovint a la literatura especialitzada apareixen tots sota l'etiqueta d'*operadors posicionals* o *de proximitat*.

Els operadors de proximitat permeten buscar un terme en funció de la posició que ocupa respecte a un segon terme. En la fonamentació teòrica d'aquest operador hi ha la idea que la proximitat entre dos termes pot representar una relació més gran entre els conceptes que representen.

Dins de la categoria d'operadors posicionals relatius hi ha una gran diversitat d'operadors específics, que poden variar molt segons cada sistema de recuperació de la informació quant a funció i expressió. La seva presència és molt més limitada que els operadors booleans i sol estar restringida a sistemes de més qualitat com, per exemple, bases de dades comercials.

Operadors posicionals relatius		
Funció	Expressió	Exemple
Recuperar documents que continguin els dos termes l'un al costat de l'altre.	NEAR, ADJ, W, N...	Desert NEAR Sahara
Recuperar documents que continguin els dos termes l'un al costat de l'altre, però delimitant el nombre de paraules entre cada terme.	NEAR n , ADJ n , W n , W/ n , N n ... (n és el nombre màxim de paraules entre cada terme)	Publicitat NEAR3 Catalunya Recuperació ADJ4 informació
Recuperar documents que continguin els dos termes en el mateix paràgraf, frase...	SAME, WITH,	Energia SAME nuclear

Variabilitat segons cada llenguatge d'interrogació

Entre els operadors posicionals relatius hi ha una gran variabilitat segons cada llenguatge d'interrogació. A continuació es mostren les diverses possibilitats que ofereixen tres sistemes diferents. Destaca que malgrat tenir una mateixa funció, la sintaxi d'un mateix operador pot variar en cada llenguatge.

Funció	Expressió	Exemple
EbscoHost		
Near. Recupera documents que continguin les dues paraules indicades amb una distància màxima entre cada mot determinada per n .	N n	Resultats N4 Revisió Recupera: Resultats de la revisió i Revisió dels resultats

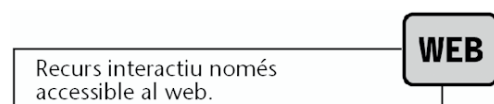
Funció	Expressió	Exemple
Within. Recupera documents que continguin les dues paraules indicades, en el mateix ordre que han estat introduïdes, amb una distància màxima entre cada mot determinada per <i>n</i> .	<i>Wn</i>	Resultats W4 Revisió Recupera: Resultats de la revisió Però no: Revisió dels resultats
Factiva		
Adjn . Recupera documents que continguin les dues paraules indicades, en el mateix ordre que han estat introduïdes, amb una distància màxima entre cada mot determinada per <i>n</i> .	Adjn	Universitat Adj Oberta
W/n . Recupera documents que continguin les dues paraules indicades, en el mateix ordre que han estat introduïdes, amb una distància màxima entre cada mot determinada per <i>n</i> . Funciona igual que <i>Adjn</i> , però <i>n</i> és obligatori.	W/n	Violència W/5 televisió
Nearn . Recupera documents que continguin les dues paraules indicades, en qualsevol ordre, amb una distància màxima entre cada mot determinada per <i>n</i> . <i>N</i> no és obligatori.	Nearn	Metadades Near5 RDF
/nn/ . Recupera documents que continguin les dues paraules indicades, en qualsevol ordre, amb una distància màxima entre cada mot determinada per <i>n</i> . Funciona igual que <i>Nearn</i> , però <i>n</i> és obligatori.	/nn/	Sociologia /n5/ comunicació
/fn/ . Recupera documents que continguin el terme buscat entre els primers <i>n</i> mots del text del document.	/fn/	Difusió de les innovacions/ f50/
Same . Recupera els documents on els dos termes de la cerca es troben en el mateix paràgraf.	Same	Radio same local
Atleastn . Permet especificar el nombre mínim de vegades que el terme buscat ha d'aparèixer en un document perquè aquest sigui recuperat.	Atleastn	Atleast6 UOC
Dialog		

Funció	Expressió	Exemple
Near. Recupera documents que continguin les dues paraules indicades, l'una al costat de l'altra, però en qualsevol ordre.	(N)	Bossa (N) Nova Recupera: Bossa Nova i Nova Bossa
With. Recupera documents que continguin les dues paraules indicades, l'una al costat de l'altra i en el mateix ordre que han estat introduïdes.	(W)	Resultats (W) Revisió Recupera: Bossa Nova Però no: Nova Bossa
nNear. Recupera documents que continguin les dues paraules indicades, l'una al costat de l'altra, però en qualsevol ordre i amb una distància màxima entre cada mot determinada per <i>n</i> .	(nN)	Resultats (4W) Revisió Recupera: Resultats de la revisió i Revisió dels resultats
nWith. Recupera documents que continguin les dues paraules indicades, l'una al costat de l'altra i en el mateix ordre que han estat introduïdes, amb una distància màxima entre cada mot determinada per <i>n</i> .	(nW)	Resultats (4W) Revisió Recupera: Resultats de la revisió Però no: Revisió dels resultats
L. Recupera documents que continguin les dues paraules indicades en el mateix encapçalament/subencapçalament.	(L)	Internet(L)ordinadors Recupera: Internet-Xarxa d'ordinadors
S. Recupera documents que continguin les dues paraules indicades en el mateix subcamp. Un subcamp pot ser un paràgraf.	(S)	Avaluació (S) informació Recupera: Documents que, per exemple, tinguin avaluació i informació en el mateix paràgraf del camp resum.

L'operador de proximitat permet precisar més la cerca que quan s'usa l'operador booleà AND, ja que aquest únicament té en compte la presència o no dels termes, però no la posició que ocupen l'un respecte de l'altre.

Avantatge de l'ús de l'operador NEAR

En l'exemple següent es mostra l'avantatge de l'ús de l'operador NEAR en alguns tipus de cerques. Cal observar la diferència en la quantitat i la qualitat dels resultats obtinguts, segons s'usi NEAR, primera cerca, o, AND, segona. L'exemple s'ha realitzat a la base de dades Factiva:



2.3. Cerca per frase

Malgrat que estrictament no es pugui considerar un operador posicional, la cerca de frase és un recurs que ofereixen molts sistemes de recuperació d'informació, en especial els motors de cerca, que comparteixen una filosofia comuna.

La cerca per frase permet cercar una cadena de caràcters, que poden formar part de més d'una paraula, exactament igual com ha estat introduïda per l'usuari. La representació d'aquest recurs se sol identificar mitjançant l'ús de les cometes ("").

Exemples de cerca per frase

"Universitat Oberta de Catalunya"

"Jocs Olímpics"

2.4. Operadors de truncament

L'ús dels operadors de truncament permet buscar cadenes de caràcter (paraules) a partir d'una mateixa arrel. Habitualment el truncament se sol representar amb un signe *, ? o \$.

Exemple de cerca amb operador de truncament al final de paraula

Biblio*

Recuperarà tots els documents que continguin qualsevol paraula que comenci per l'arrel *biblio*. Per exemple: *Biblioteca, biblioteconomia, bibliografia, bibliofília...*

El truncament se sol aplicar al final de les paraules, però hi ha sistemes que permeten truncar els mots per l'esquerra i entremig. En aquest darrer cas es parla de màscara i freqüentment el símbol que adopta és l'interrogant (?).

Exemples de cerques amb operador de truncament no al final de paraula

***informàtica**

Recuperarà els documents que continguin el terme *informàtica, teleinformàtica, microinformàtica...*

Colo?r

Recupera els documents que incloguin els mots color i colour.

L'ús del truncament és molt útil quan se cerca en sistemes de cerca per llenguatge natural –com és el cas dels motors de cerca–, ja que com que no disposa d'un vocabulari controlat, obliga a buscar totes les possibles variants d'un terme.

Malgrat la seva utilitat, l'ús d'aquests operadors ha de ser curós, ja que un ús indegut pot generar soroll documental. Un exemple típic d'aquesta circumstància es produeix quan el truncament es fa amb una arrel massa genèrica.

Exemples de cerques problemàtiques amb operador de truncament

Cine*

Recuperarà documents que continguin el mot *cinema*, *cinematografia...*, però també *cinetosi*, *cinègètic* o *cinètic*.

Els sistemes més avançats permeten determinar el nombre de caràcters que pot substituir l'operador de truncament. En el cas de la base de dades Factiva, aquesta funció s'identifica amb el símbol \$*n*, en què *n* és el nombre de caràcters que es poden substituir.

Exemples de cerques amb operador de truncament indicant nombre de caràcters

Hospita\$1

Recupera els documents que continguin el terme *Hospital*, *Hospitals*, però no *Hospitales*.

Bibliotec\$5

Recupera *Biblioteca*, *biblioteas*, *biblioteques*, però no recuperarà *biblioteconomia*.

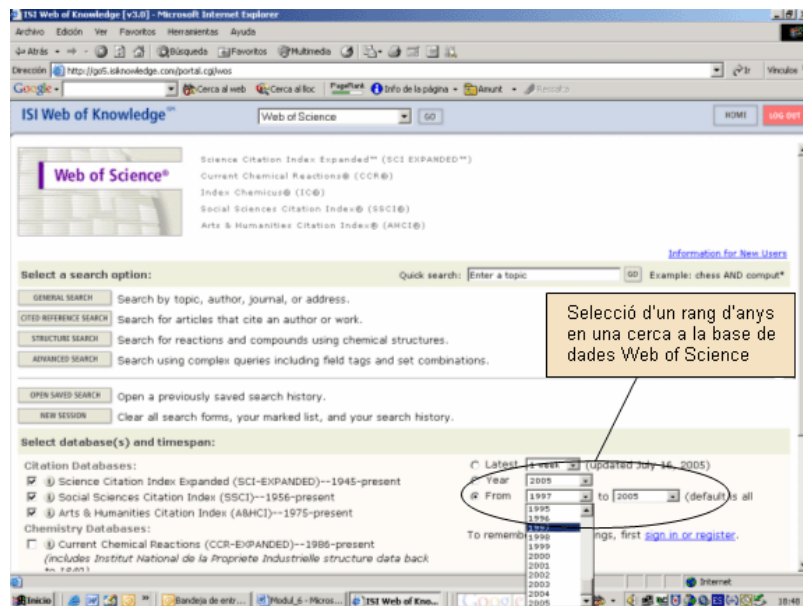
2.5. Operadors de rang

Els operadors de rang tenen un caràcter limitador i se solen aplicar amb informació numèrica. Permeten, per exemple, limitar els documents recuperats a aquells que tenen una determinada data de publicació.

Com en els operadors precedents, els símbols que adopten els operadors de rang varien segons cada sistema. En general, però, aquests són els més freqüents, tot i que en les interfícies gràfiques solen aparèixer en forma de menú desplegable.

Operadors de rang		
Significat	Expressió	Exemple
Igual que	=	= 2000
Major que	>	> 1990
Menor que	<	< 1980
Diferent	<>	<> 2004
Major o igual que	>=	>= 1990
Menor o igual que	<=	<= 1980
Entre	-	1990 - 2005

Exemple de cerca amb operadors de rang



2.6. Altres funcions

2.6.1. Discriminació de majúscules

En l'actualitat la major part de sistemes no discriminen entre majúscules i minúscules, és a dir, independentment de com estigui escrit un terme, el sistema recuperarà tots els documents que el continguin, tant si hi apareix en majúscules com en minúscules.

A vegades, però, pot ser útil poder determinar si la cerca s'ha de restringir als mots en majúscula o en minúscula. Alguns sistemes de recuperació, especialment de motors de cerca, han desenvolupat aquesta funció.

2.6.2. Discriminació dels accents

Com en el cas anterior, permet determinar si els termes que s'han de recuperar han d'estar accentuats o no.

Compte a l'hora de posar l'accent

A Altavista, si es busca la paraula *Manifestacio*, recuperarà els documents amb *Manifestació* i amb *Manifestacio*. Però si el mot s'introdueix al sistema accentuat, limita la cerca als documents que tinguin el terme amb l'accent.

3. Les estratègies de cerca

Quan s'interroguen bases de dades és convenient establir una metodologia de com s'han de fer les cerques.

Quan es realitza una cerca, normalment si és una cerca temàtica, s'han d'establir prèviament quins són els conceptes que es volen buscar i els seus sinònims, seguint les pautes que ja es van estudiar en el mòdul *Introducció a la cerca i recuperació de la informació*.

Una vegada establerts aquests conceptes, és recomanable que es facin cerques diferents per a cada un dels conceptes i després es combinin les diferents línies. D'aquesta manera és més fàcil i ràpid modificar i adaptar l'estratègia quan les cerques no tornin els resultats esperats, sense haver de tornar a escriure de nou totes les equacions de cerca. Quan les cerques són curtes, aquest "estalvi" pot semblar insignificant; tanmateix, convé pensar que de vegades un simple error ortogràfic pot obligar a haver de començar tot el procés de nou.

Exemple d'estratègia de cerca

La manera de procedir seria la següent: si es realitza una cerca sobre **formació d'usuaris en biblioteques públiques de la regió de Múrcia**. El normal seria dividir la cerca en tres subcerques i després combinar-les:

#1 Formació d'usuaris

#2 Biblioteca? Pública?

#3 Múrcia

#4 1 and 2 and 3

D'aquesta manera, si no es troba res sobre aquest tema centrat en la regió de Múrcia, es pot eliminar la línia 3, i ampliar la cerca a formació d'usuaris en biblioteques públiques en general, escrivint una nova línia del tipus:

#5 1 and 2

O fins i tot es podria especificar que aparegués Espanya i afegir-hi unes altres dues línies:

#6 Espanya

#7 #1 and #2 and #6

La raó d'afegir el sostingut es deu al fet que hi ha alguns llenguatges d'interrogació que estableixen aquest signe per a distingir el nombre de línia de cerca d'un nombre normal.

Exemple d'estratègia de cerca

La imatge següent mostra com s'ha desenvolupat una estratègia de cerca a la plataforma ERL WebSpirs:

The screenshot shows the ERLWebSPIRS search interface. The search results table is as follows:

Búsqueda	Resultados	Visualizar	Crear DSI
#9 (#1 adj #2) and ((tumour* or (tumor*)) and (child*))	6		
#8 #1 adj #2	293		
#7 ((tumour* or (tumor*)) and (child*))	18639		
#6 (tumour* or tumor*)	273684		
#5 child*	243966		
#4 tumour*	40730		
#3 tumor*	246542		
#2 therap*	420548		
#1 palliative	7642		

S'observa que les primeres cinc equacions de cerca han consistit a buscar els conceptes de manera independent, truncant-los perquè, a partir de l'arrel d'aquest terme, es trobin tots els derivats.

A continuació, en l'equació #6 s'han combinat amb l'operador OR els dos termes que identifiquen el concepte *tumor*. S'han combinat amb aquest operador perquè són termes sinònims i, per tant, que qualsevol d'aquests aparegui en el document serà suficient per a considerar-lo rellevant.

L'equació de cerca #7 ha combinat les línies de cerca #5 i #6 amb l'operador AND per a indicar que obligatòriament han d'aparèixer al document recuperat el terme *child* i qualsevol dels seus derivats i el terme *tumor* o els seus derivats i sinònim. S'ha d'observar com els parèntesis tenen un paper important en aquesta equació, ja que indiquen al sistema quins termes ha de combinar i amb quin operador booleà.

Seguidament l'equació #8 determina que els termes *palliative* i tots els derivats de *therapy* han d'aparèixer un al costat de l'altre per a recollir el concepte *teràpia pal·liativa*.

Finalment, l'equació #9 uneix les equacions #7 i #8 amb l'operador AND.

En conseqüència, aquesta estratègia trobarà documents que parlin sobre *teràpies pal·liatives de tumors patits per nens*.

Bibliografia

Tramullas, J. (1997). *Introducción a la documática*. Saragossa: Kronos.