

Collaborative Knowledge Engineering Methods and Tools for System Design

PhD Thesis

Krzysztof Kutt

supervisor: dr hab. inż. Grzegorz J. Nalepa, prof. AGH

Wrocław, 09.05.2018

OVERVIEW

Introduction

CKE Process

Methods and Tools

Evaluation

Summary

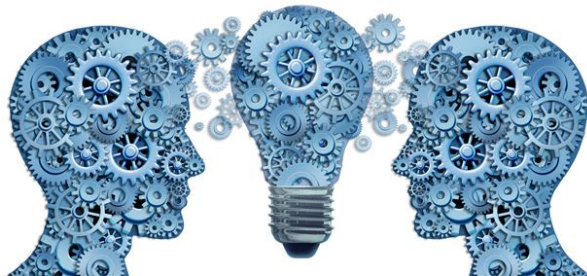
COLLABORATIVE KNOWLEDGE ENGINEERING

*an engineering discipline that involves **integrating knowledge into computer systems** in order to solve complex problems normally requiring a high level of human expertise [47]*



COLLABORATIVE KNOWLEDGE ENGINEERING

a joint involvement of participants in the project for a common purpose, although it may result from different motivations

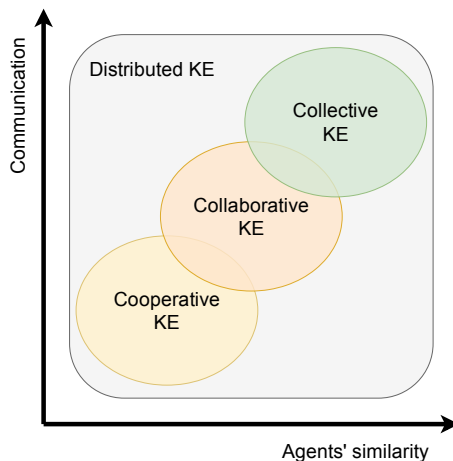


COLLABORATIVE KNOWLEDGE ENGINEERING

a joint involvement of participants in the project for a common purpose, although it may result from different motivations



COLLABORATIVE KNOWLEDGE ENGINEERING



WHY DOES IT MATTER?

PARADIGM SHIFT

	Past	Present
Design	KE expert with domain expert access	KE expert paired with domain expert(s)
Population	KE expert learns domain	KE and domain experts determine the vocabulary
Evolution	KE expert heavily involved	KE expert involved in customizing tools that domain experts use
Tool Users	Trained in Computer Science	Trained in Domain Sciences
Application Users	Well understood group	Diverse and evolving group
Reuse	Well thought out	Expect the unexpected

(adapted from [99])

WHY DOES IT MATTER?

MANY USE CASES: CATALOGUS PROFESSORUM LIPSIENSIS



Professor catalog of the University of Leipzig | catalogus professorum Lipsiensis

[Home](#)
[Epochs](#)
[Faculties](#)
[Professors of
the day](#)
[Rectors and
deans](#)
[entire
directory](#)

[Background](#)
[Information in
English](#)
[Abbreviations
Literature](#)
[Epochs > 1919-1932 and 1933-1945](#)
[Faculties > Faculty of Arts I, University of Leipzig - Faculty of Philological-Historical Division \(1920-1951\)](#)
[Data: Resource | RDF | PDF | Printer-friendly](#)

Prof. Dr. phil. Schücking *Levin* Ludwig

Life

b. 5/29/1878 in Burgsteinfurt

d. 12/10/1964 in Farchant

PND: [117124931](#)



Source: Private collection

Curriculum vitae

Study

- 1897-1901 Study: modern languages and art history in Freiburg, Göttingen, Berlin and Munich

Qualification

- 1904 Habilitation for English Language and Literature at the University of Göttingen
Title of work: broad set of shortcut in Beowulf.
- 1901 Promotion to Dr. phil. in English Philology at the University of Göttingen
Title of work: English Material Relations of the Italian comedy to Lilly.

WHY DOES IT MATTER?

MANY USE CASES: WIKIPEDIA



WIKIPEDIA
The Free Encyclopedia

Main page

[Contents](#)

[Featured content](#)

[Current events](#)

[Random article](#)

[Donate to Wikipedia](#)

[Wikipedia store](#)

Interaction

[Help](#)

[About Wikipedia](#)

[Community portal](#)

[Recent changes](#)

[Contact page](#)

Tools

[What links here](#)

[Related changes](#)

[Upload file](#)

[Special pages](#)

[Permanent link](#)

[Page information](#)

[Wikidata item](#)

Not logged in [Talk](#) [Contributions](#) [Create account](#) [Log in](#)

Main Page [Talk](#)

Read

[View source](#)

[View history](#)



Welcome to **Wikipedia**,

the *free encyclopedia* that *anyone* can edit.

5,132,278 articles in [English](#)

- [Arts](#)
- [Biology](#)
- [Geography](#)
- [History](#)
- [Mathematics](#)
- [Science](#)
- [Society](#)
- [Technology](#)
- [All portals](#)

From today's featured article



Replicas of the "Twin Sisters", cannons used in the Battle of San Jacinto

The **Runaway Scrape** was the 1836 escape of Texas residents from the encroaching Mexican Army of Operations under the command of [Antonio López de Santa Anna](#) during the [Texas Revolution](#). Civilian evacuations

began on the Gulf Coast in January after the vanguard of the Mexican army crossed the [Rio Grande](#) to quell the Insurrection of American colonists and [Tejanos](#) (Mexicans born in Texas). Weeks later, news of the [Battle of the Alamo](#) and the [Goliad massacre](#) created a state of panic. [Sam Houston](#) was the Texas commander-in-chief of raw recruits who had little or no combat experience. Fleeing civilians moved in tandem with Houston's troops for protection, as he sought a safe training camp for his soldiers. The pursuing Mexican army had orders to execute all rebel combatants, and it cut a swath of destruction in its search for them. After a mere three weeks training near

In the news

- American singer and songwriter **Prince** is found dead at the age of 57.
- At least 64 people are killed and more than 340 others injured in **an attack** in [Kabul](#), Afghanistan.
- Ethiopians [Lemi Berhanu Hayle](#) and [Atsede Baysa](#) (*pictured*) win the men's and women's **Boston Marathon**, respectively.
- **A magnitude 7.8 earthquake** hits Ecuador, killing at least 480 people and injuring more than 4,000 others.
- **Two earthquakes** kill at least 40 people and injure more than 3,000 others across [Kyushu](#), Japan.



Atsede Baysa

Ongoing events

Recent deaths: [Chyna](#) · [Victoria Wood](#) · [Estelle Balet](#) · [Bill Gray](#)

WHY DOES IT MATTER?

MANY USE CASES: WIKIPEDIA

- ▶ People want to collaborate!
- ▶ There is a need for tools that work → wiki!

But wiki is not good enough.

Knowledge engineering methods can improve it!

WHY DOES IT MATTER?

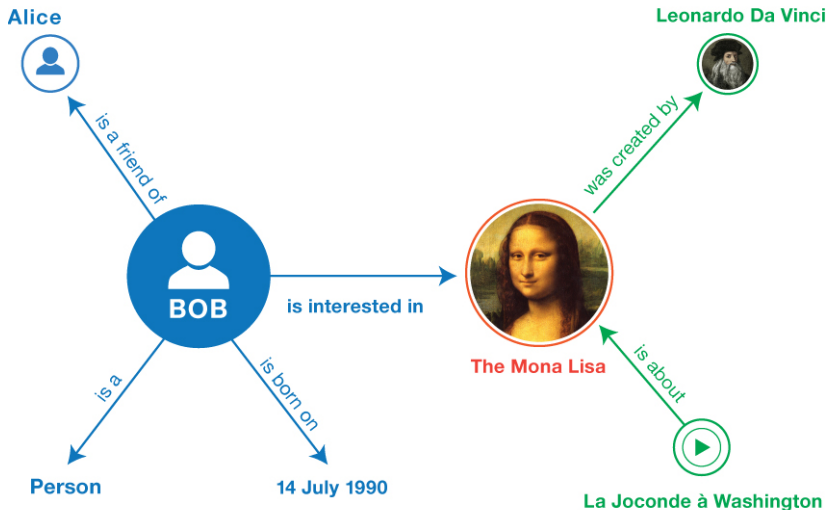
MANY USE CASES: WIKIPEDIA

- ▶ People want to collaborate!
- ▶ There is a need for tools that work → wiki!

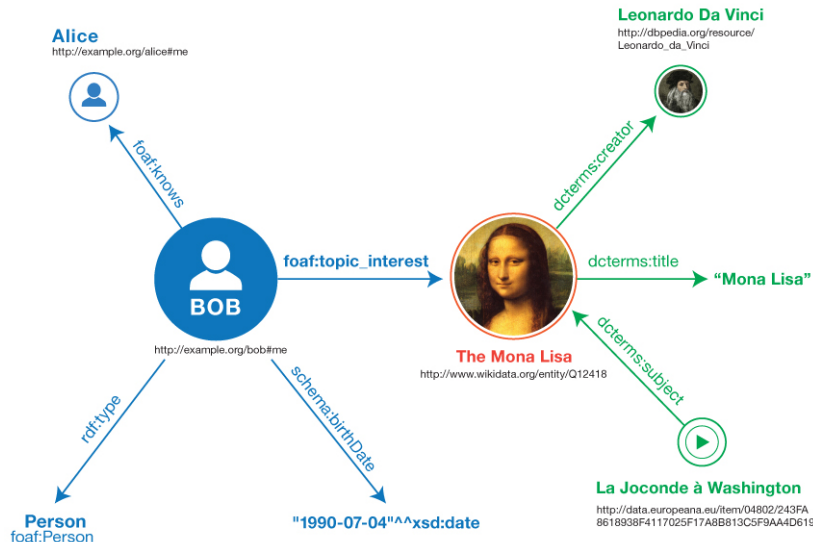
But wiki is not good enough.

Knowledge engineering methods can improve it!

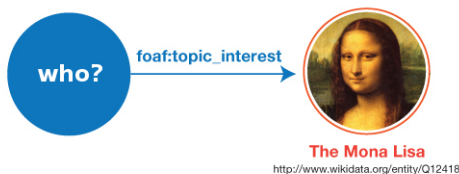
RDF



RDF



SPARQL



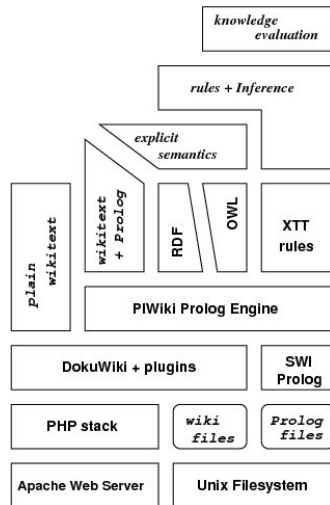
SEMANTIC WIKIS

Wiki + Semantic Web (RDF/SPARQL) = Semantic Wiki

- ▶ 2004: first semantic wiki (Platypus)
- ▶ 2005-2006: “semantic wiki explosion”
- ▶ 2017: Semantic MediaWiki, KnowWE, OntoWiki, Loki

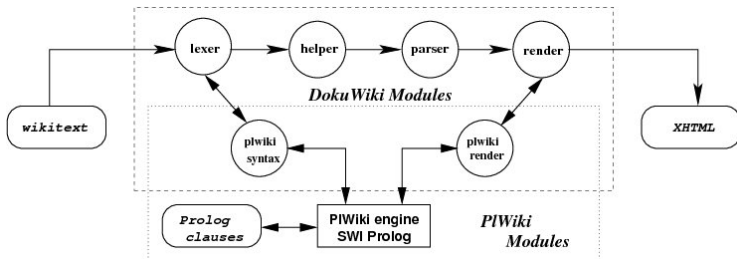
LOKI

DokuWiki + Prolog = LOKI



LOKI

DokuWiki + Prolog = LOKI



LOKI

Wiki

Title: The Call of Cthulhu

Author: [h_p_lovecraft](#)

Publisher: [iap](#)

Date: 2009

Language: english

Genre: [horror](#)

Pages: 52

Keywords: evenings

LOKI

Wiki + annotations

Title: The Call of Cthulhu****Title**:** [[title:=The Call of Cthulhu]]**Author:** h_p_lovecraft****Author**:** [[author::bookstore:author:h_p_lovecraft]]**Publisher:** iap****Publisher**:** [[publisher::bookstore:publisher:iap]]**Date:** 2009****Date**:** [[date:=2009]]**Language:** english****Language**:** [[language:=english]]**Genre:** horror****Genre**:** [[genre::bookstore:genre:horror]]**Pages:** 52**Keywords:** evenings

LOKI

Wiki + annotations = machine transformations

Title: The Call of Cthulhu

****Title**:** [[title:=The Call of Cthulhu]]

Author: h_p_lovecraft

****Author**:** [[author::bookstore:author:h_p_lovecraft]]

Publisher: iap

****Publisher**:** [[publisher::bookstore:publisher:iap]]

Date: 2009

****Date**:** [[date:=2009]]

Language: english

****Language**:** [[language:=english]]

Genre: horror

****Genre**:** [[genre::bookstore:genre:horror]]

Pages: 52

Keywords: evenings

Recommendation:

Books by this author:

at_the_mountains_of_madness

the_call_of_cthulhu

► mature agile methodology and knowledge maintenance cycle

Process

- shared repository that supports range of expertise levels
- compatibility with existing mainstream system
- easily accessible current KB status
- adapted to specific project needs
- domain knowledge owned by domain users, not a third party
- proper representation: easy and powerful
- review process
- automatic knowledge checking mechanisms
- knowledge consistency and credibility
- handling of users' conflicts
- robust versioning mechanism
- support for users' kinds and expectations
- usability consideration: experiments and surveys
- taking care of compatibility with established practices
- incentives: gamification mechanisms?

▶ mature agile methodology and knowledge maintenance cycle

Process

- ▶ shared repository that supports range of expertise levels
- ▶ compatibility with existing mainstream system
- ▶ easily accessible current KB status
- ▶ adapted to specific project needs

Tools

- ▶ domain knowledge owned by domain users, not a third party
- ▶ proper representation: easy and powerful
- ▶ review process
- ▶ automatic knowledge checking mechanisms
- ▶ knowledge consistency and credibility
- ▶ handling of users' conflicts
- ▶ robust versioning mechanism
- ▶ support for users' kinds and expectations
- ▶ usability consideration: experiments and surveys
- ▶ taking care of compatibility with established practices
- ▶ incentives: gamification mechanisms?

► mature agile methodology and knowledge maintenance cycle

Process

- shared repository that supports range of expertise levels
- compatibility with existing mainstream system
- easily accessible current KB status
- adapted to specific project needs

Tools

- domain knowledge owned by domain users, not a third party
- proper representation: easy and powerful

KRR

- review process
- automatic knowledge checking mechanisms
- knowledge consistency and credibility
- handling of users' conflicts
- robust versioning mechanism
- support for users' kinds and expectations
- usability consideration: experiments and surveys
- taking care of compatibility with established practices
- incentives: gamification mechanisms?

► mature agile methodology and knowledge maintenance cycle

Process

- shared repository that supports range of expertise levels
- compatibility with existing mainstream system
- easily accessible current KB status
- adapted to specific project needs

Tools

- domain knowledge owned by domain users, not a third party
- proper representation: easy and powerful

KRR

- review process
- automatic knowledge checking mechanisms
- knowledge consistency and credibility
- handling of users' conflicts

**Quality
management**

- robust versioning mechanism
- support for users' kinds and expectations
- usability consideration: experiments and surveys
- taking care of compatibility with established practices
- incentives: gamification mechanisms?

- ▶ mature agile methodology and knowledge maintenance cycle

Process

- ▶ shared repository that supports range of expertise levels
- ▶ compatibility with existing mainstream system
- ▶ easily accessible current KB status
- ▶ adapted to specific project needs

Tools

- ▶ domain knowledge owned by domain users, not a third party
- ▶ proper representation: easy and powerful

KRR

- ▶ review process
- ▶ automatic knowledge checking mechanisms
- ▶ knowledge consistency and credibility
- ▶ handling of users' conflicts

**Quality
management**

- ▶ robust versioning mechanism
- ▶ support for users' kinds and expectations

**Change
management**

- ▶ usability consideration: experiments and surveys
- ▶ taking care of compatibility with established practices
- ▶ incentives: gamification mechanisms?

- ▶ mature agile methodology and knowledge maintenance cycle

Process

- ▶ shared repository that supports range of expertise levels
- ▶ compatibility with existing mainstream system
- ▶ easily accessible current KB status
- ▶ adapted to specific project needs

Tools

- ▶ domain knowledge owned by domain users, not a third party
- ▶ proper representation: easy and powerful

KRR

- ▶ review process
- ▶ automatic knowledge checking mechanisms
- ▶ knowledge consistency and credibility
- ▶ handling of users' conflicts

**Quality
management**

- ▶ robust versioning mechanism
- ▶ support for users' kinds and expectations

**Change
management**

- ▶ usability consideration: experiments and surveys
- ▶ taking care of compatibility with established practices
- ▶ incentives: gamification mechanisms?

Motivation

- ▶ mature agile methodology and knowledge maintenance cycle

Process

- ▶ ~~shared repository that supports range of expertise levels~~
- ▶ ~~compatibility with existing mainstream system~~
- ▶ ~~easily accessible current KB status~~
- ▶ ~~adapted to specific project needs~~

Tools

- ▶ ~~domain knowledge owned by domain users, not a third party~~
- ▶ ~~proper representation: easy and powerful~~

KRR

- ▶ review process
- ▶ automatic knowledge checking mechanisms
- ▶ knowledge consistency and credibility
- ▶ handling of users' conflicts

**Quality
management**

- ▶ robust versioning mechanism
- ▶ support for users' kinds and expectations

**Change
management**

- ▶ usability consideration: experiments and surveys
- ▶ taking care of compatibility with established practices
- ▶ incentives: gamification mechanisms?

Motivation

KNOWLEDGE ENGINEERING AND SOFTWARE ENGINEERING

<i>Knowledge Engineering</i>	<i>Software Engineering</i>
Quality of knowledge?	Quality of software?

KNOWLEDGE ENGINEERING AND SOFTWARE ENGINEERING

<i>Knowledge Engineering</i>	<i>Software Engineering</i>
Quality of knowledge? Formal methods	Quality of software?

KNOWLEDGE ENGINEERING AND SOFTWARE ENGINEERING

<i>Knowledge Engineering</i>	<i>Software Engineering</i>
Quality of knowledge?	Quality of software?
Formal methods	Unit tests!

KNOWLEDGE ENGINEERING AND SOFTWARE ENGINEERING

<i>Knowledge Engineering</i>	<i>Software Engineering</i>
Quality of knowledge?	Quality of software?
Formal methods	Unit tests!
Methodology / Process?	Methodology / Process?

KNOWLEDGE ENGINEERING AND SOFTWARE ENGINEERING

<i>Knowledge Engineering</i>	<i>Software Engineering</i>
Quality of knowledge? Formal methods	Quality of software? Unit tests!
Methodology / Process?	Methodology / Process? Waterfall or others?

KNOWLEDGE ENGINEERING AND SOFTWARE ENGINEERING

<i>Knowledge Engineering</i>	<i>Software Engineering</i>
Quality of knowledge? Formal methods	Quality of software? Unit tests!
Methodology / Process?	Methodology / Process? Waterfall or others? Agile!

THESIS

OBJECTIVES

*The main goals were to **describe a Collaborative Knowledge Engineering process** that provides a general framework for defining roles which should be identified in a group and steps that should be taken in this process, as well as to propose **methods and tools that support the defined CKE process**, leading to the creation of good quality KB in reasonable time, through the means that will be convenient for target users.*

THESIS

SCOPE

- ▶ participants are generally motivated to work,
- ▶ development is done in an agile and decentralized way,
- ▶ group is working voluntarily (material and formal incentives are not considered),
- ▶ CKE process is understood as a creation of KB structured using the **ontology-like** knowledge representation grounded in **RDF** abstract syntax

OVERVIEW

Introduction

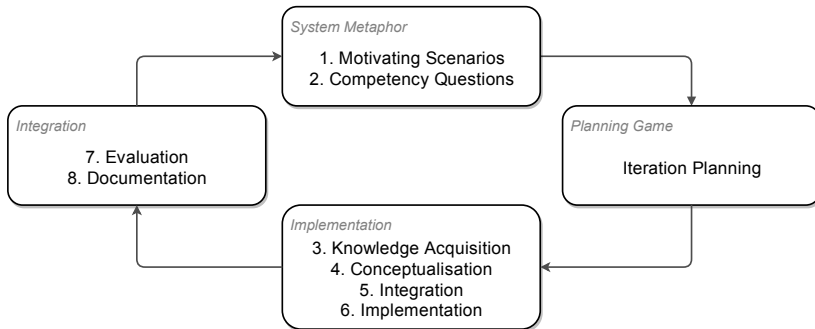
CKE Process

Methods and Tools

Evaluation

Summary

CKE AGILE PROCESS

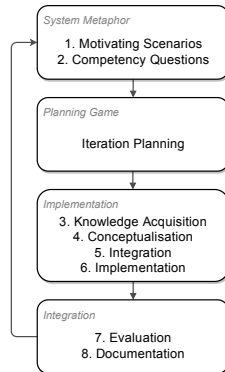
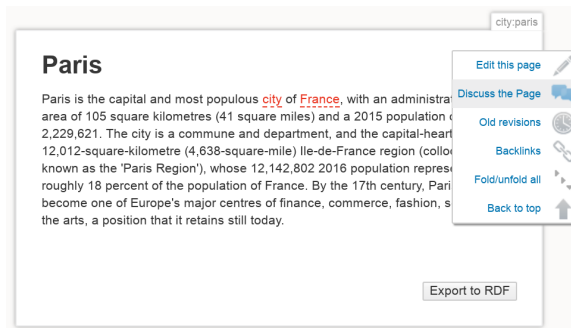


CKE AGILE PROCESS

GOOD PRACTICES

- ▶ Frist iteration: basic vocabulary as a reference point
- ▶ Subsequent iterations: from less to more formal
- ▶ Group:
 - ▶ the product owner, who has the whole system vision
 - ▶ the CKE process master, who oversees the project course
 - ▶ the team of 3-9 people: domain experts and 1-2 knowledge engineers
- ▶ Roles in the team:
 - ▶ the adders, who create a lot of material
 - ▶ the synthesizers, who take care about semantics and concepts interrelations
 - ▶ “the cops”, who are responsible for imposing standards

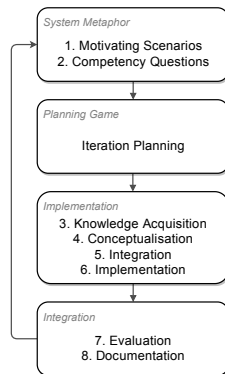
EUROPEAN WIKI



EUROPEAN WIKI

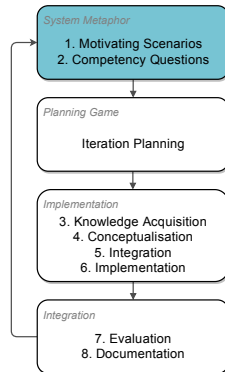
- ▶ Group of seven colleagues
- ▶ Description of all European countries
→ support during travels
- ▶ There is no “client” → one person selected as a product owner
- ▶ One selected as a CKE process master
- ▶ Basic knowledge of the KE → no knowledge engineer is needed

Here: the first iteration, three wiki pages,
two users (kkutt and yoda)



EUROPEAN WIKI: STEP BY STEP

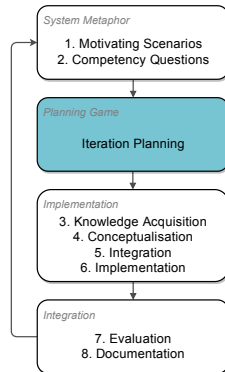
- ▶ Three use cases were proposed:
 - ▶ I want to see cities: A, B and C. In what order do I have to visit them to be able to use direct flights?
 - ▶ I want to see all European capitals. List all of them.
 - ▶ I am in city A. What is interesting here?
- ▶ Reasoning unit tests were specified



EUROPEAN WIKI: STEP BY STEP

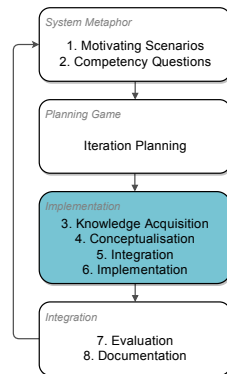
In the first iteration three tasks were defined:

- ▶ Describe London
- ▶ Describe Paris
- ▶ Describe Cracow



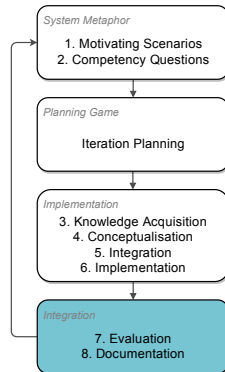
EUROPEAN WIKI: STEP BY STEP

- ▶ Two users, six changes
- ▶ Discussion of the difficult points
- ▶ The semantic changelog created
- ▶ Gamification-based incentives for users
- ▶ Task management on an iteration board



EUROPEAN WIKI: STEP BY STEP

- ▶ Reasoning unit tests: knowledge quality and the level of requirements fulfillment
- ▶ Comments with design decision made (city category)



OVERVIEW

Introduction

CKE Process

Methods and Tools

Evaluation

Summary

QUALITY MANAGEMENT

Methods and Tools

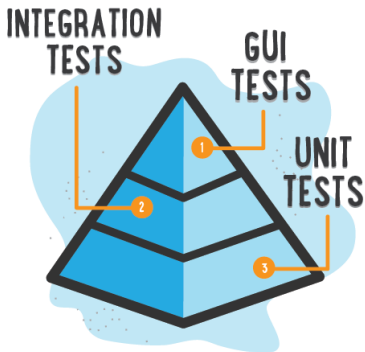
Quality Management

Change Management

User Involvement

REASONING UNIT TESTS

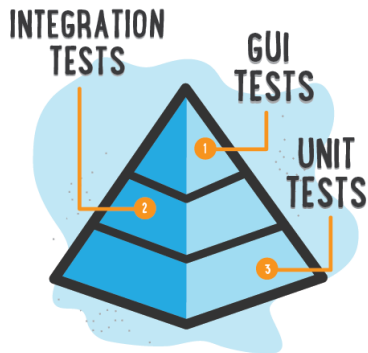
IDEA ADOPTED FROM SOFTWARE ENGINEERING



ILLUSTRATED BY SEGUE TECHNOLOGIES

REASONING UNIT TESTS

IDEA ADOPTED FROM SOFTWARE ENGINEERING



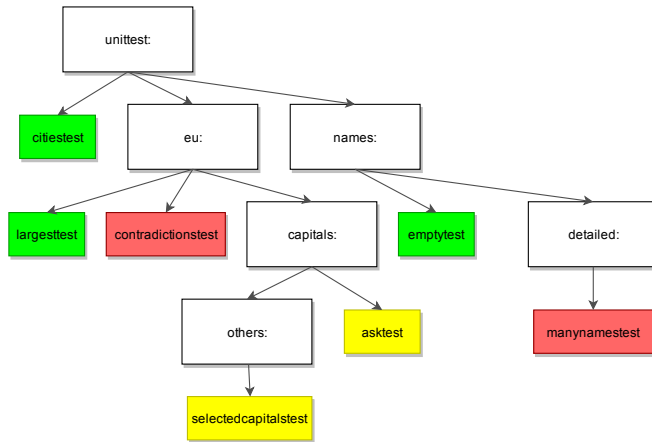
ILLUSTRATED BY SEGUE TECHNOLOGIES

Reasoning unit tests:

- ▶ Adopted by [164]
- ▶ Verify knowledge
- ▶ Define **expectations**

REASONING UNIT TESTS

TESTS HIERARCHY



REASONING UNIT TESTS

LOKI: 1. SPECIFY THE SPARQL-COMPATIBLE QUERY

```
1 <pl format="sparql">
2 PREFIX wiki: <>
3 SELECT ?page ?name
4 WHERE {
5     ?page a "city" .
6     ?page wiki: name ?name .
7     ?page wiki: largestSettlementOf wiki: organisation:eu .
8 }
9 </pl>
```

REASONING UNIT TESTS

LOKI: 2. SPECIFY THE EXPECTED QUERY RESULT

```

1 <pl format="sparql">
2 PREFIX wiki: <>
3 SELECT ?page ?name
4 WHERE {
5     ?page a "city" .
6     ?page wiki: name ?name .
7     ?page wiki: largestSettlementOf wiki: organisation:eu .
8 }
9 </pl>

```

```

1 [[unittest_assert_anequal:?name:Paris|Is Paris one of the
   largests?]]
2 [[unittest_assert_noneequal:?name:Cracow|Cracow is not one
   of the largests!]]

```

REASONING UNIT TESTS

LOKI: 3. CHECK THE RESULTS

Test results

Query result

page	name
london	Londinium
london	London
paris	Paris

Test: anyequal; Field name; Value: Paris Result: PASSED

Test: noneequal; Field name; Value: Cracow Result: PASSED

REASONING UNIT TESTS

LOKI: 3. CHECK THE RESULTS

Test results

Query result

page	name
london	Londinium
london	London
paris	Paris

Test: anyequal; Field name; Value: Paris Result: PASSED

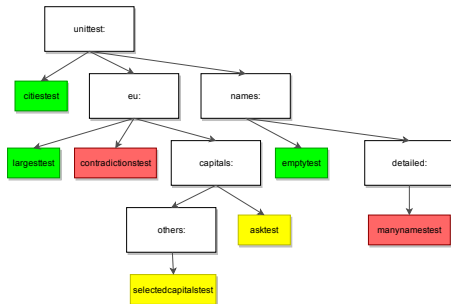
Test: noneequal; Field name; Value: Cracow Result: PASSED

Reasoning unit tests results:

Test	Status
citytest	PASSED
eu:capitals:asktest	NOT EXECUTED
eu:capitals:others:selectedcapitaltest	NOT EXECUTED
eu:contradictiontest	FAILED
eu:largesttest	PASSED
names:detailed:manynametest	FAILED
names:emptytest	PASSED

REASONING UNIT TESTS

LOKI: 3. CHECK THE RESULTS



Reasoning unit tests results:

Test	Status
citytest	PASSED
eu:capitals:asktest	NOT EXECUTED
eu:capitals:others:selectedcapitaltest	NOT EXECUTED
eu:contradictiontest	FAILED
eu:largesttest	PASSED
names:detailed:manynametest	FAILED
names:emptytest	PASSED

METRICS OF CHANGES

EXISTING ONTOLOGY METRICS

- **Attribute Richness:** an average number of attributes *att* per class *C*:

$$AR = \frac{|att|}{|C|} \quad (1)$$

- **Average Population:** the ratio of instances *I* to classes *C*:

$$AP = \frac{|I|}{|C|} \quad (2)$$

- **Size Of Vocabulary:** counts the vocabulary used in the ontology (instances *I* and attributes *att*):

$$SOV = |I| + |att| \quad (3)$$

- **Edge Node Ratio:** a ratio of all edges *E* (the number of all triples) to all nodes *N* (subjects and objects, both named and literals):

$$ENR = \frac{|E|}{|N|} \quad (4)$$

METRICS OF CHANGES

CHANGE AND NORMALIZATION

Metric	1	2	3	4	5
<i>AR</i>	$(-\infty; -0.4)$	$[-0.4; -0.1)$	$[-0.1; 0.1]$	$(0.1; 0.4]$	$(0.4; \infty)$
<i>AP</i>	$(-\infty; -0.4)$	$[-0.4; -0.1)$	$[-0.1; 0.1]$	$(0.1; 0.4]$	$(0.4; \infty)$
<i>SOV</i>	$(-\infty; -4)$	$[-4; -1)$	$[-1; 1]$	$(1; 4]$	$(4; \infty)$
<i>ENR</i>	$(-\infty; -0.4)$	$[-0.4; -0.1)$	$[-0.1; 0.1]$	$(0.1; 0.4]$	$(0.4; \infty)$

METRICS OF CHANGES

WEIGHTED AVERAGE

1. To promote **fast KB expansion**: based on *Size of Vocabulary*, which directly measures the database growth.

$$WA_1 = \frac{AR + AP + 5 \cdot SOV + ENR}{8} \quad (5)$$

2. To put emphasis on **KB coherence**: combine the values of *Attribute Richness*, *Average Population* and *Edge Node Ratio*. In this case classes with more attributes and instances are promoted, as they are indicators of richer KB.

$$WA_2 = \frac{5 \cdot AR + 5 \cdot AP + SOV + 3 \cdot ENR}{14} \quad (6)$$

METRICS OF CHANGES

THE 1st CHANGE IN THE EUROPEAN WIKI

```

1 ===== London =====
2
3 [[name:=London]] is the capital and most populous [[category:city|city]] of [[
  capitalOf::country:england|England]] and the [[capitalOf::country:uk|United
  Kingdom]].
4
5 Standing on the River Thames in the south east of the island of Great Britain,
  London has been a major settlement for two millennia. It was founded by the
  Romans, who named it [[name:=Londinium]].
6
7 London has a diverse range of people and cultures, and more than 300 languages
  are spoken in the region. Its estimated mid-2016 municipal population (
  corresponding to Greater London) was [[population:=8787892|8,787,892]], the
  largest of any city in the [[largestSettlementOf::organisation:eu|European
  Union]], and accounting for 13.4% of the UK population. London's urban
  area is the second most populous in the EU, after Paris, with 9,787,426
  inhabitants at the 2011 census. The city's metropolitan area is the most
  populous in the EU with 13,879,757 inhabitants, while the Greater London
  Authority states the population of the city-region (covering a large part
  of the south east) as 22.7 million. London was the world's most populous
  city from around 1831 to 1925.
```

$$\triangleright AR = \frac{|att|}{|C|} = \frac{4}{1} = 4.00$$

$$\triangleright AP = \frac{|I|}{|C|} = \frac{1}{1} = 1.00$$

$$\triangleright SOV = |C| + |I| + |att| = 1 + 1 + 4 = 6$$

$$\triangleright ENR = \frac{|E|}{|N|} = \frac{7}{8} = 0.88$$

METRICS OF CHANGES

THE 1st CHANGE IN THE EUROPEAN WIKI

```

1 ===== London =====
2
3 [[name:=London]] is the capital and most populous [[category:city|city]] of [[
  capitalOf::country:england|England]] and the [[capitalOf::country:uk|United
  Kingdom]].
4
5 Standing on the River Thames in the south east of the island of Great Britain,
  London has been a major settlement for two millennia. It was founded by the
  Romans, who named it [[name:=Londinium]].
6
7 London has a diverse range of people and cultures, and more than 300 languages
  are spoken in the region. Its estimated mid-2016 municipal population (
  corresponding to Greater London) was [[population:=8787892|8,787,892]], the
  largest of any city in the [[largestSettlementOf::organisation:eu|European
  Union]], and accounting for 13.4% of the UK population. London's urban
  area is the second most populous in the EU, after Paris, with 9,787,426
  inhabitants at the 2011 census. The city's metropolitan area is the most
  populous in the EU with 13,879,757 inhabitants, while the Greater London
  Authority states the population of the city-region (covering a large part
  of the south east) as 22.7 million. London was the world's most populous
  city from around 1831 to 1925.
```

$$\blacktriangleright AR = \frac{|att|}{|C|} = \frac{4}{1} = 4.00$$

$$\blacktriangleright AP = \frac{|I|}{|C|} = \frac{1}{1} = 1.00$$

$$\blacktriangleright SOV = |C| + |I| + |att| = 1 + 1 + 4 = 6$$

$$\blacktriangleright ENR = \frac{|E|}{|N|} = \frac{7}{8} = 0.88$$

METRICS OF CHANGES

THE 1st CHANGE IN THE EUROPEAN WIKI

Metric	1	2	3	4	5
<i>AR</i>	$(-\infty; -0.4)$	$[-0.4; -0.1)$	$[-0.1; 0.1]$	$(0.1; 0.4]$	$(0.4; \infty)$
<i>AP</i>	$(-\infty; -0.4)$	$[-0.4; -0.1)$	$[-0.1; 0.1]$	$(0.1; 0.4]$	$(0.4; \infty)$
<i>SOV</i>	$(-\infty; -4)$	$[-4; -1)$	$[-1; 1]$	$(1; 4]$	$(4; \infty)$
<i>ENR</i>	$(-\infty; -0.4)$	$[-0.4; -0.1)$	$[-0.1; 0.1]$	$(0.1; 0.4]$	$(0.4; \infty)$

$$\blacktriangleright AR = \frac{|att|}{|C|} = \frac{4}{1} = 4.00 \text{ (5)}$$

$$\blacktriangleright AP = \frac{|I|}{|C|} = \frac{1}{1} = 1.00 \text{ (5)}$$

$$\blacktriangleright SOV = |C| + |I| + |att| = 1 + 1 + 4 = 6 \text{ (5)}$$

$$\blacktriangleright ENR = \frac{|E|}{|N|} = \frac{7}{8} = 0.88 \text{ (5)}$$

OPINIONS AND DISCUSSION

No possibility to
automatically check all
quality aspects → user
evaluation is needed:

- ▶ Subjective
assessment of
changes (1-5 stars)
- ▶ A place for
discussion and
comments

OPINIONS AND DISCUSSION

No possibility to automatically check all quality aspects → user evaluation is needed:



- ▶ Subjective assessment of changes (1-5 stars)
- ▶ A place for discussion and comments

[city:paris_1508153067](#) created — yoda (*Master Yoda*) 2017/10/16 13:24

★ ★ ☆ ☆ ☆ from 1 votes
☐ ☐ ☐ ☐ ☐ Rate



[Export to RDF](#)

Discussion

 Krzysztof Kutt,  2017/10/16 17:41



You have created a good page about Paris, but I do not understand why you assigned it to three categories city, City and town?

[Export to RDF](#)

 Master Yoda,  2017/10/16 17:43



I did not know what category we would ultimately choose, so I decided to write a few that are potential options...

[Reply](#) [Export to RDF](#)

 Krzysztof Kutt,  2017/10/16 17:45

OK, I understand. Let's use the category city – it seems to be the most logical, because we already put the pages in a city namespace 🙄

[Export to RDF](#)

 Master Yoda,  2017/10/16 18:01

Sure. I will fix it!

I will also add some annotations to your London page, because I have some ideas how to do this.

[Reply](#) [Export to RDF](#)

Enter your comment. Wiki syntax is allowed:

[!\[\]\(881738dffb134c43b162e8cf846b0cce_img.jpg\)](#) [!\[\]\(5089ca6f12aee786c55691a4fad90c23_img.jpg\)](#) [!\[\]\(cf391b03be465bca008f0586fc582556_img.jpg\)](#) [!\[\]\(2ed3f1d2b8cf64584ba6c2ade3dd3f2e_img.jpg\)](#) [!\[\]\(7ba86b82a5e8efa7ab5cd9fe6253481a_img.jpg\)](#) [!\[\]\(cdfa2826830e42fe873188c27c78b403_img.jpg\)](#) [!\[\]\(a3f57ab912916ce6fe231544d0442f33_img.jpg\)](#) [!\[\]\(c08cc02b346871700d09660bd9cc08d1_img.jpg\)](#) [!\[\]\(cd7350e79591c702884e5fd4fa0fa861_img.jpg\)](#) [!\[\]\(f5d1e39c2769f4ed99fd943ed29a7bb3_img.jpg\)](#) [!\[\]\(27d766c3fde1fbaeb4900da66f41bad3_img.jpg\)](#) [!\[\]\(ce7f66d39ffcc3f831f3dd77b6e65048_img.jpg\)](#)

☐ Subscribe to comments

CHANGE MANAGEMENT

Methods and Tools

Quality Management

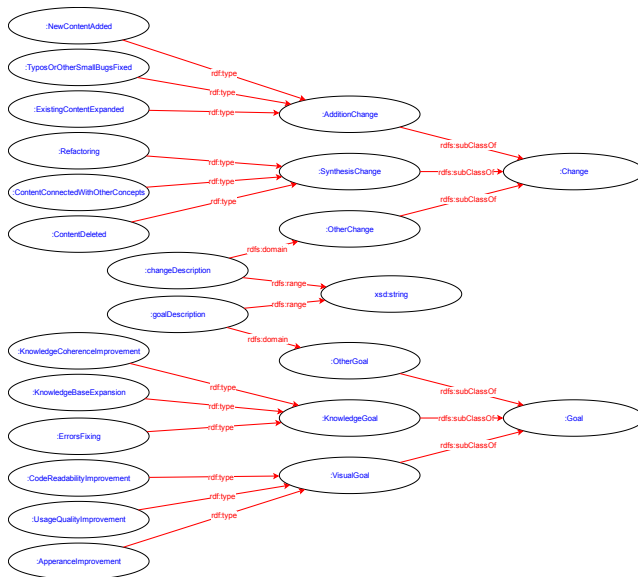
Change Management

User Involvement

CHANGE ONTOLOGIES

- ▶ Cover two aspects of change:
 - ▶ **the factual change** (What was done? e.g. Typos or other small bugs fixed, New content added)
 - ▶ **the goal** (Why it was done? e.g. Errors fixing, Knowledge database expansion)
- ▶ Created for specific purposes, e.g. for preparing conference papers

CHANGE ONTOLOGIES



CHANGE ONTOLOGIES

EXTENDED WIKI EDITION FORM

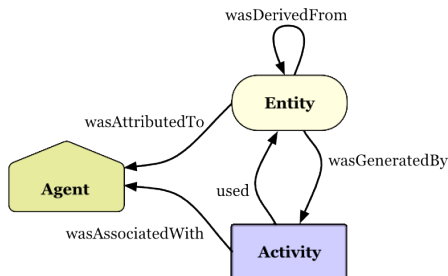
The screenshot shows the top of a Wikipedia article titled "Harrison Ford". The edit summary is "Harrison Ford biography updated". The "What have you done?" section is expanded, showing the following changes:

- (1) What have you done? Existing content expanded
- (2) Why have you done this? Other... Test for documentation
- (3) Which resources have you used? (one URL/URL in one line): Double-click to select wiki page as a resource:
 - lokipage:movies:kingdom_of_the_crystal_skull
 - http://www.imdb.com/name/nm0000148/

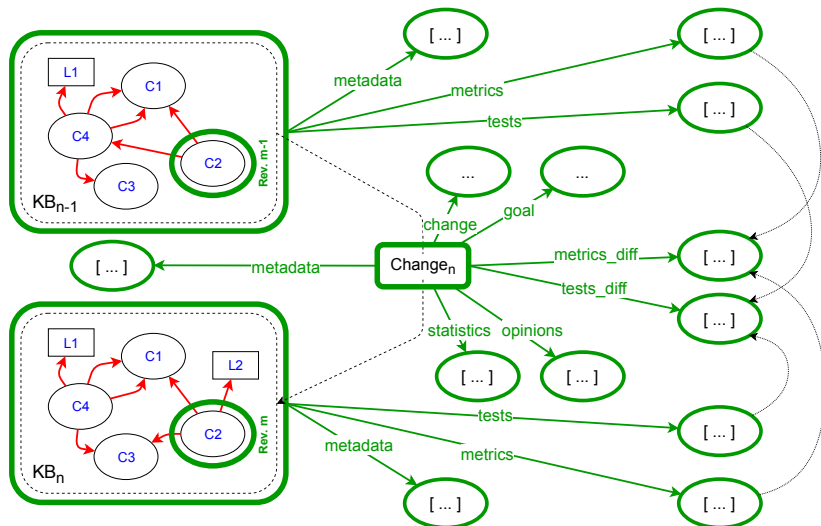
The "lokipage:movies:kingdom_of_the_crystal_skull" link is highlighted in red. The "http://www.imdb.com/name/nm0000148/" link is highlighted in yellow. The "Double-click to select wiki page as a resource:" text is highlighted in blue.

SEMANTIC CHANGELOG

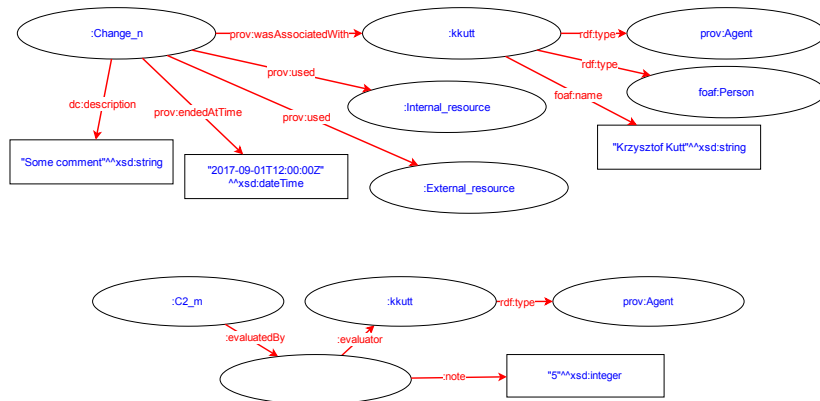
- ▶ Graph-based changelog (RDF grounded)
- ▶ Describes the relations between Agents (users), Activities (created, edit, delete) and Entities (wiki pages, revisions of pages)
- ▶ Collects all available information: metadata, tests results, metrics, opinions, change ontologies, ...



SEMANTIC CHANGELOG



SEMANTIC CHANGELOG



SEMANTIC CHANGELOG

IDENTIFICATION OF USERS ASSOCIATED WITH LOW QUALITY CHANGES

```

1 SELECT ?user (COUNT(?user) as ?maliciousChanges)
2 WHERE {
3     ?change a prov:Activity ;
4         loki:testsPassed [ loki:valueBefore ?testsBefore ;
5                             loki:valueAfter ?testsAfter ] ;
6         loki:weightedAverage ?metric ;
7         prov:wasAssociatedWith ?user .
8     FILTER (?testsBefore > ?testsAfter) .
9     FILTER (?metric < 3.0) .
10 }
11 GROUP BY ?user
12 ORDER BY DESC(?maliciousChanges)
13 LIMIT 5

```

SEMANTIC CHANGELOG

USERS' TYPES IDENTIFICATION

```

1  SELECT ?user ?type
2  WHERE {
3      {
4          SELECT ?user (COUNT(?user) AS ?additions)
5          WHERE {
6              ?change rdf:type prov:Activity ;
7                  loki:whatWasDone ?changeType ;
8                  prov:wasAssociatedWith ?user .
9              ?changeType rdf:type change:AdditionChange .
10         }
11         GROUP BY ?user
12     }
13     UNION
14     {
15         SELECT ?user (COUNT(?user) AS ?syntheses)
16         WHERE {
17             ?change rdf:type prov:Activity ;
18                 loki:whatWasDone ?changeType ;
19                 prov:wasAssociatedWith ?user .
20             ?changeType rdf:type change:SynthesisChange .
21         }
22         GROUP BY ?user
23     }
24     BIND(if(( ?additions > ?syntheses ), "Adder", "Synthesizer") AS ?type)
25 }

```

USER INVOLVEMENT

Methods and Tools

Quality Management

Change Management

User Involvement

GAMIFICATION



GAMIFICATION

Adders vs. Synthesizers?

GAMIFICATION

The simple *points* (pts) system:

- ▶ Login to the system – 5 pts once a day,
- ▶ New concept added – 25 pts,
- ▶ Concept edition – 5 pts.

Points and *badges* as a result of achievements:

- ▶ For logging in for X days in a row:

X	Points	Badge
1	10	Welcome to the Y
2	15	Y is Fun
3	15	Y is Really Fun
7	25	Y for a Whole Week
14	40	Two Weeks in a Row?
30	40	You Have to be Addicted

- ▶ For creating X new concepts ...
- ▶ For X editions ...

Levels calculated according to the equation (value is rounded to the tens):

$$P_{L+1} = 40L^{\frac{5}{3}} + 10 \quad (7)$$

GAMIFICATION

Adders vs. Synthesizers?

GAMIFICATION

- ▶ *Bounty* system – challenges (“parse the source Y into the KB”) and rewards
- ▶ *Natural tests* mechanism – random changes and concepts: mark them as good or bad
- ▶ *User’s credibility* – average of:
 - ▶ ratio of reasoning unit tests passed to all tests
 - ▶ weightedAverage metric
 - ▶ mean value of votes given
 - ▶ ratio of positive natural tests values to all ratings given

Threshold of 70% → the “Trusted user” badge

GAMIFICATION

EUROPEAN WIKI

Change	User	Achievements
Ch_1	kkutt	Login to a system: 5 pts; New concept added: 25 pts; Logging in for 1 day: 10 pts + Badge "Welcome to the European Wiki"; Creation of 1 concept: 10 pts + Badge "My First Concept"
Ch_2	yoda	Login to a system: 5 pts; New concept added: 25 pts; Logging in for 1 day: 10 pts + Badge "Welcome to the European Wiki"; Creation of 1 concept: 10 pts + Badge "My First Concept"
Ch_3	kkutt	New concept added: 25 pts
Ch_4	yoda	Concept edition: 5 pts; Edition of 1 concept: 10 pts + Badge "Going Down in History"
Ch_5	yoda	Concept edition: 5 pts
Ch_6	kkutt	Concept edition: 5 pts; Edition of 1 concept: 10 pts + Badge "Going Down in History"

GAMIFICATION

EUROPEAN WIKI

User	Ratio of unit tests passed	Weighted average metric	Votes	Final Credibility
kkutt	$\frac{2.33}{6}$ (44%)	3.71 (74%)	N/A	59%
yoda	$\frac{3}{6}$ (50%)	3.91 (78%)	4 (80%)	70%

TOP 10

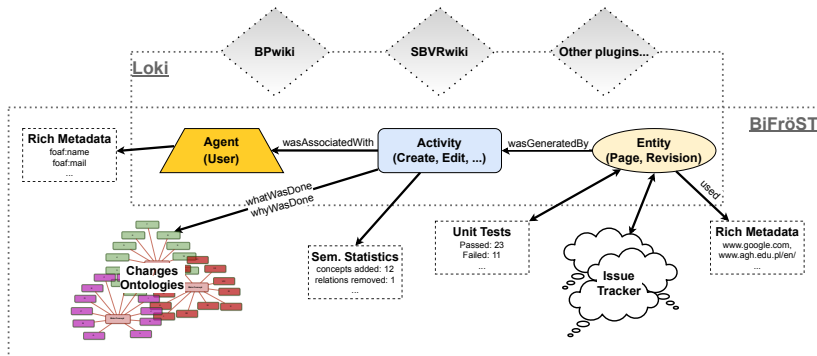
1.	kkutt (2)	90 pts
2.	yoda (2)	70 pts

city:paris

Paris

Paris is the capital and most populous [city](#) of [France](#), with an administrative-limits area of 105 square kilometres (41 square miles) and a 2015 population of 2,229,621. The city is a commune and department, and the capital-heart of the 12,012-square-kilometre (4,638-square-mile) Ile-de-

BiFröST FRAMEWORK FOR SEMANTICAL TRACKING



USABILITY

ONTOLOGY STORAGE + CODE HINT AND COMPLETION

Ontology name: Multimedia 

Classes

ID	Name	
item	Multimedia Item	edit :: delete
book	Book	edit :: delete
movie	Movie	edit :: delete
human	Human being	edit :: delete
actor	Actor	edit :: delete
author	Author	edit :: delete
ID	Name	Save

Class relations

Relation	Subject ID	Object ID	
subClassOf	book	item	delete
subClassOf	movie	item	delete
subClassOf	actor	human	delete
subClassOf	author	human	delete
Relation	Subject ID	Object ID	Save

Object properties

Property ID	Subject ID	Object ID	
hasAuthor	book	author	delete
hasActor	movie	actor	delete
isAssociatedWith	human	item	delete
Property ID	Subject ID	Object ID	Save

Data properties

Property ID	Domain	Range	
hasTitle	item	xsd:string	delete
hasName	human	xsd:string	delete
Property ID	Domain	Range	Save

Property relations

Property ID	Subject ID	Object ID	
subPropertyOf	hasAuthor	isAssociatedWith	delete
subPropertyOf	hasActor	isAssociatedWith	delete
Property ID	Subject ID	Object ID	Save

[[category:|

actor
author
book
human
item
movie
animal

[[category:actor]]
[[category:actress]]
|

“Cops” are no longer needed!

OVERVIEW

Introduction

CKE Process

Methods and Tools

Evaluation

Summary

5 EXPERIMENTS

<i>Module (plugin) / Feature</i>	Experiment				
	1st	2nd	3rd	4th	Final
Loki	+	+	+	+	+
<i>Reasoning unit tests</i> (Loki)					+
<i>Metric of changes</i> (prov)				+	+
<i>Opinions and discussion</i> (revisionsrater)					+
<i>Change ontologies</i> (prov)		+	+	+	+
<i>Semantic changelog</i> (prov)		+	+	+	+
<i>Gamification</i> (wikigame)					+
Usability extensions (Loki)			+		+
CKE agile process	+	+	+	+	+
Usability evaluation (SUMI)					+
Comparison with SMW				+	

4 SUPPORTING EXPERIMENTS

1ST. POKEMONS, SIMPSONS, ET AL.

	Phase 1: Creation	Phase 2: Annotation	Phase 3: Evaluation
Team 1	Pokemons	Simpsons	Dragon Ball Z
Team 2	Drinks	Dragon Ball Z	Simpsons
Team 3	Simpsons	Pokemons	Drinks
Team 4	Dragon Ball Z	Drinks	Pokemons

- ▶ 6 weeks
- ▶ 8 students
- ▶ 3000 changes
- ▶ 1891 triples
- ▶ 345 wiki pages

The most important conclusions:

- ▶ experts and engineers **must** collaborate at every stage
- ▶ need for more strictly controlled conditions

4 SUPPORTING EXPERIMENTS

2ND: CSP LIBRARY

► simple iterative process:

1. competency questions
2. implementation
3. evaluation

► 5 iterations

► 13 students

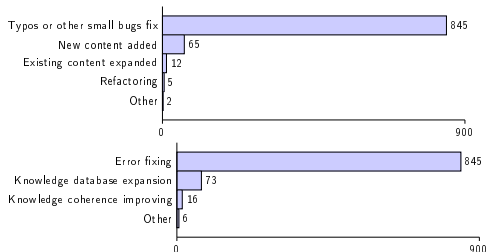
► 1186 changes

► 821 triples

► 265 wiki pages

The most important conclusions:

- important steps and roles in process identified
- reasoning unit tests idea proposed



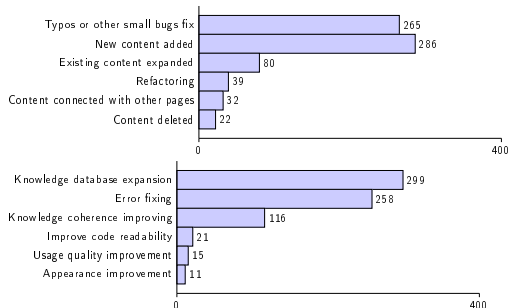
4 SUPPORTING EXPERIMENTS

3RD: PUBS IN CRACOW

- ▶ updated iterative process
- ▶ 5 iterations
- ▶ 15 students
- ▶ 665 changes
- ▶ 1031 triples
- ▶ 202 wiki pages

The most important conclusions:

- ▶ updated change ontology seems to be good enough
- ▶ a need for task board identified
- ▶ new version of plugin forced users think about change type and goal:



4 SUPPORTING EXPERIMENTS

4TH: ARTIFICIAL INTELLIGENCE CLASS

	Changes	Wiki pages	Triples
Wiki A (Loki)	2417	177	1113
Wiki B (SMW)	2629	108	1422

The most important conclusions:

- ▶ Semantic MediaWiki vs Loki
- ▶ 56 (A) + 49 (B) students
 - ▶ 1 hour introductory training is not enough?
 - ▶ group of 50 people is too big for self-management
 - ▶ a strong leader or subgroup who would make any decision had not emerged
 - ▶ motivation: focused on educational process itself
 - ▶ no significant differences between reported difficulty for SMW and Loki

FINAL EXPERIMENT

COOKBOOK AND MOVIES KB

- ▶ 16 students (Cookbook),
5 iterations
- ▶ 5 knowledge engineers
(online call, Movies KB)

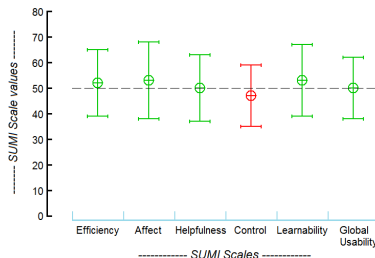
The most important
conclusions:

- ▶ even simple iteration board
is a great idea
- ▶ there were no “cops”
(thanks to the hint and
completion mechanism)
- ▶ conflicts still appeared
(“egg” vs “eggs”?) –
resolved via discussion

FINAL EVALUATION

- ▶ Observation and user questionnaires: process was clear and robust
- ▶ Each module works good enough
- ▶ All modules of BiFröST framework smoothly interact with each other
- ▶ Users generally felt satisfied (SUMI inventory)

SUMI Scale Profiles: Means with Standard Deviations



<i>Semantic Wiki</i>	<i>Requirements (max 26)</i>
SMW	17
KnowWE	17
OntoWiki	14
Loki (before PhD)	14
Loki (with BiFröST)	23

OVERVIEW

Introduction

CKE Process

Methods and Tools

Evaluation

Summary

SUMMARY

OBJECTIVES

*The main goals were to **describe a Collaborative Knowledge Engineering process** that provides a general framework for defining roles which should be identified in a group and steps that should be taken in this process, as well as to propose **methods and tools that support the defined CKE process**, leading to the creation of good quality KB in reasonable time, through the means that will be convenient for target users.*

SUMMARY

The most important results:

- ▶ Analysis of **issues and challenges** for CKE
- ▶ Definition of **CKE agile process**
- ▶ Conceptualization of **change ontology**
- ▶ Proposal of **graph-based semantic changelog**
- ▶ Formulation of **involvement metrics**
- ▶ Implementation of **prototypical toolkit** (BiFröST)

Future works:

- ▶ (semi-)automation: machine learning model that supports selection of change ontology concepts
- ▶ sources credibility and other usage scenarios for semantic changelog
- ▶ application of presented CKE process in digital humanities and software engineering

SUMMARY

The most important results:

- ▶ Analysis of **issues and challenges** for CKE
- ▶ Definition of **CKE agile process**
- ▶ Conceptualization of **change ontology**
- ▶ Proposal of **graph-based semantic changelog**
- ▶ Formulation of **involvement metrics**
- ▶ Implementation of **prototypical toolkit** (BiFröST)

Future works:

- ▶ (semi-)automation: machine learning model that supports selection of change ontology concepts
- ▶ sources credibility and other usage scenarios for semantic changelog
- ▶ application of presented CKE process in digital humanities and software engineering

PAPERS

- ▶ Grzegorz J. Nalepa, Krzysztof Kutt & Szymon Bobek (2018). *Mobile Platform for Affective Context-Aware Systems*. FGCS Journal.
- ▶ Krzysztof Kutt, Wojciech Binek, Piotr Misiak, Grzegorz J. Nalepa & Szymon Bobek (2018). *Towards the Development of Sensor Platform for Processing Physiological Data From Wearable Sensors*. ICAISC 2018.
- ▶ Grzegorz J. Nalepa, Krzysztof Kutt, Szymon Bobek & Barbara Giżycka (2018). *Development of Mobile Platform for Affect Interpretation*. Current Progress. AfCAI 2018, CEUR-WS.
- ▶ Krzysztof Kutt, Grzegorz J. Nalepa & Dominik Burdzy (2017). *Semantically Enhanced Navigation System Using Augmented Reality*. MCSS 2017.
- ▶ Grzegorz J. Nalepa, Barbara Giżycka, Krzysztof Kutt & Jan Argasiński (2017). *Affective Design Patterns in Computer Games*. Scrollrunner Case Study. Communication Papers of the FedCSIS 2017.
- ▶ Mateusz Ślażyński, Grzegorz J. Nalepa, Szymon Bobek & Krzysztof Kutt (2017). *Semantic Annotations for Mediation of Complex Rule Systems*. ICAISC 2017.
- ▶ Krzysztof Kluza, Grzegorz J. Nalepa, Mateusz Ślażyński, Krzysztof Kutt, Edyta Kucharska, Krzysztof Kaczor & Adam Luszpaj (2017). *Overview of Selected Business Process Semantization Techniques*. In: Tomasz Pelech-Pilichowski, Maria Mach-Król & Celina M. Olszak (Ed.). *Advances in Business ICT: New Ideas from Ongoing Research*, Springer.
- ▶ Krzysztof Kutt (2016). *Semantic Wikis Versioning with BiFröST*. Doctoral Consortium of AI*IA 2016.
- ▶ Grzegorz J. Nalepa, Jan K. Argasiński, Krzysztof Kutt, Paweł Węgrzyn, Szymon Bobek & Mateusz Z. Lepicki (2016). *Affective Computing Experiments in Virtual Reality with Wearable Sensors*. Methodological considerations and preliminary results. AfCAI 2016, CEUR-WS.
- ▶ Grzegorz J. Nalepa, Krzysztof Kutt, Szymon Bobek & Mateusz Z. Lepicki (2016). *AfCAI systems: Affective Computing with Context Awareness for Ambient Intelligence*. Research proposal. AfCAI 2016, CEUR-WS.
- ▶ Grzegorz J. Nalepa, Mateusz Ślażyński, Krzysztof Kutt, Edyta Kucharska & Adam Luszpaj (2015). *Unifying Business Concepts for SMEs with Prosecco Ontology*. Proceedings of the 2015 FedCSIS.
- ▶ Krzysztof Kutt, Artur Gunia & Grzegorz J. Nalepa (2015). *Cognitive Enhancement: How to Increase Chance of Survival in the Jungle?* CYBCONF 2015.
- ▶ Krzysztof Kluza, Krzysztof Kutt & Marta Woźniak (2014). *SBVRwiki (Tool Presentation)*. Proceedings of KESE10.
- ▶ Grzegorz J. Nalepa, Krzysztof Kutt & Krzysztof Kaczor (2014). *Can the Generation of Test Cases for Unit Testing be Automated with Rules?* ICAISC 2014.
- ▶ Krzysztof Kaczor, Grzegorz J. Nalepa & Krzysztof Kutt (2013). *HaDEclipse – Integrated Environment for Rules (Tool Presentation)*. Proceedings of KESE9.