

Curs “Els millors recursos d’informació digitals per a la documentació periodística”.

Col·legi Oficial de Bibliotecaris-Documentalistes de Catalunya, Barcelona, març 2009
<http://www.cobdc.org/cursos/2009/curs0709.html>

Bases de dades d’informació d’agències de premsa. Efedata

Javier Guallar
jguallar@gmail.com



¿Qué es EfeData?

EfeData es la base de datos, creada por la [Agencia EFE](#), con el mayor volumen de noticias de información periodística, a la que se puede acceder en cualquier momento.

EfeData actúa como una gran memoria inteligente que mantiene accesible y ordenada toda la información difundida por la Agencia EFE.

El universo de EfeData

Acceso directo a Noticias, Biografías, Organizaciones, Documentos, Efemérides y Información regional.

Acceso directo a todos los sectores y campos: cultura, economía, política, ciencia, deportes, salud, ecología, etc.

Acceso directo a la información Internacional, nacional y regional.

Acceso a EfeData
(Aplicación optimizada con una resolución de 800x600)

<http://www.efeservicios.com>

Descripció

Efedata es defineix com un banc de dades documental gestionat per l'Agència Efe amb el material informatiu elaborat per la pròpia agència.

Ha estat dissenyat per emmagatzemar, mantenir i recuperar la informació generada per l'Agència Efe per a la seva utilització tant per part dels professionals de l'agència com dels clients externs.

Efe és l'agència de premsa més important de l'estat, va néixer el 1939 i en l'actualitat és una de les grans agències a nivell mundial. Efedata va néixer l'any **1988**, convertint-se en aquell moment en la primera base de dades de notícies creada a Espanya. És per tant, el recurs digital d'informació de premsa de més llarga tradició a l'estat espanyol. Des de finals dels 80 va començar a ser utilitzada en els serveis de documentació de premsa dels mitjans espanyols. Durant els primers anys funcionava amb sistemes de programari propi en connexió via mòdem al servidor local fins a finals de **1996**, en què s'inicia l'accés a través d'internet amb el format web actual.

Tecnològicament utilitza el Sistema de Gestió de Bases de Dades documentals PLWeb.

Efedata consta de diversos continguts diferenciats (per tant, diferents bases de dades): Notícies, Biografies, Entitats i Documents temàtics, Efemèrides, Notícies regionals i continguts en català, anglès i portugués.

Notícies. Són les notícies seleccionades diàriament pel departament de documentació a partir de totes les notícies generades pels diferents serveis de l'Agència. La cobertura temporal actual és des de l'inici de 1988. A més, inclou continguts diferenciats de comunitats autònomes, i en el cas de Catalunya, continguts en català. Incorpora al voltant de 800.000 registres cada any.

Biografies de Persones. Conté documents amb informació biogràfica sobre persones a nivell internacional, amb una presència més destacada de l'àmbit espanyol. Conté dades de més de 18.000 personatges.

Biografies d'Entitats. En aquest cas es tracta de documents amb informació sobre institucions, partits polítics, associacions, sindicats, etc.

Documents Temàtics. Són documents confeccionats específicament per a proporcionar una informació cronològica i detallada sobre temes d'interès periodístic. Conté més de 4.500 documents.

Consulta d'Efedata

The screenshot shows the Efedata search interface. At the top, there are checkboxes for various categories: Noticias 2007-09, Noticias 1988-06, Biografías, Entidades y Empresas, Documentos, Efemérides, Noticias Regionales, Catalán, Inglés, and Portugués. Below these are input fields for 'Texto', 'Título o Nombre', 'F. Inicio', and 'Fecha Fin'. There are also dropdown menus for 'Utilizar Derivados', 'Orden por Relevancia', and 'Fechas Rápidas'. A search bar with 'Buscar' and other navigation buttons like 'Relacionados', 'Similares', and 'Diccionario' is present. The results section shows '52 Documentos encontrados: 1-20. Consulta: (ACTUALIDAD *fecha)'. Below this is a table with columns for document ID, title, and relevance percentage.

		Tamaño Relevancia(%)
20090319 000000	DOC SUBASTAS/CUADROS Y OTROS OBJETOS /RECORDS	22729
20090319 000000	DOC LA PENA DE MUERTE EN ESTADOS UNIDOS	15715

La pantalla de cerca d'Efedata té de dues estructures clarament diferenciades: a la part superior es realitza la cerca i a la part inferior es mostren els resultats i els documents.

A la primera línia tenim les diferents bases de dades que conté Efedata:

- **Noticias**, agrupades en 2007-2009 i 1988-2006
- **Biografías**
- **Entidades y Empresas**
- **Documentos Temáticos**
- **Efemérides**
- **Noticias regionales**
- Informacions en altres idiomes: **catalán, inglés, portugués**

This screenshot is identical to the one above, showing the search interface with the same filters and navigation options.

A continuació hi han diverses opcions per introduir elements de cerca.

Texto permet cercar a tots els camps de la base de dades. El sistema permet utilitzar equacions de cerca que incloguin operadors booleans (**AND**, **OR**, **NOT**), d'adjacència (**NEAR**, **ADJ**, **-**, **W**) i comodins (**?**, *****), amés d'altres funcionalitats de restriccions (vegeu més endavant la relació completa dels operadors). L'operador per defecte si no s'indica res és **AND**.

Título o Nombre funciona igual que l'anterior però la cerca es limita al camp TITULO de les bases de dades de **Noticias** o al camp NOMBRE de les **Biografías**.

Fecha Inicio i **Fecha Fin** restringeixen la cerca a un interval de dates amb format AAAAMMDD, sense que sigui necessari omplir tots dos.

A la línia inferior tenim **Fechas rápidas**, un menú desplegable amb dates marcades (Anteayer, 7 días atrás, 1 mes atrás, ...)

Altres opcions són: el botó per activar la cerca (**Buscar**); opcions de ordenar els resultats per **Relevancia** o per **Fecha**; **Relacionados** mostra una llista de paraules relacionades amb el contingut del terme de cerca i **Diccionario** mostra una llista de termes adjunts al de cerca en el diccionari de la base de dades.

L'estructura inferior de la pantalla mostra els resultats de la cerca. A la part superior dels resultats apareix el nombre de documents trobat, la sintaxi de cerca, i en el cas de llistes extenses de documents, opcions per avançar i retrocedir pàgines.

La llista de resultats té cinc columnes: **Data**, **Hora**, **Categoria** (categories temàtiques d'Efe), **Títol** i **Mida** del document.

			Tamaño Relevancia(%)
20090316	173800	SOC UNIVERSIDADES-UPF	2247
20090316	071100	SOC PROTESTAS BOLONIA	1582
20090315	093700	TRI PROTESTAS BOLONIA	1632
20090314	174000	SOC PROTESTAS BOLONIA	2451
20090314	094600	TRI PROTESTAS BOLONIA	1041
20090313	233500	SOC PROTESTAS BOLONIA	3589
20090313	223300	SOC PROTESTAS BOLONIA	2060
20090313	141100	SOC UNIVERSIDAD-BOLONIA	1971
20090313	070900	SOC PROTESTAS BOLONIA	3325
20090312	222600	SOC PROTESTAS BOLONIA	3335
20090312	185700	EDU ESPAÑA-EDUCACIÓN	2567

Al fer clic sobre el títol d'una notícia, obtenim el document, sempre dins de l'estructura inferior de la pantalla, juntament amb uns botons per tornar a la llista de resultats, anar al document anterior o posterior i saltar a la primera posició de la paraula cercada. Dins del text de la notícia apareixen ressaltats els termes de la cerca.

12.03.2009 18:57 utc EDUCACION

ESPAÑA-EDUCACIÓN

Universitarios españoles se movilizan contra el proceso europeo de Bolonia

Madrid, 12 mar (EFE). - Miles de estudiantes universitarios se manifestaron hoy en España contra el proceso europeo de **Bolonia**, que aboga por adaptar las enseñanzas a la demanda social y por la evaluación continua y las tutorías personales en detrimento de las horas lectivas.

El llamado proceso de **Bolonia** es una reforma que afecta a las universidades de la Unión Europea (UE) que tiene como objetivos armonizar los criterios de la educación superior en los países miembro y lograr un incremento del empleo en la UE.

En las **protestas** de Madrid, los manifestantes se enfrentaron con agentes antidisturbios de la Policía Nacional e interrumpieron el tráfico.

Un menor de edad fue detenido como presunto autor de la agresión que recibió uno de los agentes, golpeado en la espalda con un palo.

Los manifestantes -15.000, según los convocantes-, algunos de enseñanzas medias, exigieron un referéndum universitario a nivel nacional sobre el proceso de **Bolonia** y un debate directo, público y televisado de toda la comunidad universitaria con la ministra española de Ciencia e Innovación, Cristina Garmendia.

"Nos manifestamos para que, por una vez de verdad, Garmendia nos plante cara, porque parece que nos tiene miedo, y que debata con nosotros", dijo la portavoz de las asambleas de alumnos Olga Arnaiz.

En Barcelona, los estudiantes ocuparon un edificio de la Universitat Pompeu Fabra.

2.3. Relació d'operadors de cerca

(informació d'Efedata)

AND

Operador booleano 'Y'. Es el operador por defecto. Busca ambas palabras que separa (o todas si hay más de dos).

OR

Operador booleano 'O'. Busca una cualquiera de las palabras que separa.

NOT

Operador booleano 'Y NO'. Busca documentos que contengan la palabra que precede al operador y que no contengan la palabra que le sigue.

NEAR/n

Operador NEAR. Funciona como operador bidireccional de proximidad si se especifica un rango de palabras (p.ej deficit NEAR/3 presupuesto). Si no se especifica rango actúa como el operador adyacencia (ADJ).

ADJ (igual que - ó . ó , ó)

Operador Adyacencia. Busca documentos donde las palabras aparecen consecutivas en ese orden (descartadas las palabras vacías). P.ej. deficit-presupuestario.

W/n

Operador proximidad. Busca pares de palabras en las que la segunda aparezca separada de la primera por n términos como máximo. Ejemplo: anfibio W/5 DNA encuentra documentos en que DNA aparece a menos de 6 palabras de anfibio.

?

Operador comodín para un sólo carácter. P. ej: Generali?a? encontraría Generalitat y Generalidad.

Operador comodín para cadenas Ejemplos: micro* encuentra microscopio, microcomputador; *lacion encuentra relación, translación.

%

Operador búsqueda exacta (no buscar derivados de la palabra) Ejemplo: singular% encuentra singular, pero no singulares o singularmente.

@

Operador sinónimo; reemplaza la palabra en cuestión por su sinónimo en la lista de sinónimos de la base de datos. Ejemplo: satellite@ se cambia por sinónimos de satellite en el thesaurus.

Restricción por campos

:nombre_de_campo

Operador de búsqueda por campos para sólo la palabra anterior a los dos puntos. Ejemplo: europa:título especifica que la palabra europa se va a buscar sólo en el campo título

/F:nombre_de_campo1,nombre_de_campo2

Operador de búsqueda por campos para toda la búsqueda. Ejemplo: kirk douglas/F:nombre,texto especifica que tanto kirk como douglas necesitan estar en el campo texto, o nombre.

nombre_de_campo=número (también >, <, <=, >=)

Operadores numéricos.

P.ej: 19961223<=fecha<=19971201

Exemples de consultes

Documents temàtics

No existeix un llistat per consultar tots els documents temàtics elaborats, biografies o institucions. Però una manera de saber quins documents hi ha d'una temàtica determinada es demanar el tema en el camp Título /nombre de la base de dades de Documents.

Per exemple, podem mirar quins documents ha creat Efedata dels esports més importants. Per exemple, de tenis hi ha una sèrie de textos elaborats amb la història i dades més rellevants de les competicions més destacables informativament: els quatre torneigs grans, (Wimbledon , Austràlia, Roland Garros i l'Obert dels EUA), el trofeu Godó , i la Copa Davis.

Un exemple d'Internacional: El conflicte de l'Irak. Demanem en el camp título documents sobre l'Irak o la guerra del Golf. Obtenim diversos articles molt interessants. Els ordenem per data, ja que el conflicte és llarg i ens interessa mirar primer els més recents. Aquest són:



Ordenar por Fecha	Tamaño	Relevancia(%)
20090310 000000 DOC MUERTOS EN IRAK	4854	100
20090309 000000 DOC IRAK POSGUERRA: ATENTADOS MAS SANGRIENTOS	7312	100
20090128 000000 DOC IRAK POSGUERRA: PAISES CON TROPAS DESPLEGADAS EN IRAK / REPLIEGUE	7212	100
20090114 000000 DOC IRAK: DATOS BASICOS Y EVOLUCION POLITICA	1542	100
20081209 000000 DOC IRAK POSGUERRA / RECONSTRUCCION	4202	100
20071211 000000 DOC CONSEJO DE COOPERACION DE LOS ESTADOS ARABES DEL GOLFO	2002	100
20071015 000000 DOC 4-IRAQ-GUERRA-POSGUERRA: PERIODISTAS MUERTOS Y SECUESTRADOS (MUERTE-SECUESTRO-PRENSA)	1212	100
20070905 000000 DOC SELECCIÓN NACIONAL DE FÚTBOL DE IRAK	672	100
20070625 000000 DOC JUICIO SADAM. DETENCIÓN Y JUICIOS CONTRA SADAM HUSEIN	1322	100
20061102 000000 DOC JUICIO SADAM EL CASO DUYAIL. LA PRIMERA MATANZA MASIVA DEL REGIMEN DE SADAM Y LA QUE LE COSTO EL PRIMER JUICIO	1122	100
20061020 000000 DOC IRAK-POSGUERRA: ATENTADOS E INCIDENTES CONTRA LAS FUERZAS DESPLEGADAS EN IRAK	2512	100
20061019 000000 DOC LA MATANZA DE ANFAL. LA MAS SANGRIENTA DEL REGIMEN DE SADAM	442	100
20060214 000000 DOC IRAK POSGUERRA: TORTURAS EEUU UNIDO ABU GHRAIB	872	100
20060118 000000 DOC IRAK-POSGUERRA: PRINCIPALES SECUESTROS DE IRAQUIES	212	100
20060118 000000 DOC IRAK POSGUERRA: SECUESTROS DE EXTRAJEROS EN IRAK	512	100

Exemple de consulta d' Efemèrides

La base de dades ens mostra les efemèrides de cada dia, amb dades històriques, de naixements i morts. Per exemple, si volem saber les efemèrides del 23 de març:



The screenshot shows the EFE website interface. At the top, there are navigation tabs for 'Noticias 2007-09', 'Noticias 1988-06', 'Biografías', 'Entidades y Empresas', 'Documentos', 'Efemèrides', and 'Noticias Regionales'. The 'Efemèrides' tab is selected. Below the tabs, there is a search bar with 'Texto' and 'Título ó Nombre' fields, and a date field set to '23 marzo'. There are also buttons for 'F. Inicio' and 'Fecha Fin'. Below the search bar, there are buttons for 'Buscar', 'Relacionados', 'Similares', and 'Diccionario'. There is also a link for 'Instalar Certificado' and a menu for 'Ayuda | Sugerenci@s | Consumo | Más Vistos en 3 Días'. The main content area shows the search results for '23.03.2009 00:00 utc: EFM'. The title of the result is 'EFEMERIDES DEL 23 DE MARZO'. Below the title, there is a section for 'Última Actualización 20090323 Accn 177' and 'Clasificación'. The main text of the result is: 'Madrid, 22 mar (EFE-Documentación).- Santoral para mañana, lunes 23 de marzo de 2009: Santos Toribio de Mogrovejo, Fidel, Félix y Rebeca. Día Meteorológico Mundial. El sol sale mañana a las 06,13 y se pone a las 18,30 (gmt). La luna sale mañana a las 04,32 y se pone a las 15,26 (gmt). 1508.- Capitulaciones de Fernando el Católico con Juan Díaz de Solís y Vicente Yáñez Pinzón, por las que éstos se comprometen a buscar un estrecho entre el Atlántico y el Pacífico.'

Madrid, 22 mar (EFE-Documentación).-

Santoral para mañana, lunes 23 de marzo de 2009: Santos Toribio de Mogrovejo, Fidel, Félix y Rebeca. Día Meteorológico Mundial.

El sol sale mañana a las 06,13 y se pone a las 18,30 (gmt).

La luna sale mañana a las 04,32 y se pone a las 15,26 (gmt).

1508.- Capitulaciones de Fernando el Católico con Juan Díaz de Solís y Vicente Yáñez Pinzón, por las que éstos se comprometen a buscar un estrecho entre el Atlántico y el Pacífico.

1534.- Clemente VII declara excomulgado a Enrique VIII de Inglaterra si persiste en divorciarse de Catalina de Aragón.

1748.- Un terremoto convierte en un montón de escombros el castillo de Montesa (Valencia), sede de la orden militar de este nombre.

1766.- Comienza en Madrid el motín de Esquilache, originado por la orden, bajo penas de multa y cárcel, de este ministro de Carlos III de sustituir el sombrero gachó por el de tres picos y la capa larga por la corta para dejar ver mejor el rostro y si se lleva espada bajo la capa.

1848.- Proclamación de la República de San Marcos, en Venecia, que se libera de Austria.

1860.- Batalla de Wad-Ras, que puso fin a la guerra de Africa, ganada por las tropas españolas, dirigidas por O'Donnell, y en la que Prim se cubrió de gloria.

1901.- Estados Unidos entrega a España el precio de la venta de las islas de Joló (Filipinas).

1902.- La Cámara de Italia eleva la edad legal para el trabajo: de 9 a 12 años para los niños y de 11 a 15 para las niñas.

1918.- Proclamación de la independencia de Letonia.

1919.- Se fundan en Milán los "Fasci di combattimento" (fascio de combate), por iniciativa de Benito Mussolini.

1927.- El poeta Antonio Machado es elegido para ocupar un sillón en la Real Academia Española.

- 1933.- El Reichstag aprueba una ley por la que se conceden plenos y excepcionales poderes al Gobierno de Hitler.
- 1945.- Los japoneses realizan una matanza de españoles en Manila.
- 1956.- Pakistán se proclama República Islámica.
- 1959.- Günther Grass publica la novela "El tambor de hojalata".
- 1981.- Golpe de Estado en Bangladesh, que derriba al Gobierno.
- 1987.- Willy Brandt dimite como presidente del Partido Socialdemócrata Alemán. Le sucede Hans Jochen Vogel.
- 1989.- Se legaliza el derecho a la huelga en Hungría, segundo país de la Europa del Este, tras Polonia, que lo hace.
- 1990.- La trapecionista española Pinito del Oro gana el Premio Nacional de Circo 1990, otorgado por vez primera por el Ministerio de Cultura.
- 1992.- Presentación a los medios de comunicación del nuevo tren de alta velocidad (AVE), que recorrió 317 kilómetros y alcanzó una velocidad de 300 kilómetros por hora.
- 1998.- El presidente ruso Boris Yeltsin destituye al Gobierno en bloque y nombra a Serguéi Kirienko primer ministro en funciones.
- "Titanic", del estadounidense James Cameron, gana once Oscar.
- 2000.- Un equipo de investigadores desvela la secuencia completa del genoma de la mosca del vinagre, el más complejo descubierto hasta la fecha.
- 2001.- Rusia hunde en el Pacífico la estación espacial Mir, tras quince años en órbita.
- La Agencia Española del Medicamento autoriza la comercialización de la píldora "del día después".
- 2003.- "Hable con ella" de Pedro Almodóvar consigue el único Oscar al mejor guión original ganado por una película española, un premio destinado habitualmente a filmes de habla inglesa.
- Los eslovenos aprueban en sendos referendos su ingreso en la UE y la OTAN.

(...)

Nacimientos

- 1749.- Pierre Simon Laplace, sabio francés.
- 1814.- Gertrudis Gómez de Avellaneda, poetisa española.
- 1887.- Juan Gris, seudónimo de José González, pintor español.
- 1898.- Ricardo Molinari, escritor y académico argentino.
- 1900.- Erich Fromm, psicoanalista germano-estadounidense.
- 1903.- Alejandro Casona, dramaturgo español.
- 1906.- Joan Crawford, actriz estadounidense de cine.
- 1910.- Akira Kurosawa, director japonés de cine.
- 1931.- Viktor Korchnoi, ajedrecista soviético.
- 1937.- Robert Gallo, médico estadounidense, codescubridor del virus del SIDA.
- 1944.- Michael Nyman, compositor, pianista y director de orquesta británico.
- 1957.- Nieves Herrero, periodista española.
- 1963.- Amparo Larrañaga, actriz española.
- 1963.- Ana Fidelia Quirot, atleta cubana.
- Amparo Larrañaga Merlo, actriz española.
- 1968.- Fernando Hierro, futbolista español.

...

Exemple de cerca a Notícies

Volem saber la xifra exacta d'incendis forestals durant l'estiu passat a Catalunya (sabem que és una informació que es fa pública el mes de setembre)

Demanem a la base de dades de notícies regional Catalunya les paraules clau: incendios , forestales , i acotem per dates el mes de setembre. Obtenim 2 documents, el primer ens dona la dada que cerquem: 1.187 incendis.



The screenshot shows a search interface for a news database. The search criteria are: Text: "incendios forestales", Title: "", Start Date: "20080901", End Date: "20080930". The search results show a document titled "CAT INCENDIOS FORESTALES (Previsión) Verano acaba con un descenso histórico de incendios pero crecen los rescates" dated 19.09.2008 14:46 utc. The document content discusses the number of forest fires and the area burned during the summer, as well as the increase in rescues.

Doc. Número: 1 de 2. Base de datos regional

19.09.2008 14:46 utc CIENCIA

CAT INCENDIOS FORESTALES (Previsión)
Verano acaba con un descenso histórico de incendios pero crecen los rescates

Barcelona, 19 sep (EFE) - El verano que acaba se ha saldado con el número de **incendios forestales** y superficie quemada más bajo de los últimos 20 años -apenas 101,5 hectáreas en 1.187 fuegos-, aunque las salidas de los Bomberos de la Generalitat por rescates y salvamentos ha aumentado un 49%.

El conseller de Interior, Joan Saura, ha presentado hoy en rueda de prensa la campaña de verano, que en su opinión se cierra con "cifras excepcionales" en lo que respecta a **incendios**, ya que se ha quemado un 93% menos de superficie y la mitad de fuegos que en el año 2007.

Junto al positivo balance en **incendios forestales**, la campaña estival arroja una "cifra preocupante", según Saura, en rescates y salvamentos de personas, que han crecido un 49%, especialmente en canales, ríos y pantanos.

Por este motivo, el conseller ha anunciado que la Generalitat enviará a partir de octubre la factura a las personas que tengan que ser rescatadas por tener conductas negligentes, aunque de momento no se les cobrará el importe, y de cara al año que viene impulsará una campaña informativa para pedir prudencia en las actividades de riesgo al aire libre.

De momento, el Departamento no se plantea cobrar por los rescates de montaña, pero "a medio plazo", ha dicho Saura, en los casos de negligencia evidente se podría obligar a pagar parte o toda la factura, que actualmente supera los 2.200 euros por hora.

Concretamente, en una operación de salvamento se invierten 2.271 euros la hora en el caso de intervenir un medio aéreo, 39 euros por hora en los vehículos terrestres, y otros 30 por cada persona que participe en el rescate.

Según la directora general de Prevención, Extinción de **Incendios** y Salvamento, Olga Lanau, la meteorología de este verano, que animaba a hacerse a la montaña, es uno de los factores que explica el aumento de rescates registrado, pero también el desconocimiento del entorno por parte de los excursionistas.

La mayoría de salvamentos, ha aclarado Lanau, ocurre debido a que los rescatados no han tenido en cuenta las condiciones meteorológicas antes de iniciar su excursión, o han salido a la montaña sin el equipamiento adecuado o en unas condiciones físicas insuficientes.