



AUTOMATED INFORMATION ENRICHMENT FOR A BETTER SEARCH

JOSE LUIS PREZA

What is information enrichment?

Adding metadata

**Metadata can include: Tags, Description, Keywords,
Categories (SKOS, Getty), etc.**

What is metadata good for?

searching

LET US SEARCH

PHAIDRA User-ID: Password:

Browse

Search

Search in all fields

Sort by

Upload date

Advanced Search

Search Results

[Imprint](#) | [Contact](#) | [Service Page](#) | [Terms of use](#)
 University of Vienna | Universitätsring 1 | 1010 Vienna | T +43-1-4277-0

✓ Search through all available fields

- Owner
- Full-text search
- Object type
- Title
- Abstract
- Upload date
- Identifier
- General Identifier
- Title
- Subtitle
- Alternative Title
- Language
- Description or Additional Data
- Keywords
- Coverage
- Institutional Repository
- Identifiers
- Resource Identifier
- Lifecycle
- Version
- Status
- Peer Reviewed
- Contribute
- Role
- Entity / Personal data
- Firstname
- Lastname
- Institution
- Title
- Student ID
- ORCID
- VIAF

HOW IS METADATA CURRENTLY ADDED?

During the process of storing a digital asset in a Repository, the user has the ability to

manually describe the asset using
“Metadata” Fields.

RISKS OF MANUAL PROCESSES

Incompleteness

Lack of time

Lack of imagination

Typos

Incorrectly classified

AUTOMATING METADATA

TEXT ANALYTICS

COGNITIVE COMPUTING

Technology platforms for Artificial Intelligence and Signal Processing:

machine learning, reasoning, natural language processing, speech and vision, human-computer interaction, dialog and narrative generation, and more

wikipedia

COGNITIVE COMPUTING SERVICES



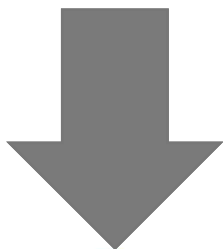
IBM Watson

MonkeyLearn

AUTOMATED METADATA: IBM WATSON TEXT ANALYTICS API



TITLE
ABSTRACT



Watson
Text analytics

Concepts
Entities
Keywords

EXAMPLE: TITLE AND DESCRIPTION OF THE OBJECT (TEXT I SEND TO WATSON)



Overview of work-life balance discourse and its relevance in current economic scenario

During the second half of the 20th century, with work demands increasingly encroaching on family and personal time at a faster pace, employers acknowledged the need of work-life balance programmes to facilitate employees maintain a healthy balance between the conflicting demands of their work and personal life. Availability of work-life balance facilities to employees witnessed a phenomenal growth between the late eighties of the 20th century and early years of the 21st century. This growth has been abruptly interrupted by the current economic downturn. Increasing numbers of organisations, in the name of cost cutting, have either curtailed work-life balance facilities or are contemplating to do the same. This paper analyses the emergence of work-life balance discourse, from the days of early communal living till the present day theories, and presents a macro level model of work-life balance. Further, a detailed analysis of proven and anticipated benefits of work-life balance is presented to justify the need of work-life balance initiatives at organisational level during the present economic downturn.



IBM WATSON KEYWORDS

Entity	Keywords		
Entities	numbers, cost cutting, economic downturn, work-life balance facilities, analysis, employees, personal time, conflicting demands, macro level model, early communal living, present economic downturn		
Keywords	Availability, present day theories, phenomenal growth		
Taxonomy	current economic downturn		
Concepts	family, organisational level, late eighties, 20th century, work-life balance programmes, work-life balance discourse		
Document Sentiment	half, Overview, work-life balance		
Targeted Sentiment	employers, 21st century, faster pace, healthy balance		
Document Emotions (Beta)	relevance, emergence, work-life balance initiatives		
Relations	current economic scenario		
Language	need, personal life, work demands		
Title			
Author	Keyword		
Text	Relevance		
Feeds	Sentiment		
Microformats			
	work-life balance	0.917959	mixed
	work-life balance discourse	0.772748	positive
	work-life balance facilities	0.730362	mixed
	work-life balance programmes	0.623306	positive
	work-life balance initiatives	0.582295	positive
	healthy balance	0.395272	positive
	economic downturn	0.388243	mixed
	current economic downturn	0.380741	negative
	current economic scenario	0.376483	neutral
	20th century	0.370891	positive
	present economic downturn	0.370468	positive
	early communal living	0.34989	positive
	macro level model	0.346475	positive
	conflicting demands	0.312725	positive
	work demands	0.303198	positive
	faster pace	0.297767	positive
	late eighties	0.294348	positive

IBM WATSON CONCEPTS

Entities	Concept	Relevance	Linked Data
Keywords	21st century	0.952892	dbpedia freebase yago
Taxonomy			
Concepts	Recession	0.723387	dbpedia freebase
Document Sentiment			
Targeted Sentiment	20th century	0.718918	dbpedia freebase yago
Document Emotions (Beta)			
Relations	Present	0.709026	dbpedia freebase
Language	Unemployment	0.699621	dbpedia freebase website
Title			
Author	Modern history	0.691225	dbpedia freebase yago
Text			
Feeds	Time	0.654916	dbpedia freebase
Microformats			
	Health	0.626727	dbpedia freebase website

IBM WATSON ENTITIES

[Click here to learn more about entities.](#)

Visual

JSON

API

Entities	economic downturn					
Keywords						
Taxonomy						
Concepts						
Document Sentiment						
Targeted Sentiment						
Document Emotions (Beta)						
Relations						
Language						
Title						
Author	Entity	Relevance	Sentiment	Type	Subtypes	Linked Data
Text	economic downturn	0.933285	mixed	FieldTerminology		
Feeds						
Microformats						

AUTOMATED METADATA FOR A BETTER SEARCH

Top 50 - Watson Categorisation (in progress)

Concepts Graph Concepts Keywords Entities

Search Concept

Zoom with mouse to cluster. or DOUBLE Click on a circle to drill in

