



iSOCO

intelligent software for the networked economy

A Semantic Portal for the International Affairs Sector

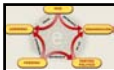
Contreras, **Benjamins**, Blazquez, Losada, Salle,
Sevilla, Navaro, Casillas, Mompo, Paton, Corcho
(iSOCO)

Tena, Martos
(Real Instituto Elcano)



www.esperanto.net
MCYT, PROFIT

EKAW, 2004



The Potential of Semantic Web Technology

- Enable a paradigm switch in searching information
- From
 - Information Retrieval
- To
 - Question Answering
- If you want ...
 - Information Overload on Demand

- This paper illustrates an application in this line
 - For one particular domain
 - Promising, but still a lot of work to do

Forward

iSOCO

intelligent software for the networked economy

© Intelligent Software Components, S.A. 2003

2/27

Real Instituto Elcano
de Estudios Internacionales y Estratégicos

Quiénes somos | Publicaciones | Calendario | Prensa | Recursos

Búsqueda avanzada | Ayuda

4 de Octubre de 2004 549 Analistas del Real Instituto 135 Documentos de Trabajo
902 Publicaciones 224 Materiales de Interés
Última actualización: 4.10.2004 Visitas totales: 1022132 - Hoy: 1629

El Real Instituto Elcano de Estudios Internacionales y Estratégicos es una fundación privada e independiente cuya tarea es servir de foro de análisis y discusión sobre la actualidad internacional, y muy particularmente sobre las relaciones internacionales de España. El Real Instituto Elcano no comparte necesariamente las opiniones manifestadas en los documentos firmados por sus analistas o colaboradores y difundidos en su página web o en cualquier otra publicación.

Materiales destacados

Darfur Letter
Texto de la carta escrita por el equipo directivo

Irak, la tercera ofensiva (ARI)
Manuel Conia 30.9.2004
La sensación de empeoramiento y caos en que parece estar sumido Irak es en gran parte producto de una gran ofensiva, la tercera desde que terminó la guerra, de la informal coalición suní-sadamista-yihadista, orientada a influir en las elecciones americanas

As eleições municipais no Brasil (ARI)
Lourdes Sola 28.9.2004
Prospeção das eleições municipais no Brasil, seguida de uma avaliação das principais tendências. As mudanças decorrentes dos eventuais resultados no cenário pós-eleitoral são indicadas, em termos das condições de governança política, da viabilização da agenda econômica do governo Lula (*ARI em português, próximamente disponible en español)

El idioma español en África subsahariana: aproximación y propuestas (ARI)
Frigidiano Álvaro Durántez Prados 28.9.2004
El español es un idioma prácticamente ausente en África subsahariana, con la excepción formal de la República de Guinea Ecuatorial. Sin embargo, la existencia de cinco países de lengua portuguesa, Estados oficialmente iberohablantes, apunta la oportunidad actual de la promoción del español en un marco iberoafricano

El Marruecos que no despega (ARI)
Haizam Amirah Fernandez 27.9.2004
La llegada al poder de Mohamed VI en 1999 despertó grandes esperanzas dentro y fuera de Marruecos. Cinco años más tarde, el malestar y la decepción crecen a mayor ritmo que las reformas prometidas

El faccionalismo en la política exterior iraní hacia Irak (ARI)
Luciano Zaccara 23.9.2004
Desde la caída de Sadam Husain la política exterior iraní hacia Irak está dando paso a diversas políticas exteriores desarrolladas por diferentes organismos gubernamentales y facciones políticas en el interior del sistema político iraní

Una Constitución para España

Plan Marco Asia - Pacífico

Balance - Perspectivas

ASIA-EUROPE MEETING
Grupo de Sabios de ASEM

Las Políticas culturales en Europa el caso español

Google

Web Images Groups News Froogle more

¿A qué organizaciones pertenece España

Search Advanced Search Preferences

What Organizations is Spain a member of?
"A" is a very common word and was not included in your search. [details]

Web Results 1 - 10 of about 40,400 for ¿A qué organizaciones pertenece España. (0.34 seconds)

Organizaciones Juveniles - [Translate this page]
... y libertades de los extranjeros en España y su ... Xenofobia; Reglamento del Registro de Organizaciones no Gubernamentales ... por la Entidad a la que pertenece, no se ...
www.cdrcampos.es/pjtc/orgjuven.htm - 13k - Cached - Similar pages

campana de ESPAÑA ROJA contra la tauromaquia - [Translate this page]
Quienes deseen unirse, lo mejor que pueden hacer es ... de los Animales) --a la cual pertenece el director ... ESPAÑA ROJA-- oa alguna de las organizaciones hermanas. ...
www.eroj.org/paginas/toros.htm - 10k - Cached - Similar pages

Ministerio de Defensa de España - [Translate this page]
... de ayuda humanitaria y operaciones de paz y de gestión de crisis que realicen las organizaciones internacionales y europeas a las que España pertenece, ...
www.mde.es/multiContenido.jsp?id_nodo=4045& navegacion=T&&keyword=&auditoria=F - 26k - Cached - Similar pages

ESPAÑA Vacios juridicos agravan derrame de combustible - [Translate this page]
... Las organizaciones sindicales subrayaron la necesidad de "exigir soluciones efectivas y ... pero la Unión Europea (UE), a la que pertenece España, asumió una ...
www.tierramenca.net/2002/1124/noticias1.shtml - 23k - Cached - Similar pages

Escuelas de Budismo Tibetano - [Translate this page]
... Nyingma.org Pertenece a Nyingma Centers International ... Buddhism organización budista internacional que pone hincapié ... Mati Dorje Son dos organizaciones de gays ...
inicia.es/de/es_Tibet/religion.html - 32k - Cached - Similar pages

Organizaciones interregionales de ámbito europeo en las que ... - [Translate this page]
... II.- Participación en organizaciones regionales de ámbito europeo. ... basada en la participación de España en estas ... este tipo a los que pertenece la Comunidad ...
www.mundojoven.org/cpre/ue/doc6.htm - 15k - Cached - Similar pages

Excite España - Directorio - Organizaciones - [Translate this page]
... Información de sus actividades, pertenece al distrito 4290. ... y enlace a las organizaciones que trabajan con ... Proyecto que busca remover una barrera al desarrollo ...
www.excite.es/directory/World/Español/ Países/Colombia/Sociedad/Organizaciones - 38k - Cached - Similar pages

Real Instituto Elcano de Estudios Internacionales y Estratégicos

Especial ASEM. Ya disponible Informe final con recomendaciones

Búsqueda avanzada | Ayuda

Quiénes somos | Publicaciones | Calendario | Prensa | Recursos

Inicio: **Buscador**

BUSCADOR
Spain member Organizations
 Resultados para **organizaciones pertenece España** en todos los documentos

ANÁLISIS DEL REAL INSTITUTO (HISTÓRICO)

Atlas mundial de financiación internacional - Iñigo Moré
 3.11.2003
 Según el BIS, al cierre del primer trimestre de 2003, los bancos franceses tenían concedidos créditos a entidades basadas en Iraq por valor de 415 millones de dólares, de 549 millones de dólares a entidades basadas en Cuba y de 2.554 millones de dólares a Irán. Estas cantidades convertían a París en el principal origen de créditos privados internacionales de los ayatolás chitas, Sadam Husein y los compañeros cubanos. En realidad, los bancos franceses también son el primer financiador de buen número de países enemistados con EEUU, tanto presentes como pasados, incluyendo a Vietnam, Somalia, Camboya y el Líbano. Podría pensarse que los banqueros parisinos esperan a que EEUU tenga un tropiezo internacional para sacar su chequera, aunque luego veremos que eso es solo una parte de la historia. Curiosamente, los principales financiadores del resto de países más o menos enemistados con EEUU resultan ser... los bancos alemanes, que lideran la financiación a países como Corea del Norte, Libia y Siria, entre otros.

Globalización y enfermedades infecciosas - David Alonso
 8.5.2003
 En un mundo cada vez más interconectado e interdependiente, el brote de una enfermedad infecciosa en un país realmente podría convertirse en una preocupante situación de emergencia sanitaria para el mundo entero. El siguiente análisis evalúa el riesgo potencial de diversas enfermedades infecciosas emergentes y reemergentes. Dicho análisis ofrece como conclusión una lista de medidas que deberían adoptarse para prevenir, controlar y minimizar esta creciente amenaza.

Elecciones anticipadas en Argentina - Pablo Gerchunoff
 22.7.2002
 De no mediar un preocupante agravamiento de la crisis, el próximo 30 de marzo habrá elecciones presidenciales en Argentina. Ante el descrédito generalizado de los políticos, una cuestión importante es conocer a las fuerzas políticas y a los candidatos que se presentan a la contienda. En este sentido, una fecha clave es la del 24 de noviembre, momento en el que los partidos políticos celebrarían elecciones internas para elegir sus candidatos.

DOCUMENTOS DE TRABAJO

Una política de defensa española para Europa (DT) - Pedro Méndez de Vigo y Montojo

Real Instituto Elcano de Estudios Internacionales y Estratégicos

Español - Inglés

Búsqueda avanzada | Buscador semántico

¿Qué es el Real Instituto Elcano?

Memoria 2002

Estudios e Investigaciones

Barómetro del RIE

Publicaciones del Instituto

Materiales de Interés

Calendario de Actividades

Notas de prensa

Países y Líderes del Mundo

Enlaces

Suscripción al Boletín

Inicio

Inicio: **Buscador Semántico**

BUSCADOR SEMÁNTICO

Búsqueda por temas | Hazme una pregunta

Hazme una pregunta: **¿A qué organizaciones pertenece España**

What Organizations is Spain a memberof?

Frase Entendida: Otcher Organización Internacional Organización de Destino Lugar Pertenece a Organización País y Nombre es España

RESULTADOS DE LA BÚSQUEDA:

Se han encontrado 25 resultados.

Organizaciones Internacionales

Nombre	Corporación De Las Finanzas Internacionales
Abreviatura Organización	IFC
Objetivo Organización	Empresa Privada De La Ayuda De o En El Desarrollo Económico Internacional, Una Agencia Especializada La O.n.u.y Afiliado De Ibrd
Tipo Organización	Organización Internacional
Caracter Organización	Internacional
Acortamientos en los que Participan	
Acuerdos suscritos	
Miembros organización	
Países que la forman	República De Africano Central La India Bosnia Herzegovina

Of what Organizations is Spain a member?

<http://www.cia.gov/cia/publications/factbook/appendix/appendix-b.html>
<http://elcano.mad.isoco.net/rie/analisis/312.asp>
<http://elcano.mad.isoco.net/rie/analisis/265.asp>
<http://elcano.mad.isoco.net/rie/analisis/184.asp>
<http://elcano.mad.isoco.net/rie/documentos/40.asp>
<http://elcano.mad.isoco.net/rie/analisis/186.asp>
<http://elcano.mad.isoco.net/rie/analisis/58.asp>
<http://elcano.mad.isoco.net/rie/analisis/101.asp>
<http://elcano.mad.isoco.net/rie/analisis/106.asp>
<http://elcano.mad.isoco.net/rie/analisis/141.asp>
<http://elcano.mad.isoco.net/rie/analisis/222.asp>
<http://elcano.mad.isoco.net/rie/analisis/25.asp>
<http://elcano.mad.isoco.net/rie/analisis/127.asp>
<http://elcano.mad.isoco.net/rie/analisis/209.asp>
<http://elcano.mad.isoco.net/rie/analisis/49.asp>
<http://elcano.mad.isoco.net/rie/documentos/37.asp>

Nombre	Cámara De Comercio Internacional
Abreviatura Organización	ICC
Objetivo Organización	o Promueve Libre Cambio y La Empresa Privada y Representar Intereses De Negocio En Los Niveles Nacionales e Internacionales
Tipo Organización	Organización Internacional
Carácter Organización	Internacional
Acortamientos en los que Participan	

A Semantic Portal for the International Affairs Sector



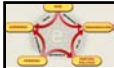
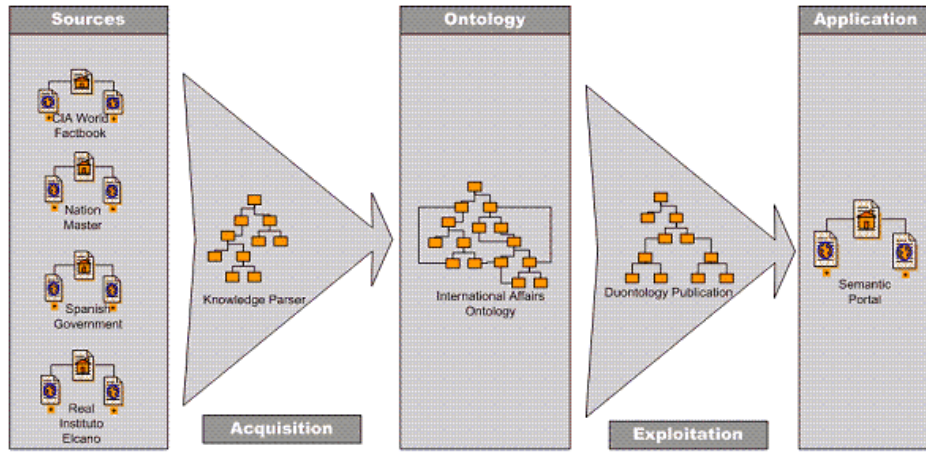
How Does it Work?

- Ontology of International Affairs**
- Knowledge Acquisition**
- Exploitation**
- Conclusions**



The Overall Process

How does it work?



Ingredients

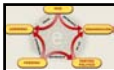
How does it work?

- Multiple, heterogeneous sources
- Ontology
- Knowledge Acquisition “Engine”
 - Knowledge Parser®
- Exploitation of knowledge
 - Publishing the results
 - Duontology®
 - Semantic Navigation
 - Semantic Search Engine
 - NLP queries
 - Enriched documents with ontological information
 - *Smart tags*

A Semantic Portal for the International Affairs Sector



How Does it Work?
Ontology of International Affairs
Knowledge Acquisition
Exploitation
Conclusions



Ontology of International Affairs

Construction

- Constructed in collaboration with experts from the Real Instituto Elcano
- Inspired by CIA World-Fact Book
- Ontology metrics (after population)
 - 60 concepts
 - 124 properties
 - 20.000 instances
 - > 60.000 facts
 - 20Mb in RDF(S) files
- Ontology access and management functionalities built with **KPOntology**.
 - Common API for: JENA1.6, JENA 2.0, Sesame, WebODE



elcano Protégé 2.1 (file:C:\Proyectos\Ongoing_Projects\elcano_mad\elcano\buscador\doc\Ontology\Domain\elcano.pprj, RDF Schema)

Project Edit Window Help

Classes Slots Forms Instances Queries Classes & Instances

Relationship Superclass

THING A
SYSTEM-CLASS A
Entidades
Acontecimientos
Actores A
Organizacion
Diplomatica
Gobierno
Organizacion_Internacional
Organizacion_Terrorista
Partido_Politico
Persona
Acuerdos
Economia
Ejercito
Lugar A
Ciudad
Continente
Edificio
Pais
Region
Relaciones A
Relacion_Acontecimiento_Relacionada_non_Anon
Relacion_Grupo_Ejecuta_Ataque
Relacion_Participa_Acontecimiento A
Relacion_Persona_Ejecuta_Ataque
Relacion_Pertenencia A
Relacion_Suscribir_Acuerdos A
Relacion_Tiene_Gobierno
Relacion_Tiene_Postura
Relacion_Tiene_Rol
Sociedad

Superclasses
Lugar A

PAIS (type=STANDARD-CLASS)

Name: Pais
Documentation:
Role: Concrete

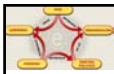
Template Slots

Name	Type	Cardinality	Other Facets
Capital	Instance	single	classes=(Ciudad)
Fronteras	Instance	multiple	classes=(Pais)
Superficie	String	single	
Tiene_Economia	Instance	single	classes=(Economia)
Tiene_Ejercito	Instance	single	classes=(Ejercito)
Tiene_Partidos_Politicos	Instance	multiple	classes=(Partido_Politico)
Tiene_Sociedad	Instance	single	classes=(Sociedad)
Nombre	String	single	
Referencia_Documental	String	multiple	

A Semantic Portal for the International Affairs Sector



Semantic Portal Definition
Ontology: International Affairs
Knowledge Acquisition
Exploitation
Conclusions



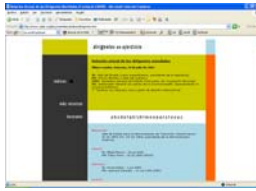
The Sources

Knowledge Acquisition

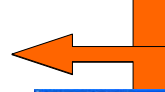
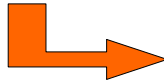
NationMaster



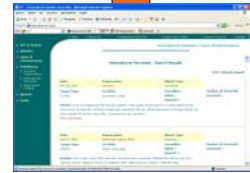
CIDOB



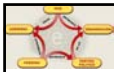
CIA World Factbook



Supervision



International Policy Institute for Counter-Terrorism

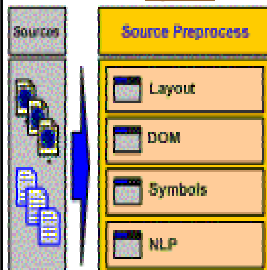


Knowledge Parser® Architecture

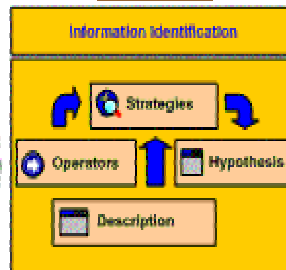
Knowledge Acquisition

- Source Pre-processing
- Information identification
- Ontology population

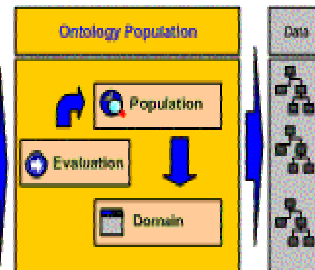
Pre-Processing Types



Pluggable Strategies



Intelligent Population of Ontologies





Linking Ontology to Sources

- In **Ontology**: add link to document where information was found
 - Allows navigation from ontology to source
- In **sources**: add link to ontology concepts
 - Allows navigation from sources to ontology
 - Only for keywords, based on:
 - “Keyword Detection in Natural Languages and DNA”, Ortuno et al, 2002
 - Distances between successive word occurrences



isoco

intelligent software for the networked economy

© Intelligent Software Components, S.A. 2003

17/27

A Semantic Portal for the International Affairs Sector



Semantic Portal Definition
Ontology: International Affairs
Knowledge Acquisition
Exploitation
Conclusions

Selection in Portal

Publishing in a Semantic Portal

Exploitation

Materiales de Interés

Calendario de Actividades

Notas de prensa

Países y Líderes del Mundo

Enlaces

Suscripción al Boletín

Inicio

Economía

Ejército

Acotamientos

Sociedad

Lugar

Continente

Edificio

País

Ciudad

Region

Acuerdos

Actores

Nombre

Capital> Nombre

Fronteras> Nombre

Tiene Economía>

Deuda Externa

Exportaciones

Importaciones

Industrias

Inflacion

Moneda

PIB

PIB Agricultura

PIB Industria

PIB Servicios

Renta Per Capita

Tasa Desempleo

Tiene Ejercito>

Tiene Partidos Politicos>

Tiene Sociedad>

Relaciones

Ninguna Relacion


Relacion Pais Subscribe Acuerdos

Relacion Pais Pertenece Continente

Relacion Lugar Pertenece Organizacion

Relacion Lugar Participa

Relacion Tiene Gobierno



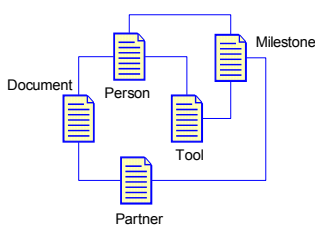
“Traditional” Publishing

Exploitation


Semantic Web Publication: Semantic Portal

Need for SW information publication on WWW (for humans)

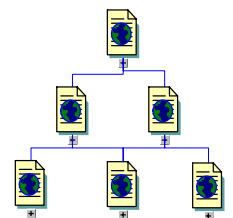
Knowledge Base



publication




Browsable Web Site



Inconveniences of direct publication/translation

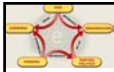
- Semantic model is not necessary user-friendly (relations, control attributes)
- Interface change entails model change
- Model publication is not always desired



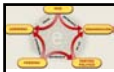
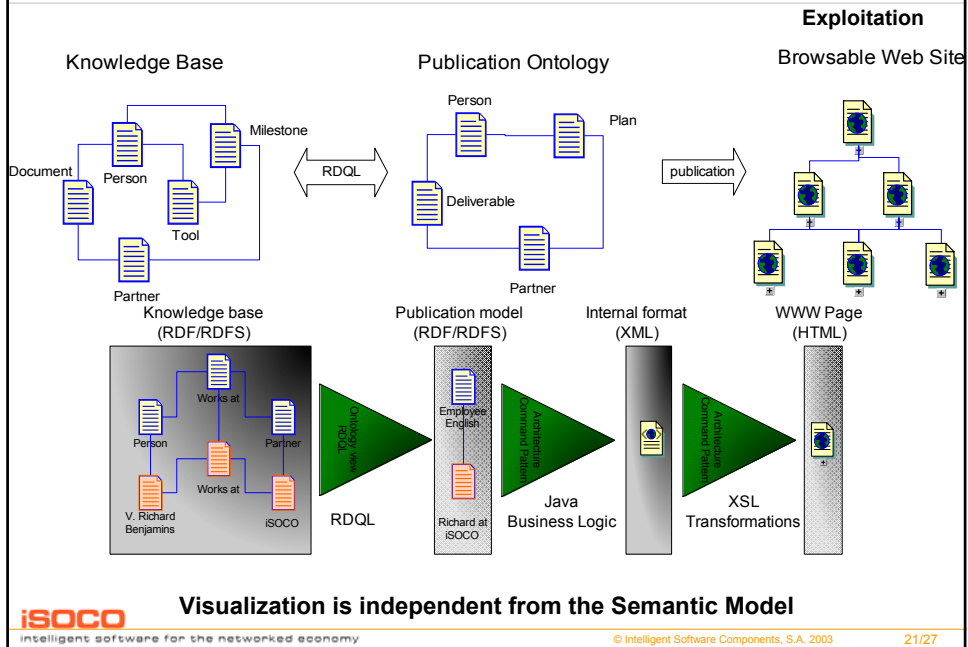
intelligent software for the networked economy

© Intelligent Software Components, S.A. 2003

20/27



Decoupled Publishing



Semantic Search Engine

Exploitation

Advanced Search

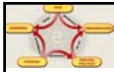


- Searching for data, not for documents (Q/A)
- When *relations* are key
- In addition to keyword-based search engines
- For well-defined domains



Examples

- GDP of Spain
- What countries have participated in the Iraq war?
- To what political party belongs the president of France?
- In which cities Hamas has performed bomb attacks?
- Who is Bush?



- **Example:** “Países fronterizos con Serbia”
- Returns instances
- Allows reference consulting

Pais

Norm: **RESULTADOS DE LA BÚSQUEDA:** Visualizar en 3D

Capi: **Países**

Frontier: *Se han encontrado 7 resultados.*

Tiene Econon

Tiene Ejerci

Tiene Partidos Politic

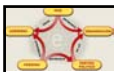
Tiene Sociad

Nombre	Bulgaria
Nombre Capital	Sofia
Superficie	110,910
	Ejército de Bulgaria
	Economía de Bulgaria
	Sociedad de Bulgaria
Acontecimientos Relacionados	
Continente al que Pertenecce	Europa

Relacio

ISOCO
intelligent software for the networked economy

© Intelligent Software Components, S.A. 2003 23/27



- Linking Back to the Sources
 - From results (answers) to documents (sources)

RESULTADOS DE LA BÚSQUEDA: Visualizar en 3D

Se han encontrado 7 resultados.

Países

Nombre	Bulgaria
Nombre Capital	Sofia
Superficie	110,910
	Ejército de Bulgaria
	Economía de Bulgaria
	Sociedad de Bulgaria
Acontecimientos Relacionados	
Continente al que Pertenecce	Europa
Referencias Documentales	http://www.cia.gov/cia/publications/factbk

Introduction Bulgaria Top of Page

Legend: Definition Field Listing Rank Order

Background:

The Bulgars, a Central Asian Turkic tribe, merged with the local Slavic inhabitants in the late 7th century to form the first Bulgarian state. In succeeding centuries, Bulgaria struggled with the Byzantine Empire to assert its place in the Balkans, but by the end of the 14th century the country was overrun by the Ottoman Turks. Bulgaria regained its independence in 1878 but, having fought on the losing side in both World Wars, it fell within the Soviet sphere of influence and became a People's Republic in 1946. Communist domination ended in 1990, when Bulgaria held its first multiparty election since World War II and began the contentious process of moving toward political democracy and a market economy while combating inflation, unemployment, corruption, and crime. Today, reforms and democratization keep Bulgaria on a path toward eventual integration into NATO and the EU - with which it began accession negotiations in 2000.

Geography Bulgaria Top of Page

Location: Southeastern Europe, bordering the Black Sea, between Romania and Turkey

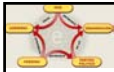
Geographic coordinates: 43° 00' N, 23° 00' E

Map references: Europe

Area: total: 110,910 sq km

ISOCO
intelligent software for the networked economy

© Intelligent Software Components, S.A. 2003 24/27



■ Browsing between ontology and sources

La adquisición de armas nucleares por parte de [redacted] tiene su origen en las abundantes reservas de uranio altamente enriquecido (UAE) que posee el país (es el tercer productor mundial). [redacted] había convertido de hecho, en las postrimerías de la Segunda Guerra mundial, en un importante proveedor de uranio para los ensayos nucleares de [redacted] y [redacted] (en concreto, el proyecto Manhattan).



A Semantic Portal for the International Affairs Sector



Semantic Portal Definition
Ontology: International Affairs
Knowledge Acquisition
Exploitation
Conclusions



- Towards a paradigm switch in searching?
- More work to be done
- Detailed failure analysis needed
- Why does Search Engine fail?
 - KA limitation
 - Not in ontology
 - Missing/wrong instances
 - Query construction
 - NLP result (too complex)
 - SeRQL query construction
- Scalability

BACK UP

CIA - The World Factbook -- Spain - Microsoft Internet Explorer

Archivo Edición Ver Favoritos Herramientas Ayuda

Dirección <http://www.cia.gov/cia/publications/factbook/geos/sp.html>

Google cia world factbook Buscar en la Web PageRank 494 bloqueado(s) Opciones cia world factbook

Geography Spain Top of Page

Location: Southwestern Europe, bordering the Bay of Biscay, Mediterranean Sea, North Atlantic Ocean, and Pyrenees Mountains, southwest of France

Geographic coordinates: 40 00 N, 4 00 W

Map references: Europe

Area:
total: 504,782 sq km
water: 5,240 sq km
note: there are 19 autonomous communities including Balearic Islands and Canary Islands, and three small Spanish possessions off the coast of Morocco - Islas Chafarinas, Penon de Alhucemas, and Penon de Velez de la Gomera
land: 499,542 sq km

Area - comparative: slightly more than twice the size of Oregon

Land boundaries:
total: 1,917.8 km
border countries: Andorra 63.7 km, France 623 km, Gibraltar 1.2 km, Portugal 1,214 km, Morocco (Ceuta) 6.3 km, Morocco (Melilla) 9.6 km

Coastline: 4,964 km

Internet

Detailed Country Profile: Spain - Microsoft Internet Explorer

Archivo Edición Ver Favoritos Herramientas Ayuda

Dirección <http://www.nationmaster.com/country/sp/>

Google Buscar en la Web PageRank 494 bloqueado(s) Opciones

Spain is a country in South Europe and member of the European Union. The main religion is christianity. The main language is Spanish, Catalan and Basque. Spain has a lot of linguistic minorities. In the Basque provinces an active separatist movement is present. In 1037 the Castilla became an independent state, it merged with Aragon in 1479 as Spain. The country is a federal parliamentary democratic monarchy. Inside the federation the communities of Basque and Catalonia have a far going autonomy. The Basque provinces has a strong and violent independence movement. Spain also includes the overseas communities of the Canary Islands and the Balears and the overseas exclaves of Ceuta and Melilla.

Spain's powerful world empire of the 16th and 17th centuries ultimately yielded command of the seas to England. Subsequent failure to embrace the mercantile and industrial revolutions caused the country to fall behind Britain, France, and Germany in economic and political power. Spain remained neutral in World Wars I and II, but suffered through a devastating civil war (1936-39). In the second half of the 20th century, Spain has played a catch-up role in the western international community, it joined the EU in 1986. Continuing concerns are Basque Fatherland and Liberty (ETA) terrorism and further reductions in unemployment.

Borders: [Andorra](#) 63.7 km, [France](#) 623 km, [Gibraltar](#) 1.2 km, [Portugal](#) 1,214 km, [Morocco](#) (Ceuta) 6.3 km, [Morocco](#) (Melilla) 9.6 km

Population: 40,217,413 (July 2003 est.)

GDP per capita: \$21152.52 per person

Top Rankings

- 1st for [Crime - Robberies \(per capita\)](#)
- 1st for [Mortality - Analgesics, antipyretics and anti-inflammatory drugs](#)
- 1st for [Mortality - Atrioventricular and left bundle-branch block \(per capita\)](#)

[... More Top Rankings](#)

Bottom Rankings

- Last for [Mortality - Contact with nonpowered hand tool](#)
- Last for [Mortality - Exposure to smoke, fire and flames, undetermined intent \(per capita\)](#)
- Last for [Mortality - Congenital malformations of trachea and bronchus](#)

[... More Bottom Rankings](#)

Popular articles

- [Spain](#)
- [History of Spain](#)
- [Philip II of Spain](#)
- [Spanish colonization of the Americas](#)

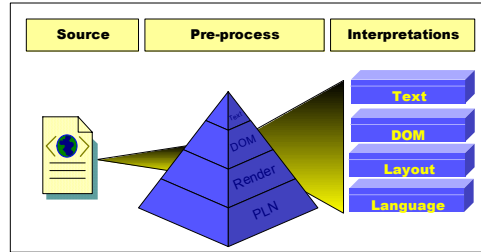
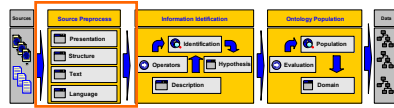
Internet



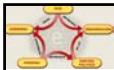
Different Types of Pre-processing

Knowledge Acquisition

- Plain Text Model
 - Regular Expression Check and Retrieval
 - Offset References
- DOM/Hypertext
 - HTML object identification
 - HTTP control and navigation
- NLP
 - Basic NLP: Tokenizer, Morphology and Chunk Parsers
 - Retrieve phrases using head driven approach
 - Basic semantic relations (synonyms, hyponyms, etc.)
- Layout
 - Rendered result of a HTML source: (X,Y) coordinates
 - Visual Operators: SAME_ROW, NEAR, etc...

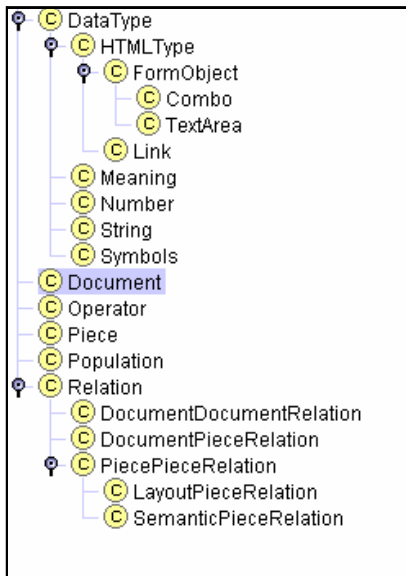


[Back](#)



Explicit Extraction Knowledge

Knowledge Acquisition



Wrapping ontology

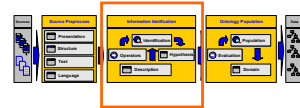
- Documents
- Pieces
- Relations
 - Semantic
 - Layout
- Data Types
 - Meaning
 - Basic Types
 - HTML





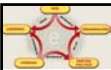
Operators and Strategies

Knowledge Acquisition



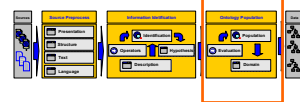
- Operators
 - Check: data types, relations, constraints
 - Retrieve: obtains piece or document (precondition)
 - Execute: navigate, select, etc...
- Strategies (operators applied for hypothesis construction)
 - Greedy: quick but not optimal
 - Heuristics: hypothesis construction and pruning
 - Optimal Backtracking: covering all search space

Back

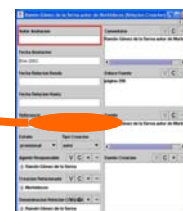
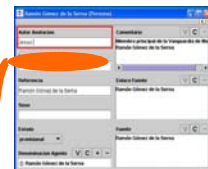


Ontology Population

Knowledge Acquisition



- Actions:
 - Create new instance
 - Modify existing instance
 - Remove existing instance
 - Relate existing instances
- Process:
 - Hypothesis evaluation
 - Population simulation
 - Lowest cost simulation algorithm



Back



Intelligent Software Components, S.A.

For more information on iSOCO, please contact us at

isoco@isoco.com

www.isoco.com

iSOCO Barcelona
Tel +34 93 5677200
C/ Alcalde Barnils 64-68
St. Cugat del Vallès
08190 Barcelona
Spain

iSOCO Madrid
Tel +34 91 3349797
C/Pedro de Valdivia 10
28006 Madrid
Spain

iSOCO Valencia
Tel +34 96 3467143
Oficina 305
C/ Prof. Beltrán Bágüena 4,
46009 Valencia
Spain