

INFOPRAGMATICS:

An Efficient Method for Information Retrieval

UNAM - DGB Rafael Ibarra & Silvia Ballesteros

Quantitative and Qualitative Methods in Libraries International Conference, Chania, Crete, Greece, 26-29 of May 2009.



To: Helena Da Silva



Definition

- Infopragmatics is an efficient method based on linguistic perspectives that allows Spanish Native Speakers to get pertinent – not "relevant" information from commercial academic data bases (*ProQuest, Ebsco, Elsevier*).
 - It consists of three elements:
 - A linguistic storm
 - A linguistic algorithm
 - An uncertainty theorem



Spanish speakers panorama





Information systems issues





Lack of method current effects



relevant records





Infopragmatics issues

lingstorm dictionary/thesauri correct affixation syntax

- warrants entrance to controlled vocabulary
- avoids confusing results (nouns -> verbs -> adjectives)
- syntax observed

<u>semantics</u> technical language synonyms antonyms homophones

manages controlled vocabulary

- enriches results
- negotiates meaning through verbal aids

pragmatics keyword descriptor hypertext hyperlink

opens pragmatic new paths

- profits previous steps
- dynamises external links



Infopragmatics steps





The linguistic algorithm



 This algorithm consists of a series of linguistic operations to solve a variety of problems rised when trying to get *pertinent* information, as opposed to the relevant one.

Lingstorm

a) database name

b) research question

Ideal AND Idea2 AND Idea3



Notes:

- A *linguistic storm*. The lingstorm represents an essential bridge between natural language and controlled vocabulary. It allows the users' lexical availability.
- A billingual dictionary, and an appropriate thesaurus are essential in this phase. They can be digital or printed.



When users finish their fruitless information search, they end up, commonly, frustrated and full of uncertainty. How can they know what they missed or did not do well?



Uncertainty theorem : II.- If the information does not fit is due to...

Key words or descriptors are not such

Inappropriate technical language

Uncertainty theorem

- I.- If there is no information is due to...
 - •User lacks lexical abilities
 - •Syntax is wrong
 - •Used terms are not equivalent in meaning
 - •Context is no the appropriate
 - •There is nothing in the used data base
- II.- If the information does not fit is due to...
 - Inappropriate technical language
 - Incorrect spelling
 - Key words or descriptors are not such

Differences between <u>pertinent</u> and <u>relevant</u> results I...

Original Quest :	Protección <mark>⇒</mark> Radiológica (syntax)	Spa. (Noun + Adj.)	ProQuest Cambridge Medline	EBSCO Health Source: Nursing/Academic: Edition	ELSEVIER ScienceDirect
First try	Radiology Protection	English (Noun + Noun)	1396	1	14,645
Second try	Radiologic Protection (wrong suffix)	English (Adj. +Noun)	5	15,519	4,027
Third try	Radiolog <mark>ical</mark> Protection	English (Adj. + Noun)	1,336	67	16,152

Original Quest :	Rana <mark>78</mark> Verde (syntax)	Spa. (Noun + Adj.)	ProQuest Biological Sciences	EBSCO Academic Search Premier	ELSEVIER ScienceDirect
Nat. Lang.	Frog Green	(Noun+Adj.)	11	1	19,574
Nat. Lang.	Green <mark>F</mark> rog	(Adj. + Noun)	309	64	19,574
Ctrld. Voc.	Lithobates Clamitans	(Noun + Adj)	1	52	13
Alternative	Rana Clamitans	(Noun + Adj)	262	40	550

Differences between <u>pertinent</u> and <u>relevant</u> results II

Original Quest :	Bolsa mexicana de valores	ABI/ INFORM	EBSCO Academic Search Premier	ELSEVIER ScienceDirect
Nat. Lang.	Mexican bag of values	4	1	1,475
Ctrld. Voc.	Mexican stock exchange	610	164	2,727

Original Quest :	Spa. odontología	Cambridge Medline	EBSCO Academic Search Premier	ELSEVIER ScienceDirect	
Nat. Lang.	Eng. odontology	1,893	1,394	13,563	
Ctrld. Voc.	Eng. dentistry	134,846	29,404	108,143	

Lexical availability Specialised dictionaries/thesauri Correct affixation Syntax	 Lingstorm Frequently unavailable Radiolog(y) ic/al Protection Radiological 		pertinent
Technical language	•Guanabana/annona muricata	Database	records
Synonyms	 Green frog /lithobates clamitans/rana clamitans 		displayed results
Antonyms	•Literate / illiterate		 -restrict -enhance
Homophones	 Eyes/ice some/sum 		-read abs.
Keyword	Odontology		
Descriptor	Dentistry	HORITCLEVAR	
Hypertext	• Go 4 it		
Hyperlink	• Go 4 it		
пуреннк	• 00 + 11		

Cases & solutions









The three scenarios: a) current situation; b) alone, but following infopragmatics and c) assisted and following infopragmatics



E-Resources & queries universe

Fecha	echa LIBRUNAM		SERIUNAM	REVISTA-E	TESIUN	AM	CLASE		PERIODICA		HELA	MAPAMEX	Bases de Datos	Recursos Libres	TOTALES
	1.036.317	6.656	53.170	17.325	391.874	91.870	288.963		317.554		3.402	14.261	174	2.692	2.125.732
								21.336							
eng	249.039		9.372	10.272	2.254							198	150	795	272.080
spa	266.046		4.687	189	42.118						2.636	433	23	1.684	317.816
otro/desc	521.232		39.111	6.864	347.502		288.963		317.554		766	13.630	1	213	1.535.836
	Búsquedas	Texto			Búsquedas	Texto									
may 2008	59.096	4.690		57.593											116.689
jun 2008	206.609	4.320		94.590											301.199
jul 2008	54.035	3.019		58.907											112.942
ago 2008	211.627	6.390		108.853											320.480
sep 2008	169.006	7.981	25,190	77.155	65.359		24.547		7.287		7.234				375.778
oct 2008	229.874	5.665	71.730	75.285	161.624		54.727		12.585		5.825				611.650
nov 2008	301.439	4,169	55.215	94.174	164.984	26.864	39.939		8.059		2.243				666.053
dic 2008	154.743	4.347	18.915	58.276	85.544	29.664	19.011		3.489		962				340.940
ene 2009	241.554	7.958	37.008	70.612	144.414	44.985	26.737		5.025		2.092				527.442
feb 2009	383.268	9.696	52.259	67.950	219.378	60.643	41.637		6.998		2.440				773.930
mar 2009	564.958	8.556	64.011	98.524	227.096	68.057	45.595		7.635		2.716	7.765	28.688	14.915	1.061.903
abr 2009	313.980	5.112	24.156	66,180	117.561	33.899	20.439		2.969		1.193	4.928	13.307	7.727	572.440
	2.890.189	71.903	348.484	928.099	1.185.960	264.112	272.632		54.047		24.705	12.693	41.995	22.642	3.942.708
															21.604

Current web page

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO Mapa de Sitio Dirección General de Bibliotecas **Biblioteca Digital Biblioteca Central** Sistema Bibliotecario **Dirección General** English version Periodísticas BiDi UNAM Académica Metabuscadores OA-HERMES Comunidad - a. Nuevo Método Para Producir Proteínas Catálogos con recursos Útiles de acceso libre Recursos electrónicos en evaluación Un nuevo método permitirá la obtención de Colecciones proteínas para uso medicinal o industrial, sin Recupera información Digitales tener que emplear células vivas. Muchas de pertinente en inglés esas proteínas no pueden producirse mediante la biotecnología convencional. Envíanos tus recursos LIBROS Entrega tu tesis - b. Amazon Unveils a Large-Screen Kindle electrónica Aimed at Textbooks and Newspapers REVISTAS Recomendación Acceso Remoto Comonal Actualización: 0:00a.m. - 14/May/09 TESIS. BASES DE DATOS. Recursos por suscripción Reportes Fallas del sitio 4. WILEY RECURSOS LIBRES. Problemas de conexión Wiley ofrece acceso a artículos en texto completo, tablas de contenido y resúmenes, de temática multidisciplinaria. Temáticas: antropología, 2466780 Acceso por token biología, biotecnología, física, medicina, guímica, psicología, entre otras. Sólo UNAM. Consulte a su bibliotecario. Bit UNAM Solicitud de clave 1 2 3 4 5 6 7 8 9 10 11 12 13 14 Siguiente > para acceso remoto a

 ¿Quiénes somos?
 Contacto
 Proyectos
 Acerca de
 Preguntas frecuentes

 Ciudad Universitaria, 04510 México, D.F. Tels. (52-55) 5622 3984, 5622 3976, 5622 3977 y 5622

 4021

Biblioteca Digital





rafael@dgb.unam.mx
silviabe@dgb.unam.mx

http://bidi.unam.mx/ http://dgb.unam.mx/