

Semantic Successes and Challenges for Search Based Applications 24 September 2010

Gregory Grefenstette Chief Science Officer Exalead

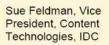


Sue Feldman:

Search Based applications "employ search as a central component, but they solve a specific business problem or support an information-intensive process"

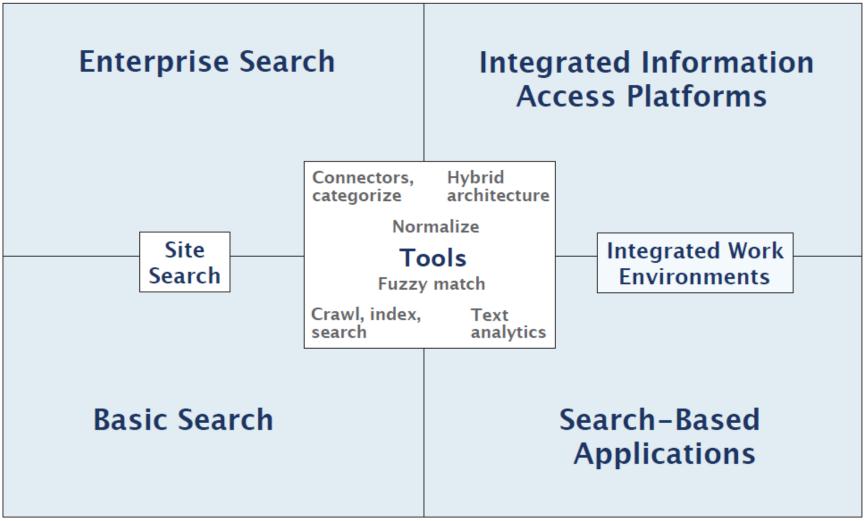
"Search-based applications as a group are growing much faster than enterprise search because their value in accomplishing a task is immediately recognized by business users, because they are quick to deploy, and because they either control costs or increase revenue. Good examples of those are call center support, ediscovery, marketing, sales, R&D research to locate opportunities and e-commerce."







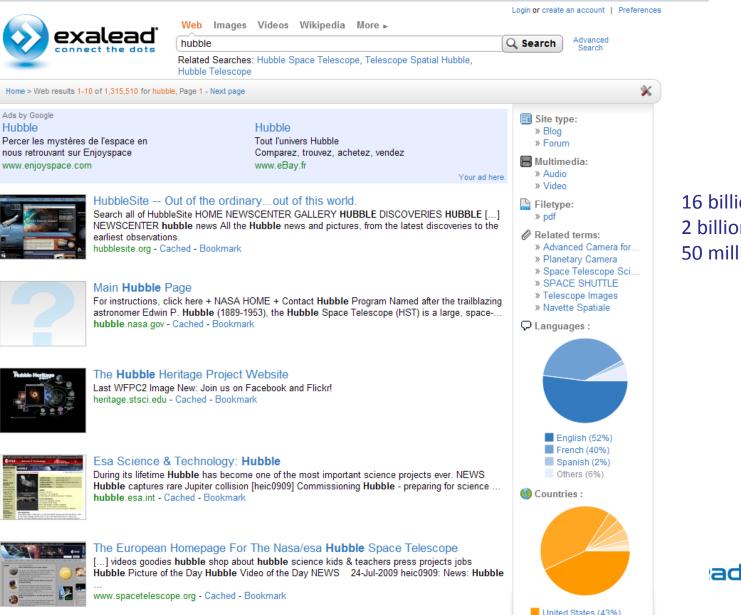
Search and Discovery Software: 2009 Market Map







www.exalead.com/search, labs.exalead.com



16 billion web pages2 billion images50 million videos

Semantics is three things to the ordinary person

Is this an example of that?

Are these two things the same?

What is the relation between these two things?

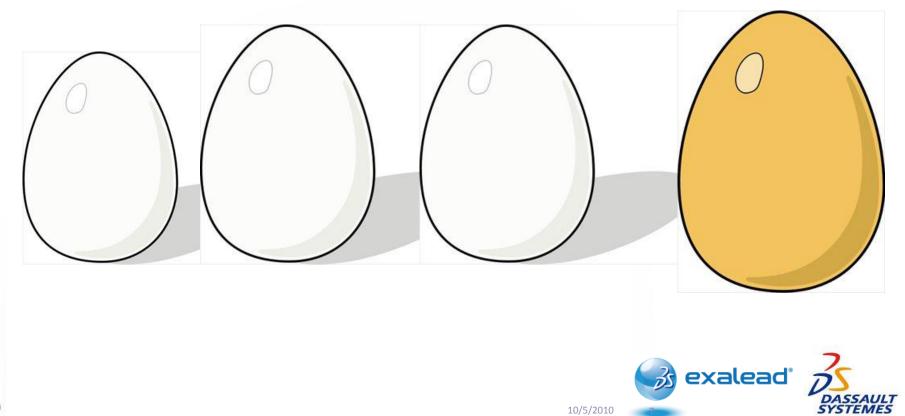


Success 1



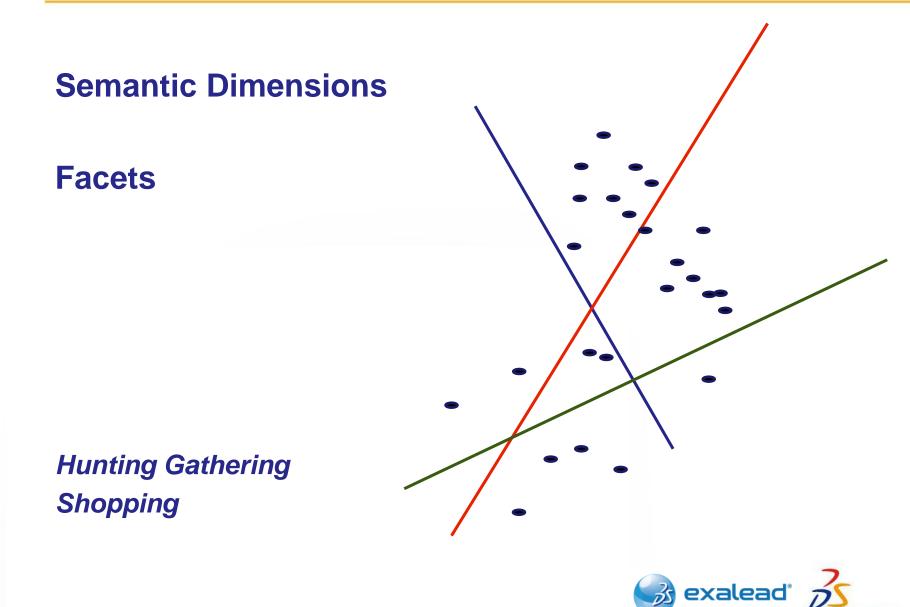


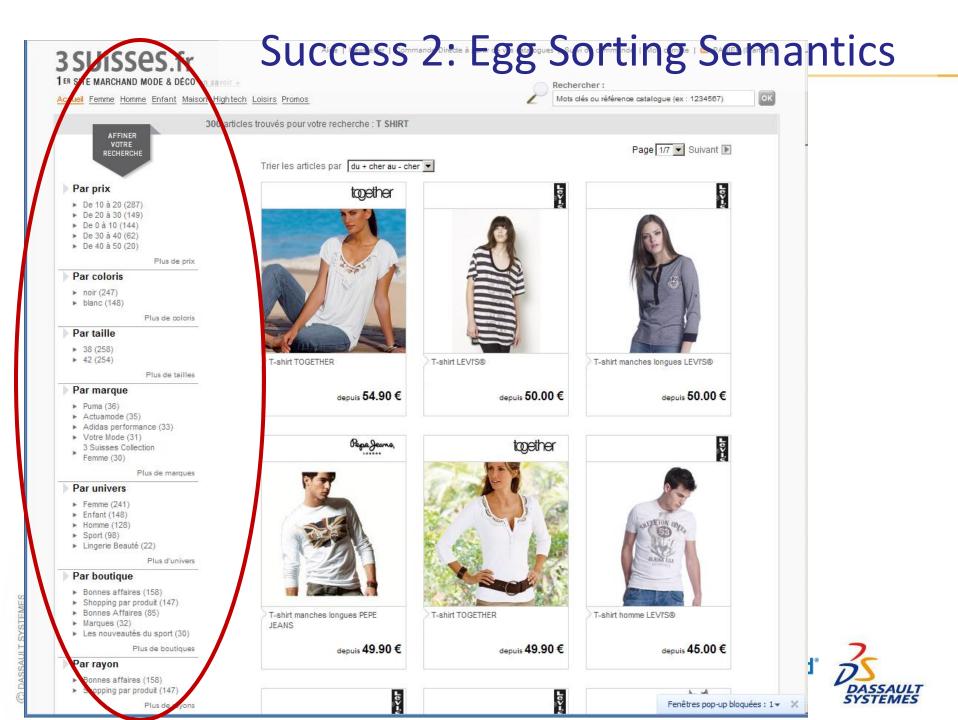












exolead	developing countries	OFFICE\ggrefens <u>Sign out</u> <u>Profile</u> <u>Arteren</u>
Search Results	Results 1-10 of about 840 for developing countries	Refine your search
	Sorf by relevance - Sort by date - Sort by size View: 📃 🛄	Source
		W Vie société (402) excludi
	IDC WW Search Discovery Forecast 2007-2011.pdf The search market is in for another tumultuous year of fast growth. ! Whether this	Source Documentation (221) exclude
	The electron makes is in for all other containable year or lake growth : where the time consolication will be in conjunction with business intelligence applications,	Commercial (81) exclude ex
and the second se	19 May 2008 - 220k	Marketing (36) exclude
- T	1 Bourse: Vie société : Formations Nouveaux Anivants : Resources : Additional Interesting documents	<u>Client</u> (20) <u>exclud</u> <u>exclud</u>
	🐨 Famous people: Susan Feldman, Basel II	Interpretation (222) Interp
	Goographie losation: <u>Oss</u> , <u>Americas</u> , <u>Asis</u> , <u>Speen Breet</u> , <u>URA</u> , <u>United Bistes</u> , <u>Geographic Region</u> , <u>While North</u> Gorganization: <u>Microsoft, IBM, Oracle</u> , <u>Google, Ask</u> , <u>Verees</u> , <u>ALAX</u>	Contrats (93) exclud
		Sales and communication kit exclude
	Research Areas, FP7	(59)
	International collaboration with Developing Nations and Emerging Economies PAGEREF	Image: Second state
-	Toc193707850 'h Editors Further, since the combination of media and 21 Mar 2008 - 656k	Competition (12) exclude
		Sommercial (89) exclud
	Source: Vie société : Documentation : Contrats : Consortium_Projects : z_FP6 : CHORUS Famous people: Federico Noro, Isloro Laso Ballesteros, Petros Daras	喻 <u>KK</u> (66) <u>exclud</u>
	Organization: Europe, Brussels, Novel, Japan Organization: FMI, EU, European Commission, European Union, Information Society, The European Commission, Gartner Group, Enabling New Mobile, Virtual Community	Philippe DELTENRE (19) exclude
		Sylvain CHENE (2) exclud
	SIXTH FRAMEWORK PROGRAMME	
	search engines with the goal of developing an open source prototype of a distributed,	Shared Folders (88) exclud
	semantic-based search goals, developing a system that is both	Sales and communication kit (58) exclude
16-	24 Apr 2007 - 3109k	Sender
Y	Source: Vie société : Documentation : Contrats : Consortium_Projects : z_FP7 : SeenOnTV Francus people: Alma BioInformatica, Karl Aberer, Henry Tirri, Wray Buntine, Alvis Architecture, Adeline Nazarenko	 Ameer Rizvi (2) exclud
	Geographie losation: Paris, Europe, US, Pu, Denmark, France, Finland Gorganization: INRA, EU, Exalead, EPFL, Google, Microsoft, MB, Project Management Committee, Université Paris, European Commission	Carole Offredo (16) exclud Cella Revol (11) exclud
	Organization: INPA, EU, Exaleso, EPFL, Soogle, Microsoft, Me, Project Management Committee, Universite Paris, European Commission	Ella Maria Barragan (4) exclud
		 Florian Douetteau (4) exclud
	2 2 Individual participants (Maximum length efficiency and satisfaction, and developing	Gluseppe Contissa (6) exclud Julien Gledel (23) exclud
	applications and platforms with which ageing people feel more at	Manufacturing News (2) exclud
doc	12 May 2007 - 102k	 gcandela@almacq.com (15) exclud
	2 Source: Vie société : Documentation : Contrats : Consortium_Projects : z_FP7 : SeenOnTV : Admin	 mgiraud@aimacg.com (5) exclud
	So Famous people: Avaia Rahav, Simon van Dam, Daniel Ockeicen, Florian van der Velde, John De Vet, Kees Tuinenbreijer Geographio looation: France, Athens, The Netherlands, Europe, Israel, Greece, Halta	Recipient
	🎲 Organization: IBM, Exalead, Philips, Ibm Halta, EU, Magellan Partners, European Commission, Philips Consumer Electronics, Digital Ty, Carnegle	 dies-alis@atosorigin.com (20) exclud
		francols.lagunas@exalead.com (12) exclud Lpitt@imperial.ac.uk (11) exclud
	untitled Preparing for Digital Natives Joining Your Workforce November 3-7, 2008 access to the	 jobs-fr@exalead.com (22) exclud
	repairing for bright readers downing from reference or 2000 a boost to the internet, which will be nearly fully democratized in most countries.	Iucretia.maiorescu@siveco.ro (11) exclud exclud
	10 Oct 2008 - 752k	 paolo_franceschini@sineura.com exclud (11)
par	🤯 Boures : Vie société : Documentation : Sales_and_communication_kit : Powerpoint_Presentation : Event_presentations : Gariner Symposium Cannes 2008	pelsse-dupichot-zirah@wanadoo.fr exclud
	Sig Famous people: Monica Basso, Marc Prensky, Chris Anderson, Abraham Maslow, Will They Be Goographio location: Cannes, France, Africa	(11) evolution over (11)
	Organization: Gartner, Inc., Gartner, Communities and Social Networks, Disney, Fun Community, LEGO, Diversity it and Mobile Skills Information Retrieval Con- social Networks, Amazon, Apple	 presse@exalead.com (11) exclud r.attlas-dq@ort.asso.fr (11) exclud
	Decent received, interest operations	sartor(Ocirild.unibo.it (11) exclud
	untitled	Document type
	The Big Bang: Demand for Advanced Leadership rather than directing, by developing self-	 Acrobat (.pdf) (503) exclud
247 V	organizations structures rather than reporting structures and by	Word (.doc) (321) exclud HTML (.html) (24) exclud
pdf	17 Nov 2008 - 520k	Rich Text Format (.rtf) (21) exclud
1	Source: Vie société: Documentation: Sales_and_communication_kh: Powerpoint_Presentation: Event_presentations: Gartner Symposium Cannes 2008 Famous people: Diane Morello, Marine Corps, Will Explode	 PowerPoint (.ppt) (19) exclude
	Organization: Gartner, Inc., Gartner, IBN, Zappos.com, America's, NASA, Internal Parties External Partners, Homogeneous Limited	Excel (xis) (18) exclud E-Mail (attachments) (11) exclud
		Text (.txt) (10) exclud
	untitled	E-Mall (6) exclud
	Retirement. In developed countries, retirement will be pushed back, as business activities	MS Open XML Word (.docx) (6) exclud
1 M	move within developed, emerging and developing countries.	Date
pdf	17 Nov 2008 - 622k	2009 exclud 2008 (276) exclud
X	Bourne: Vie societé : Documentation : Bales,and_communication_kt : Powerpoint_Presentation : Event_presentations : Gartner Bymposium Cannes 2008 Famous people: Young Genoves, Mark Trensity	March 2008 exclud
	@ Geographio location: Cannes, France, India, U.S., Applications Building, Berkeley, Belling, Hyderabad	April 2008 exclud
	Organization: <u>Gartner</u> , Inc., <u>Gartner</u> , Intosys, Population Division, <u>United Nations</u> , <u>Duke University</u> , <u>University</u> of California	October 2008 exclud
50°-	NTRODUCTION	2007 (213) exclud exclud exclud
-	rom NIC newly Industrialised countries) Destinations Strongly concentrated (75%) In	April 2007 exclud September 2007 exclud
	territori nengi independinge bertitake) bertitake derengi okterna ale (non) in berlippet economies High in different countries states may inhibit	October 2007 exclud
	24 Mar 2002 - 57k	2005 exclud
and other states and the	🐨 Bourse: Vie société : Commercial : KK : SEXALEAD NYCIE : IDEES : SARCHIVES 2000 : TEMP MBA : SARCHIVES 2000 : GT : MBA : INTERNATIONAL : from First class	Language

Success 3: Semantic Type Extraction

- MUC: named entities
- Affect: emotive content
- ML: classification, clustering

The proposed merger between Mega, Inc. and CNA Systems, Incorporated, has been postponed, Mega CEO Joe Smith said in an analyst call. "CNA's 1st quarter revenue dropped by 32%, and they lost 23 million dollars," Smith explained. CNA Systems sources blame weak sales in China. CNA shares (CNAI) fell 47 percent to \$9.84 on May 12, the first trading day after the announcement.

Company	Mega, Inc., CNA Systems, Incorporated							
Date	May 12							
Person	Joe Smith							
Person Position	Mega CEO							
Currency	23 million dollars, \$9.84							
Measurement	32%, 47 percent							
Country	China							
Concept	proposed merger, analyst call, 1st quarter revenue weak sales, first trading day							
Event: M&A	The proposed merger between Mega, Inc. and CNA Systems, Inc. has been postponed							





Web My PC Images Wikipedia Video More »

"sea turtle"

Sea turtle

Wikipedia

Results 1-10 of about 332 for "sea turtle"



Sign in | Preferences

in your browser

Sea turtles (Superfamily Chelonioidea) are turtles found in all the world's oceans except the Arctic Ocean. There are ... the body plan of sea turtles has remained ...

en.wikipedia.org/wiki/Sea turtle - Add to shortcuts

Categories: Sea turtles, Marine reptiles See also: 💩 People: Pope John Paul II, Marine Laboratory, Marine Mammal Center

Wikipedia Search

Advanced search



Loggerhead Sea Turtle

As with other sea turtles, females return to lay their eggs on or near the same beach where they hatched. ... 1758 url http://www.seaturtle.org/mtn/archives/mtn108 ... en.wikipedia.org/wiki/Loggerhead Sea Turtle - Add to shortcuts

Categories: Sea turtles, Reptiles of Australia, more » See also: 💩 People: June Haimoff, Marine Sanctuary



Sea Turtle Association of Japan, Kuroshima Research Station

research into the nesting of sea turtles, especially confirmation of the nesting of the hawksbill turtle for the ... Specimens of corals, sea turtles, and sea shells ... en.wikipedia.org/wiki/Sea Turtle Association of Japan, Kuroshima Research Station - Add to shortcuts

Category: Laboratories See also: 🚵 People: Kuroshima Research Station



ARCHELON, the Sea Turtle Protection Society of Greece

founded in 1983. It aims to protect sea turtles and their habitats in Greece through research, public awareness campaigns, ... References Archelon's official website en.wikipedia.org/wiki/ARCHELON, the Sea Turtle Protection Society of Greece - Add to shortcuts

Category: Conservation in Greece



Ridley sea turtle

Ridlev Sea Turtles (Lepidochelys) are a genus of sea turtles with two species; the Kemp's Ridley and the Olive Ridley. ... Appearance Adult Ridley Sea Turtles grow to a ... en.wikipedia.org/wiki/Ridley sea turtle - Add to shortcuts

Category: Sea turtles See also: 💩 People: Ridley Sea Turtles, Olive Ridley, more »



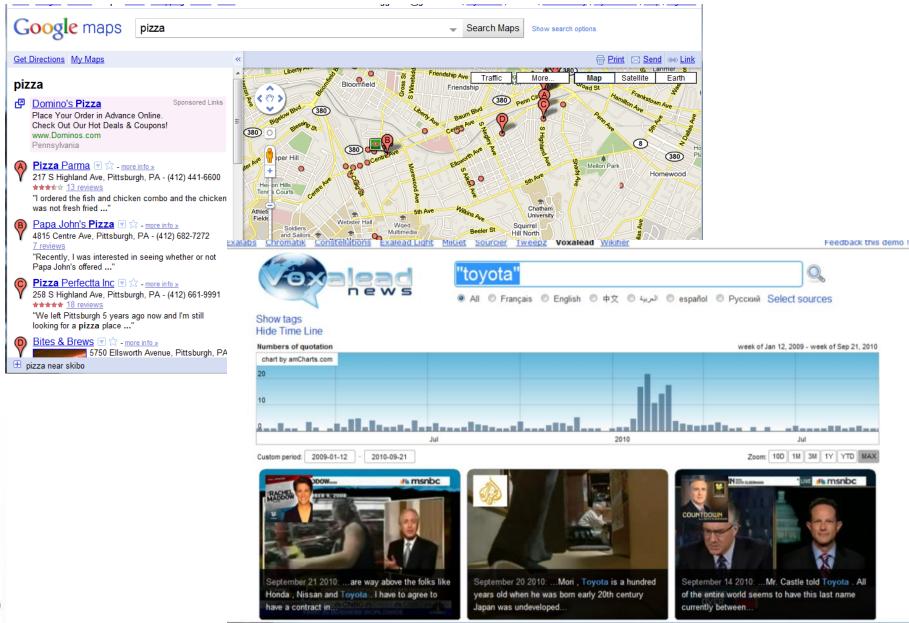
Dermochelys coriacea Endangered fauna of Australia Eretmochelys imbricata Familia Cheloniidae Fauna of Bangladesh Fauna of Delaware and Maryland Fauna of Guyana Fauna of India Fauna of Southeast Asia Habitat Hawaii Hawksbill turtles

Indonesia IUCN June Haimoff Lepidochelys kempii Loggerhead turtle Malaysia Marine Laboratory Marine reptiles Marine Sanctuary Marine turtles Mexico Natator depressus NOAA Olive Ridley Pakistan Papua New Guinea Philippines Pixar Red-eared slider Reptiles of Asia Reptiles of Australia Reptiles of Europe Reptiles of Pakistan Reptiles of South America Reptiles of the Philippines Reptiles of the United States Reptiles of Western Australia Ridley Sea Turtles Samuel Garman Sea Turtle Restoration Project Sea turtles Society of Greece Southeast Asia Texas Torres Strait Trachemys scripta elegans Truong Son Muntiac Walt Disney Imagineering Western Australia World **Conservation Union**

🖉 Categories 🗒 Related terms 🍐 People 🖺 Locations 🖻 Organizations



Success 4 : Geolocalised search and Timelines





A search based application, including social graph

URBANIZER.com

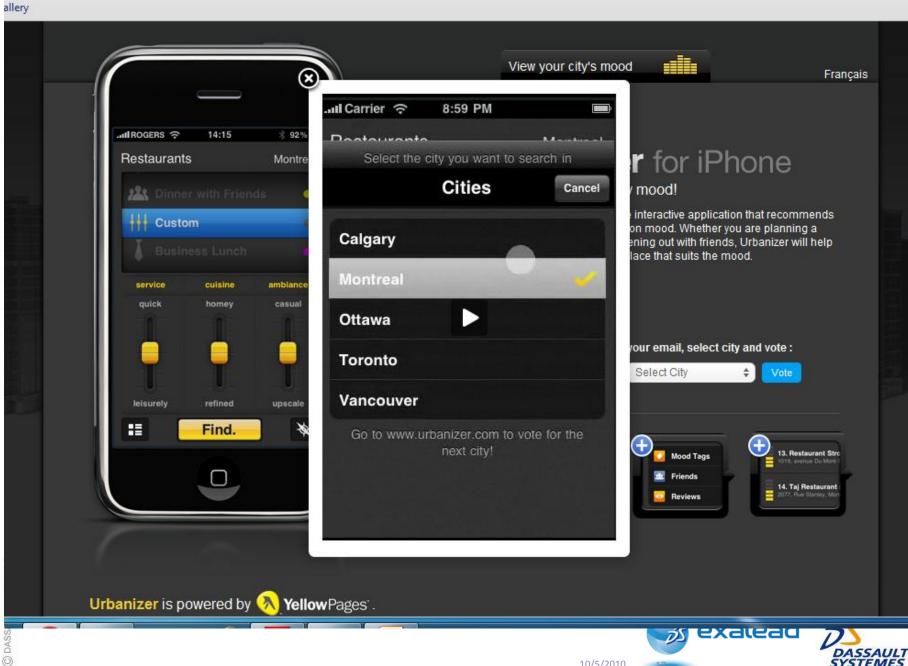


S www.urbanizer.com

Gallery



() www.urbanizer.com



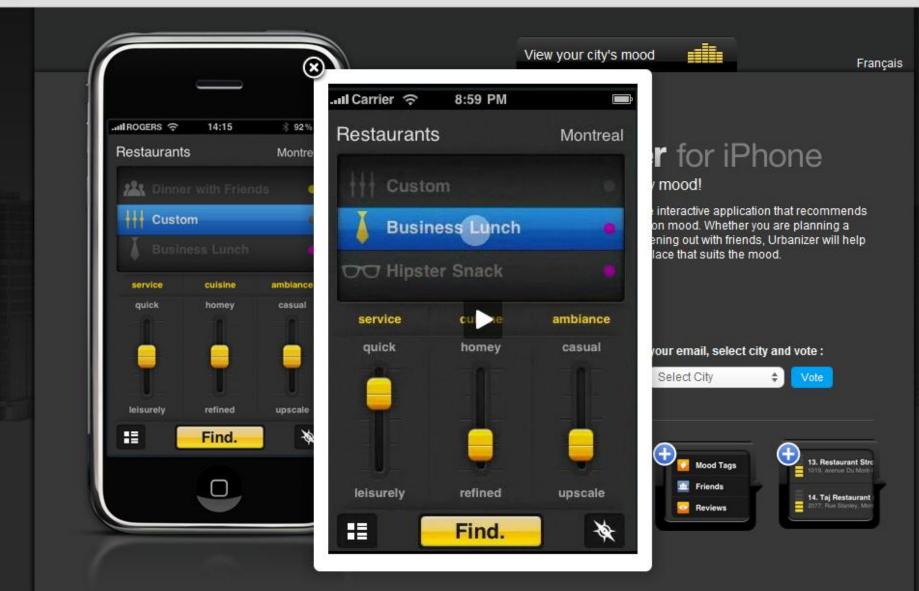
10/5/2010

SYSTEMES

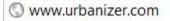
llery

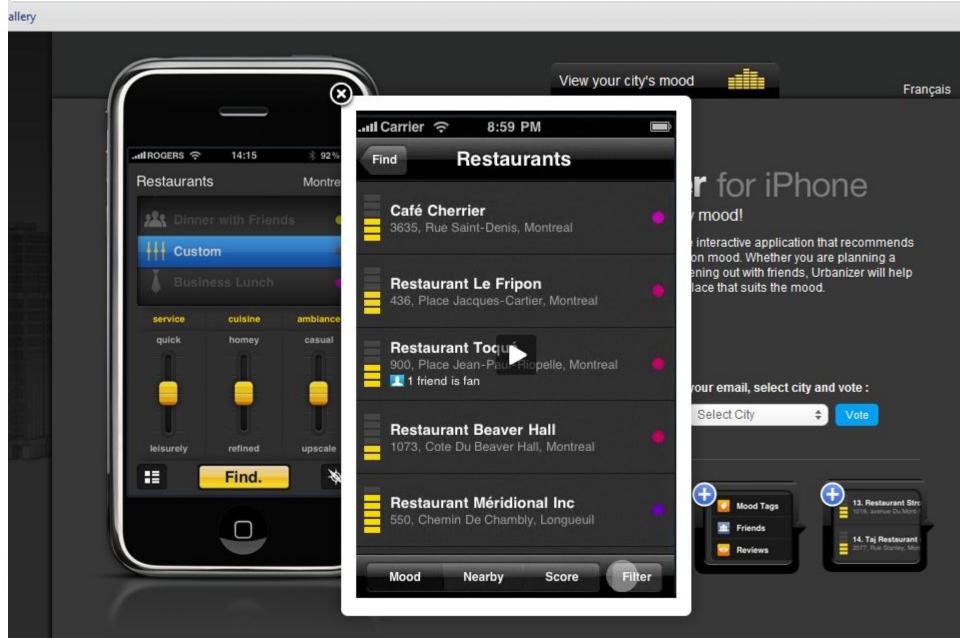


iallery



Urbanizer is powered by 🔊 Yellow Pages .

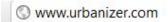


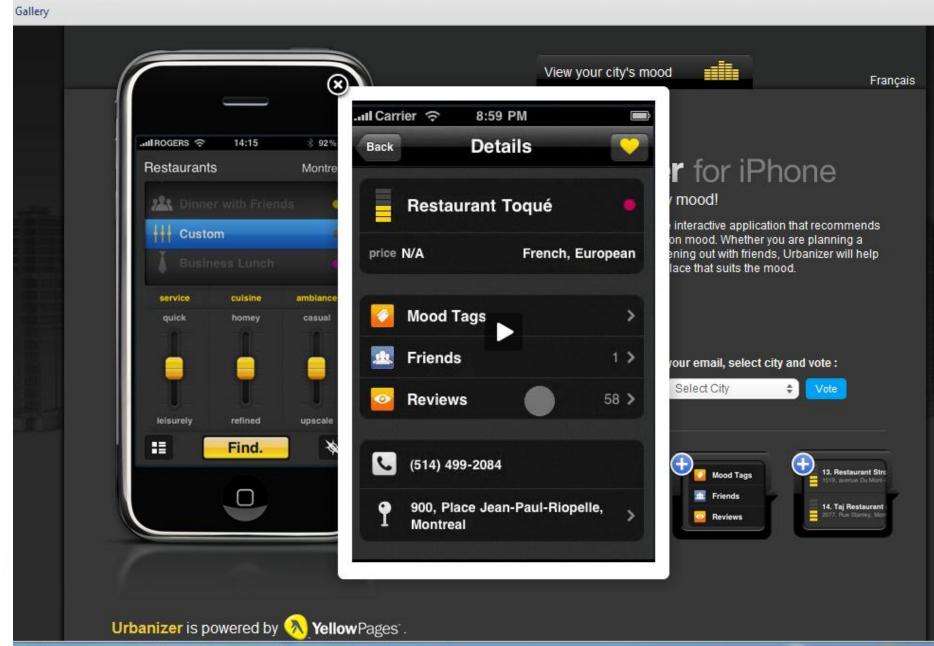


Urbanizer is powered by 💦 Yellow Pages .











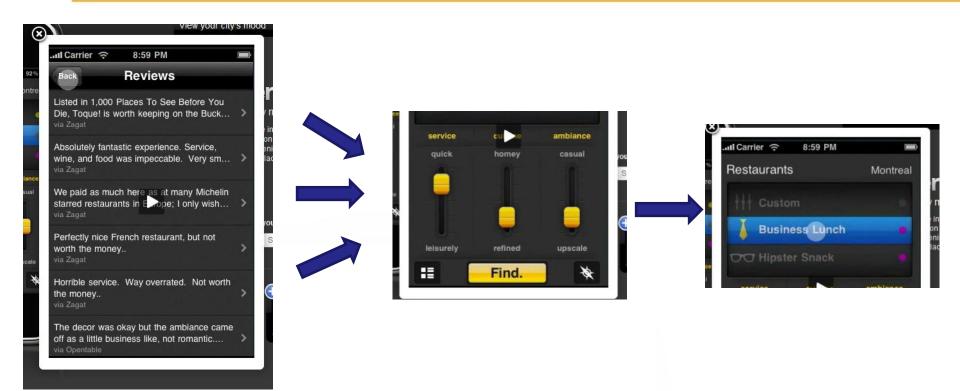
www.urbanizer.com



Urbanizer is powered by 🔊 Yellow Pages .



Extraction + Abstraction + Abstraction



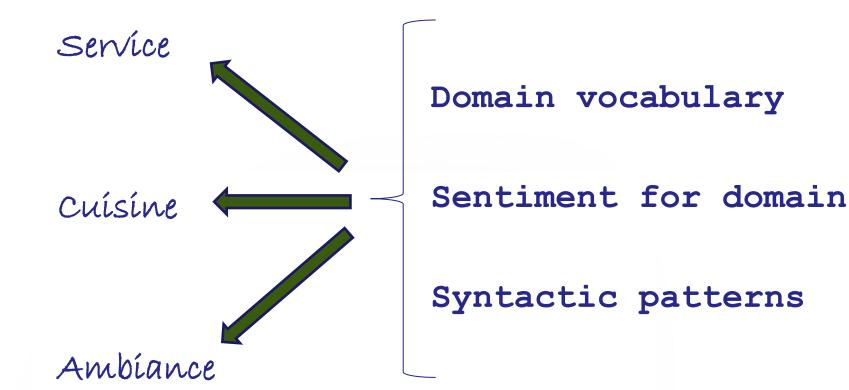
"...decor is not particularly elegant..."

ambiance \downarrow 70

Business Lunch \approx service ϵ [0,40] V cuisine ϵ [45,100] V ambiance ϵ [45,100]



Domain modeling

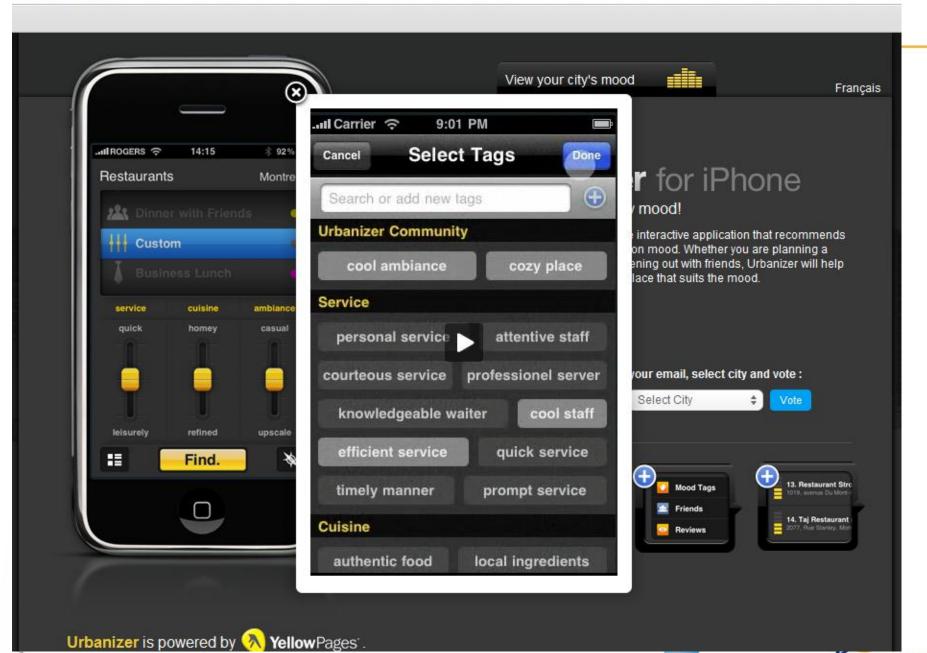






Urbanizer is powered by 🤼 Yellow Pages







00

S www.urbanizer.com

allery



Urbanizer is powered by 🔊 Yellow Pages .



Successes : Semantics for search

- Information Extraction
- Faceted Search
- Sentiment Analysis
 - → Domain Modeling
- Friends' opinion

- Drawbacks
 - Domain modeling economically feasible en masse



Challenge: searching ad hoc, idiosyncratic spaces

Video Stream from Meetings



<u>Jean Carletta, Human</u> <u>Communication Research Centre,</u> <u>School of Informatics, University of</u> <u>Edinburgh</u>

10/5/2010

Steve Renals, "Recognition and Understanding of Meetings", keynote for "Human Language Technologies: The 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics", June 1–6, 2010, Los Angeles, CA.



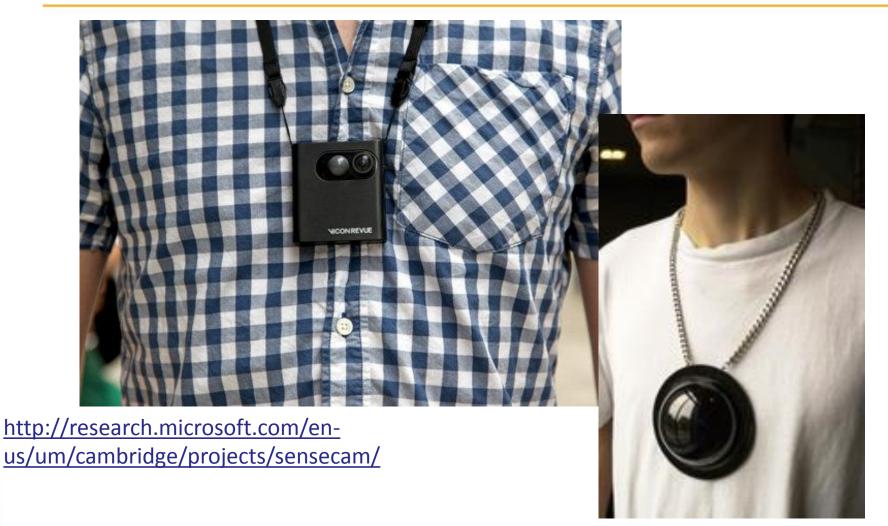
Challenge 1: searching ad hoc, idiosyncratic spaces



http://www.imrc.kist.re.kr/wiki/LifeLog



Microsoft SenseCam



http://justgetthere.us/blog/plugin/tag/cameras

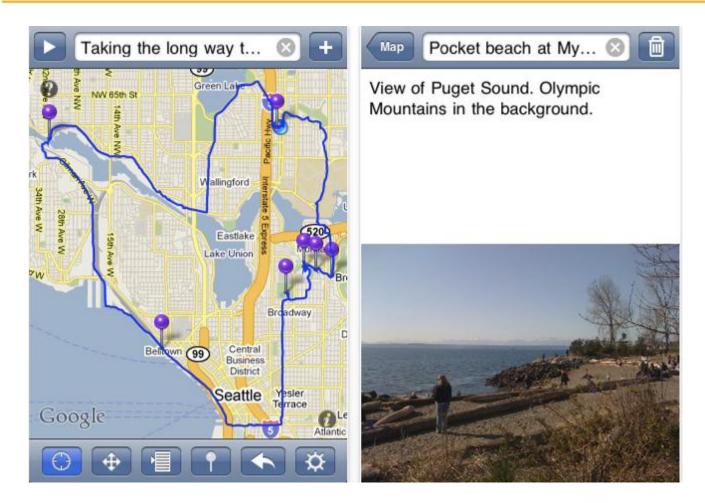




http://gadgets.infoniac.com/head-worncamera-for-special-events.html



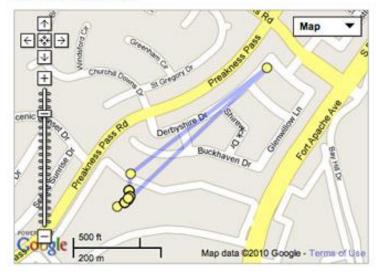
Bike My Way





Insta Mapper

Jarel's iPhone



Click on positions to select track endpoints and to see other options.

2010-03-11 01:28:59	2663 ft	0 mph	0
2010-03-11 01:30:01	2663 ft	0 mph	
2010-03-11 01:31:02	2663 ft	0 mph	
2010-03-11 01:32:03	2663 ft	0 mph	
2010-03-11 01:33:05	2663 ft	0 mph	
2010-03-11 01:34:11	2663 ft	0 mph	
2010-03-11 01:35:12	2663 ft	0 mph	
2010-03-11 01:36:13	2663 ft	0 mph	
2010-03-11 01:37:13	2663 ft	0 mph	
2010-03-11 01:38:16	2663 ft	0 mph	
2010-03-11 01:39:21	2663 ft	0 mph	
2010-03-11 01:40:25	2663 ft	0 mph	
2010-03-11 01:41:29	2663 ft	0 mph	
2010-03-11 01:42:31	2663 ft	0 mph	
2010-03-11 01:43:32	2663 ft	0 mph	- 1
Trip boundary			
2010-03-11 07:07:56	2627 ft	0 mph	
2010-03-11 07:27:59	236 ft	0 mph	
2010-03-11 07:28:31	2926 ft	0 mph	
2010-03-11 07:28:40	2926 ft	0 mph	
2010-03-11 07:36:00	2903 ft	0 mph	
2010-03-11 07:37:01	2857 ft	0 mph	
2010-03-11 07:38:01	2857 ft	0 mph	
2010-03-11 07:39:01	2857 ft	0 mph	
2010-03-11 07:40:02	2857 ft	0 mph	
2010-03-11 07:41:02	2857 ft	0 mph	
2010-03-11 07:42:03	2857 ft	0 mph	
2010-03-11 07:43:03	2857 ft	0 mph	
2010-03-11 07:44:04	2857 ft	0 mph	
2010-03-11 07:45:04	2857 ft	0 mph	1 A
2010-03-11 07:46:04	2857 ft	0 mph	¥



Mar 11 2010 Go

Challenge 2: Mapping Pertinent User information





NOVEMBER 2009 ATLANTIC MAGAZINE Filtering Reality

How an emerging technology could threaten civility

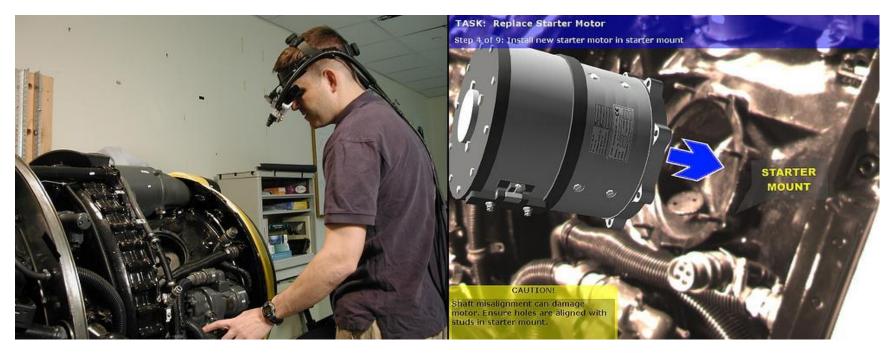
By Jamais Cascio Presented By IMAGE CREDIT: GLUEKIT



© DASSAULT SYSTEMES

Augmented Reality for Maintenance and Repair (ARMAR)

http://graphics.cs.columbia.edu/projects/armar/index.htm



(Left) A mechanic wearing a tracked head-worn display performs a maintenance task on a Rolls Royce DART 510 Engine. (Right) A view through the head-worn display depicts information provided using augmented reality to assist the mechanic.



Augmented Reality for Maintenance and Repair (ARMAR)

http://graphics.cs.columbia.edu/projects/armar/index.htm



(Left) A mechanic wearing a tracked head-worn display performs a maintenance task inside an LAV-25A1 armored personnel carrier. (Right) The AR condition in the study: A view through the head-worn display captured in a similar domain depicts information provided using augmented reality to assist the mechanic.

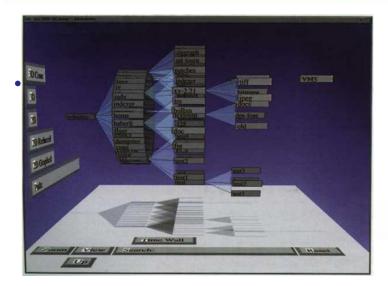


A user manipulates 3D virtual buttons while receiving haptic feedback from the underlying grooves of an engine compression section

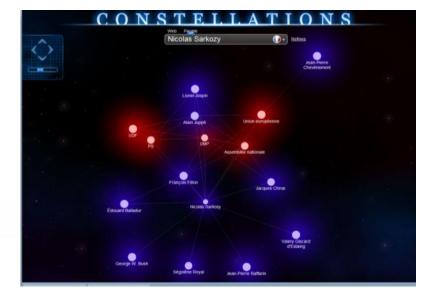




Challenge 3 : liberation from real world maps

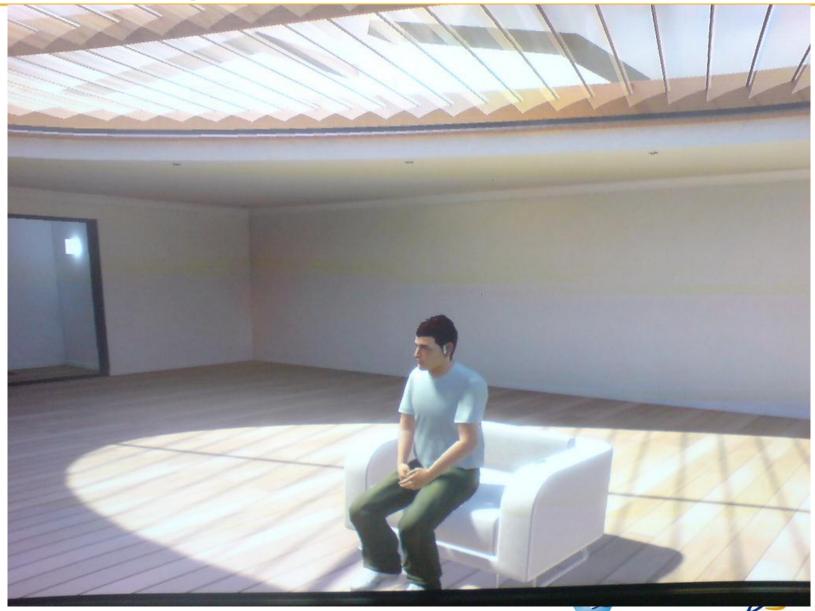


				1						1	-	-	_	_	_		-		1
F.	F.	E.	F.	·	÷.	•	÷.	٠.	•	•	•	•	•	•	·	•	G	G	G
F.	F.	F	F.	+	•	•	-	•	•	•	•	•	•			•	•	G	G
F'	F.	F	E.	÷.,	•		+	· ·	•	•	•		•	•			-	G	G
F	F.	F	•	•		•	•		•	•		•	-	•		•	-	G	G
E.	E.	•	•	•	•	+	1	+	•	A	A	A	A	•		•	•	•	•
E.	E.	E.	•	÷	•	•	•	•	•	A	A.	A	A.	A'	•	·	•	•	•
E.	E.	E.	+	+	•	·	·	·	•	₫.	×	A	A	A.	·	•	•	•	•
E.	E.	E.	+	÷	·	·	·	·	·		A'	A	A	A.	•		•	·	•
E.	E.	E.	E.		•	•	•	•	·	•	•	·	•	•	•	•	•	•	·
E.	E.	E.	E.	÷	·	·	•	•	·	•	·	÷	·	•	·	·		•	
E.	E.	E	E.	•	•	•	•	•	÷	•	·	•	•	•	•	•	•		•
D	÷.	E	•	•	÷.	•	+	•	+	·	·	•	·	·		•	-	•	•
D.	D.	D		•	•	•	•	•	+	•	•	•	•	•		·	•		•
D	D.	D.	D.	•	•	•	•	•	•	c	c	c'	¢.	•	·	•	•	•	•
D.	D.	D.	D	D.	•	•	•	•	+	c	c	c	c	•		•	-		•
÷	D.	D	D.	D.	D.	•	•	•	c'	c	c'	c	c'	•	•	•	-	•	
c.	•	D	D	D.	D	•	•	•	c	c	c	c	c	·	•	H	н	H	H
c.	c.	c	c.	c.	C.	c.	c.	c'	c'	c'	c'	c	c	•	H	H	H	H	н
c.	c.	C.	c.	c.	c.	c.	c.	c'	c'	c.	ċ.	c'	c		H	H	н	H.	н
c.	c.	c.	C.	c.	c.	с	c.	c'	c.	c'	c'	c.	c.		н	H	н	H.	H





What does space mean in a virtual world



http://twomeneatingforthree.com/



Conclusions

Successes

- Entities, Affect
- Faceted Search
- Local Search

Challenges

- Idiosyncratic volume
- Idiosyncratic abstraction
- Assigning meaning to virtual space



Thank you

ds.com