IntelliSemantic: company and solutions

AIDB 2014

Mission and focus

- Mission: document retrieval and information extraction for the professional user:
  - Quick and intuitive user interaction.
  - Precision and accuracy in results.
- Technology focus: our most innovative Natural Language and Semantic technologies applied to Business Intelligence.
- Market focus: target Patent Informatics applications and technical / scientific-related areas.

The company

- Founded in April 2005; hosted by the Incubator of the Politecnico di Torino I3p from 2005 to 2007.
- CEO/founder Alberto Ciaramella background:
  - Researcher and Research Manager at CSELT, the research branch of Telecom Italia, for speech and and Natural Language Processing.
  - Competitive Intelligence Manager at Loquendo SpA, the CSELT spin-off for speech and language processing.
- Key persons come from companies like Olivetti, Enginering and HP.

Some recognitions

- Most cited paper on Word patent Information since 2009, on Semantic technologies applied to Patent Informatics.
- Invited speaker at PDG 2013, on TOPAS Eu-Project.
- Invited speaker at AIDB 2013 and 2011: semantic technologies applied in Patent Informatics, and TOPAS.
- Invited speaker at Cebit Future talk 2010, on results of a research on semantic categorization.

Recent research activities

- Partner of the TOPAS EU-funded Project (2011-2013) on Semantics applied to Patent Informatics.
- R&D consulting on two Veneto-funded Projects (MRS and CLOUD, 2011-2013): development of a semantic Enterprise Search/Categorization engine for SMEs.
- Partner of the Piemonte-funded project Web-BI (2012-2013) for a demo intelligence portal on stem cell research.

Solutions

	Patent informatics	Other verticals
Application	MyIntelliPatent	on project request
Engines	IntelliCategory, TOPAS engines	IntelliCategory & IntelliFacet
Projects	based on MyIntelliPatent, IntelliCategory, TOPAS engines	Based on IntelliCategory, Intellifacet , MyIntelliPatent

Text mining applications

The IntelliSemantic platftorm includes:

- documents entities extraction.
- navigation on the whole set of documents.
- navigation inside a specific document.

It **enables applications** for the better use of

- structured documents, as:
  - company documents, news.
  - scientific, technical, medical literature.
- unstructured documents:
  - e-mails, tweets, tickets.

Any application requires a feasibilty study.

Patent documents engines (1)

IntelliSemantic can provide also:

- Patent entities extraction:
  - measurements, citations, substances, methods.
- Patent structure extraction:
  - large and small grain structure of the description.
  - large and small grain structure of claims.
- References between claims and description.
- in English, German and French.

Part of these technologies are evolutions of the EU funded research project TOPAS.

Patent documents engines (2)

These engines:

- Are easy to be integrated into third party platforms through RESTAPIs.
- Will be integrated in the new versions of MyIntelliPatent to extend:
  - The navigation inside a set of patents, e.g. to visualize patent documents by objectives, by background art and so.
  - The navigation inside a single patent, e.g. to access specific sections or to relate claims and description.

MyIntelliPatent

MyIntelliPatent is a **complete** solution for providing an **efficient** and **high quality** support in the most relevant patent related usecases, from **prior-art searches** to **validity searches**, from **technology analyses** to **competitors analyses.**

MyIntelliPatent is **a personalized** solution provided as a password protected Software as a Service and repository. Any registered company user can access to his/her **specific patent repository, topic vocabularies, patent annotations and scores.**

MyIntelliPatent benefits also from advanced Natural Language processing technologies by IntelliSemantic**.**

MyIntelliPatent strenghts

- **Time efficient** in supporting different patent usecases, from prior-art searches to technology/competitor analyses.
- **High recall and high precision**, since it separates the collect (for recall) from the results analysis (for precision).
- **Better insight ,** since it can use key information extracted from text (tags) besides metadata.
- **Customizable,** since it locally stores patents and vocabularies of interest for the user, besides users notes and scores.
- **Easy to learn,** with a smart web-inspired interface, based on tag-cloud views and facet-selection paradigm.
- **Easy to use,** directly browser accessed (Chrome, Firefox Safari) - so no local installation is required.

Patent usecases motivations

- Assessment of the novelty of your technical ideas.
- Evaluation of patent literature before writing a patent.
- Identification of the strength and weakness of a patent of you, or of your partners or of your competitors.
- Identification of the relevance of a patent to a product or a standard.
- Identification of technical suggestions not covered by patents in force.
- Identification of trend analyses for technologies.
- Monitoring competitors, including complementarities and evolutons.
- Identifying opportunities for your patents or known-how.
- Supporting the value analysis for a patent or a set of patents.

Different motivations can be of interest for different departments, as legal, R&D, M&S, general management.

Use cases-related pages

	Prior art search	Validity search	Technology analyses
Collect phase	Collect by similar text	Collect by similar patent	Collect by query
Tag phase	Tag	Tag	Tag
Search phase	Search	Search	Search
Analyze phase	Analyze	Analyze	Analyze

This slide relates pages, uses cases and phases.

Phases and usecases (1)

- Collect phase is for setup of the patent-set to be analyzed.
  - collect by similar patent is e.g. for validity search usecase, starting from one patent.
  - collect by similar text is used e.g. for a prior art search, starting from a text as an input (paper, document, proposed text for a patent).
  - collect by query is used for starting e.g. a portfolio analysis (eg. starting from a patent list), a technology analysis, a competitors analysis.
- Tag phase is for extracting and navigation the patent terminology. Tags (e.g. “adaptive modulation”) are multi-word expression that
  - can be learnt from a collected patent set and validated.
  - or can reuse a customer specific or topic specific vocabulary.

Phases and usecases (2)

- Search phase is for patent-set navigation.
  - Facets allow to identify and select patents characterized by high family size, high number of citations, first use of a new concept.
  - Different graphs and tables are available.
  - Graphs and tables can be dynamically reshaped by suitably selecting a facet (e.g. to obtain graphs and tables of a specific applicant.
- Analyze phase for a single patent allows to
  - identify other patents similar to a given one and rank them by a decreasing order of similarity.
  - add user specific annotations and scores to a patent.
  - directly access to different patent sections.

An example: tags distribution by year

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2
(no tags)	2	2	12	3	20	22	36	35	47	49	19	19	40	20	16	
audio signal	12	2		14	29	29	48	69	62	35	9	13	26	14	13	
auto dialer	10															
case	42	47	90	157	89	179	163	214	163	141	89	96	84	54	66	
data	76	66	115	202	132	249	259	309	252	206	146	142	156	118	135	
electronic pet					15											
emotional state	3	8		28	1	52	13	36	23	8	1	4	8	7	8	
flow chart	41		35	66	33	91	49	93	39	53	30	21	20	27	33	
information	60	68	107	197	111	232	215	297	257	214	157	140	159	114	125	
input	63	55	64	185	106	207	219	288	242	193	147	134	140	103	105	
input pattern matrix layer							3		5							
processing	49	59	97	148	117	196	185	262	225	170	139	131	130	102	104	
result of voice					15	5		1	1						2	
set	65	55	91	181	110	212	193	221	200	165	119	100	122	73	87	
speaker	79	58	110	198	118	207	178	236	201	172	125	117	137	110	113	
speaker recognition	14	35	34	76	24	79	36	43	28	38	27	13	31	23	30	
speaker verification	14	26	42	106	39	96	29	62	30	30	17	24	25	10	22	
speech	69	57	108	181	112	228	197	255	225	167	147	127	143	113	113	
speech data	34	17	44	50	20	68	22	63	23	46	18	16	18	21	21	
speech recognition	44	23	57	108	61	133	108	132	128	105	90	81	46	53	57	
speech signal	36	15	56	82	25	88	45	72	76	39	25	21	34	27	20	
standard pattern matrix layer							3		5							
steps	56	50	68	146	87	171	166	194	148	130	94	89	102	69	82	
value	31	46	68	107	95	162	129	184	135	130	66	65	111	53	84	
voice signal	22	14	26	34	21	60	44	100	18	50	35	13	39	25	26	
weighting vector		3		1					5	3						

The distribution of tags by year allows to identify first patents using a concept and the most popular concepts now.

Custom projects examples

MyIntelliPatent architectural flexibility and IntelliSemantic know how allow to implement custom projects as:

- Migration to MyIntelliPatent of
  - user collected patents, vocabularies, annotations.
  - user collected non patent literature (NPL).
- Customized postprocessing according to rules or examples agreed with the customer for
  - categorizing patents by user defined technologies.
  - ranking patents in a portfolio according to technical quality, commercial quality, relevance to a standard.

IntelliSemantic srl
Contacts and references

E-mail: info@intellisemantic.com

Tel.: +39 011 9550 380

Website: http://www.intellisemantic.com