# Semantic Technologies in Idea Management Systems:
## A Model for Interoperability, Linking and Filtering

PhD candidate: Adam Westerski

Supervisor: Carlos A. Iglesias

Grupo de Sistemas Inteligentes
Departamento de Ingenieria de Sistemas Telematics
Universidad Politecnica de Madrid

February 2013

# Table of Contents

# What is Idea Management?

## [Introduction]

- Origins
  - Open Innovation | ask others for innovation
    Chesbrough, 2003
  - Crowdsoucring | connect to large audience
    Howe, 2004
- Goals
  - communicate with community and allow deliberation
    Bailey and Horvitz, 2010
  - select best ideas available
    Bailey and Horvitz, 2010; Hrastinski, 2010
- Motivations for use in organizations
  - react to fast changing markets & customer profiles
    Bailey and Horvitz, 2010
  - capture wider scope of ideas (as opposed to closed innovation)
    Gassmann, 2006; Riedl 2009

# Idea Management Problems

[Introduction: thesis research motivation]

- Information overflow | lots of data gathered over long time
  Jouret 2009; Belecheanu 2009
- Noisy data | similar ideas, lots of simple or obvious input
  Kornish 2010; Jouret 2009; Belecheanu 2009
- Peaks of data | lots of ideas over short period of time
  Baumgartner 2008; Turrell 2002
- Rating innovation | lack of proper metrics, lot of effort required
  Hrastinski 2010; Gangi 2009

**Approach:** research on new data modeling approaches and analysis of gathered metadata

# Different perspectives on IMS

## [Introduction: IMS examples]

www.ideastorm.com

**IdeaStorm**
Where Your Ideas Reign

There are currently no Storm Sessions active.
Stay tuned!

IdeaStorm | Storm Sessions

**View** All posted ideas by the community

**Post** Your idea for Dell products or services

**Vote** Promote or demote ideas

**See** Your ideas in action

Sort Ideas By: Popular Ideas | Recent Ideas | Top Ideas — All

Comments

118030 Promote Demote
**Pre-Installed OpenOffice | alternative to MS Works & MS Office**
By dhart, Feb 17, 2007

Provide OpenOffice.org for free pre-installation alongside Microsoft Works and Microsoft Office. OpenOffice.org is more capable than Microsoft Works, and a serious competitor to Microsoft Office, at a fraction of the cost (it's free!)

OpenOffice.org can open, create, edit and save Microsoft Word, Excel and PowerPoint files.

106450 Promote Demote
**Pre-Installed Linux | Ubuntu | Fedora | OpenSUSE | Multi-Boot**
By dhart, Feb 16, 2007

Offer the 3 top *free* **Linux** versions for *free* pre-installation on all Dell PCs.

Quality free and open source software drastically lowers the cost of new PCs, and helps prevent

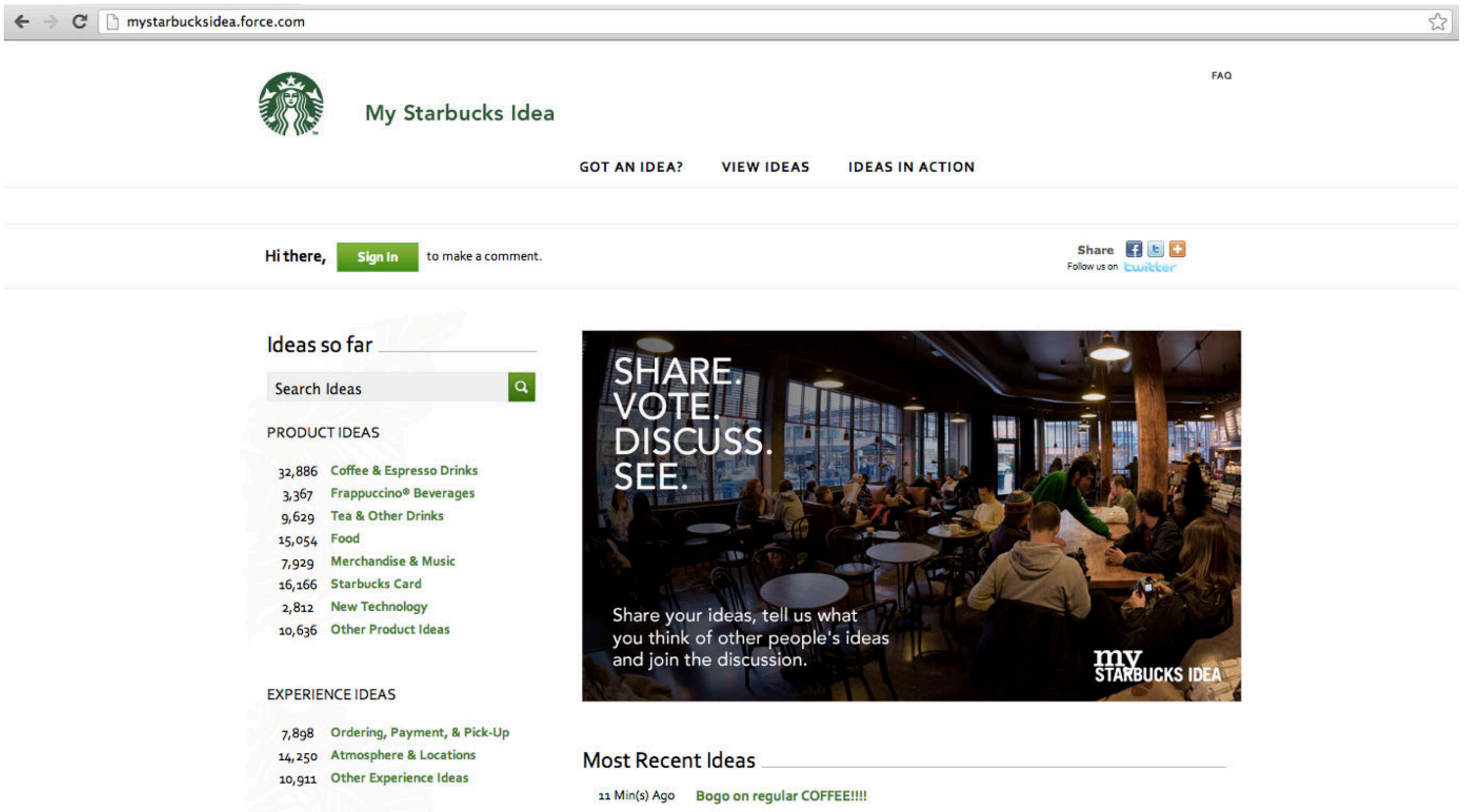
Login to IdeaStorm
Don't have an IdeaStorm account? Register Now.
Username:
Password:
Login
Forgot Password?

**Your Ideas in Action**

**IdeaStorm Recap - 1/22/2010**
Happy 2010 everyone! I know I'm a little late with the holiday greetings, but there is a lot to share on...

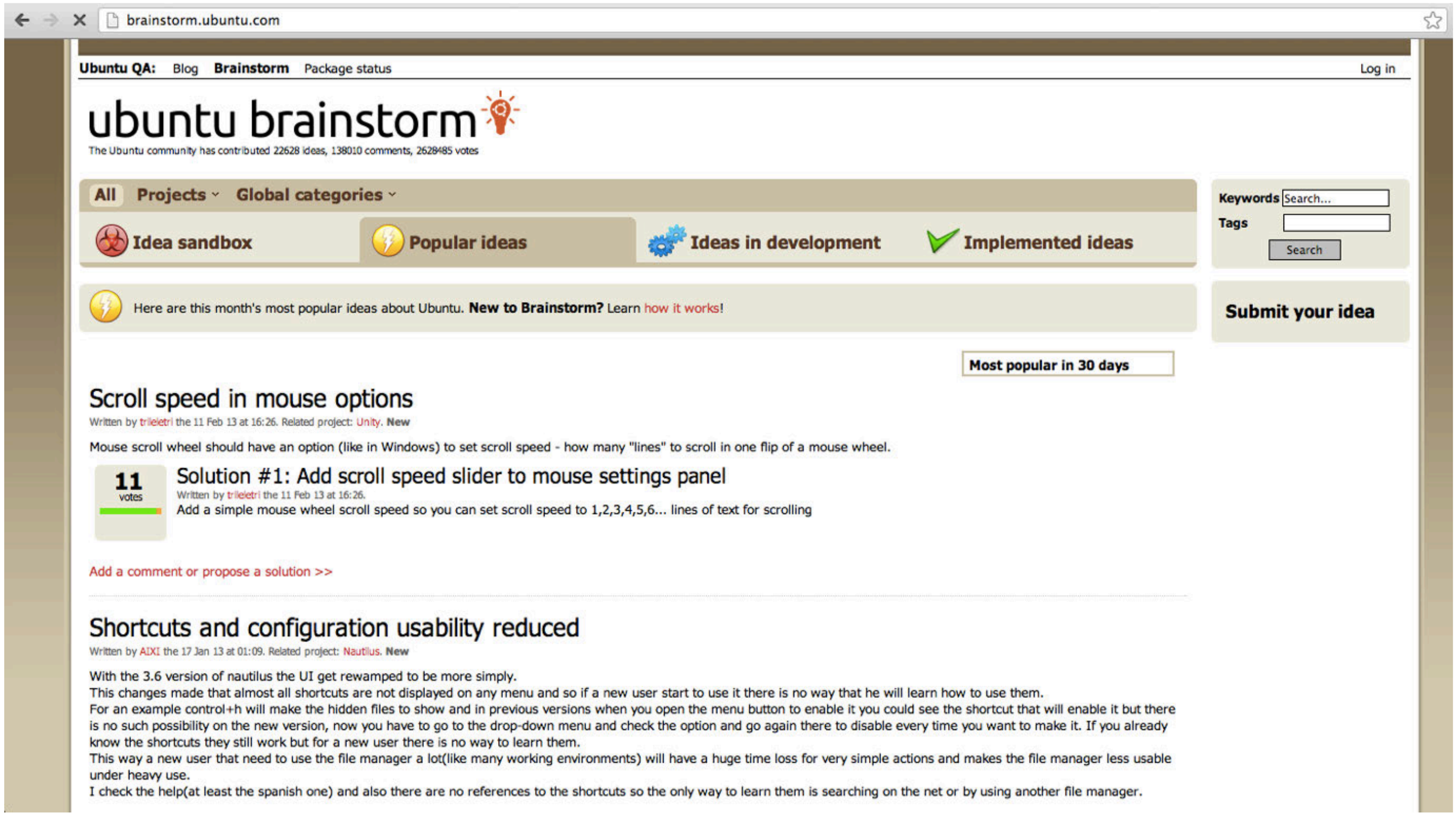
# Different perspectives on IMS

## [Introduction: IMS examples]



# Different perspectives on IMS

## [Introduction: IMS examples]



# Thesis objectives

[Introduction: from motivation to goals]

- propose a **single conceptual model** for Idea Management Systems
- **summarize data** of Idea Management Systems
- deliver **indicators for idea assessment**

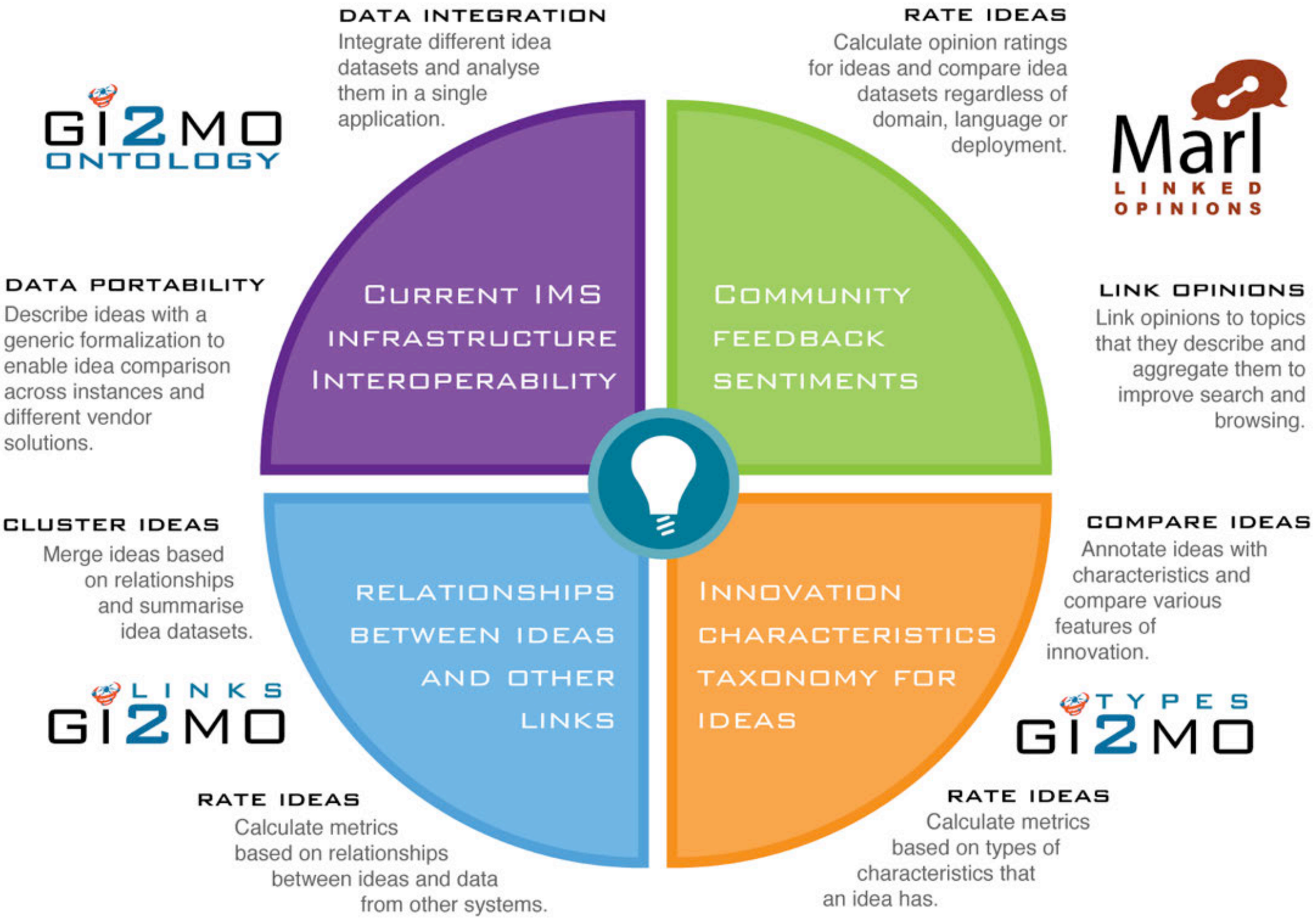
# Thesis research questions

[Introduction: research motivation]

- *can the contemporary IMSes be generalized into a* ***single model*** *?*
- *can* ***community activity*** *related to ideas be modeled, summarized and measured regardless of the system or domain ?*
- *can ideas be modeled,* ***summarized and compared*** *independently of the domain or use of IMS ?*
- *can the content of IMS be summarized based on the basis of* ***relationships between ideas*** *?*

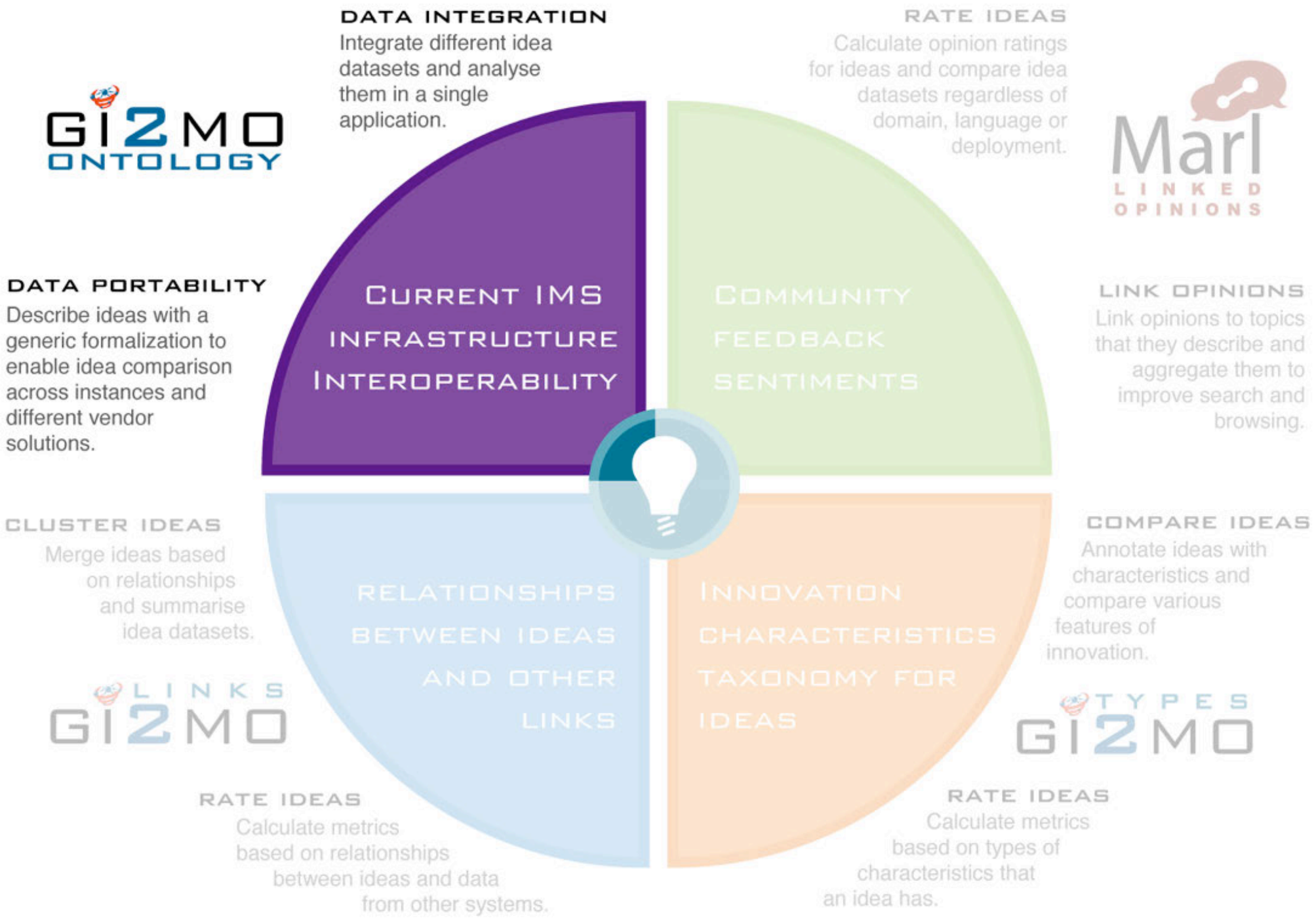
# Thesis Contribution Areas

## [Introduction: solution architecture]



# Generic Model for Idea Management

## [Solution - Part I]



# Generic Model for Idea Management

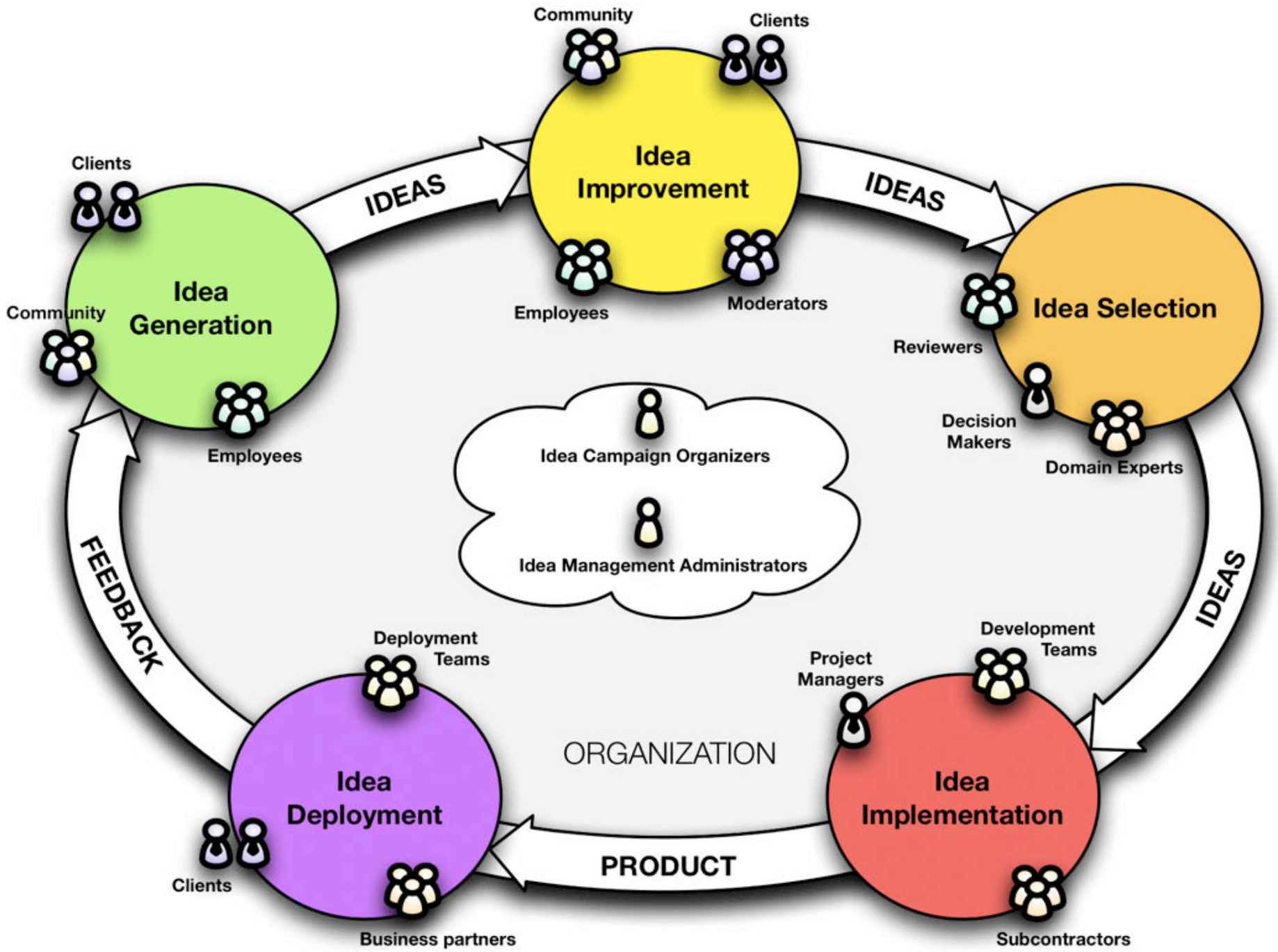
[Solution - Part I]

**Methodology:**

1. Analyse Industrial Systems and State of the Art in research
2. Capture common elements and dynamics in a form of **Idea Life Cycle**
3. List all **common data properties** for each stage of the Idea Life Cycle
4. Formalize the Life Cycle as an **ontology**
5. Evaluation for multiple systems

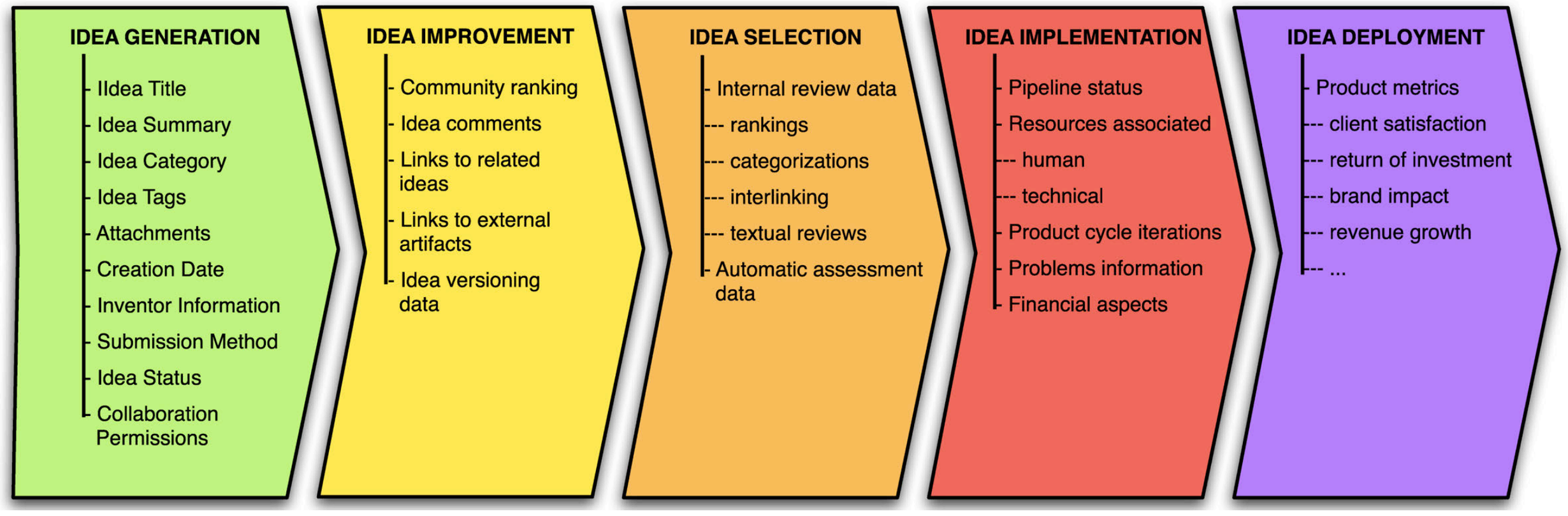
# Idea Life Cycle

## [Solution Part I: Generic Model for Idea Management Systems]



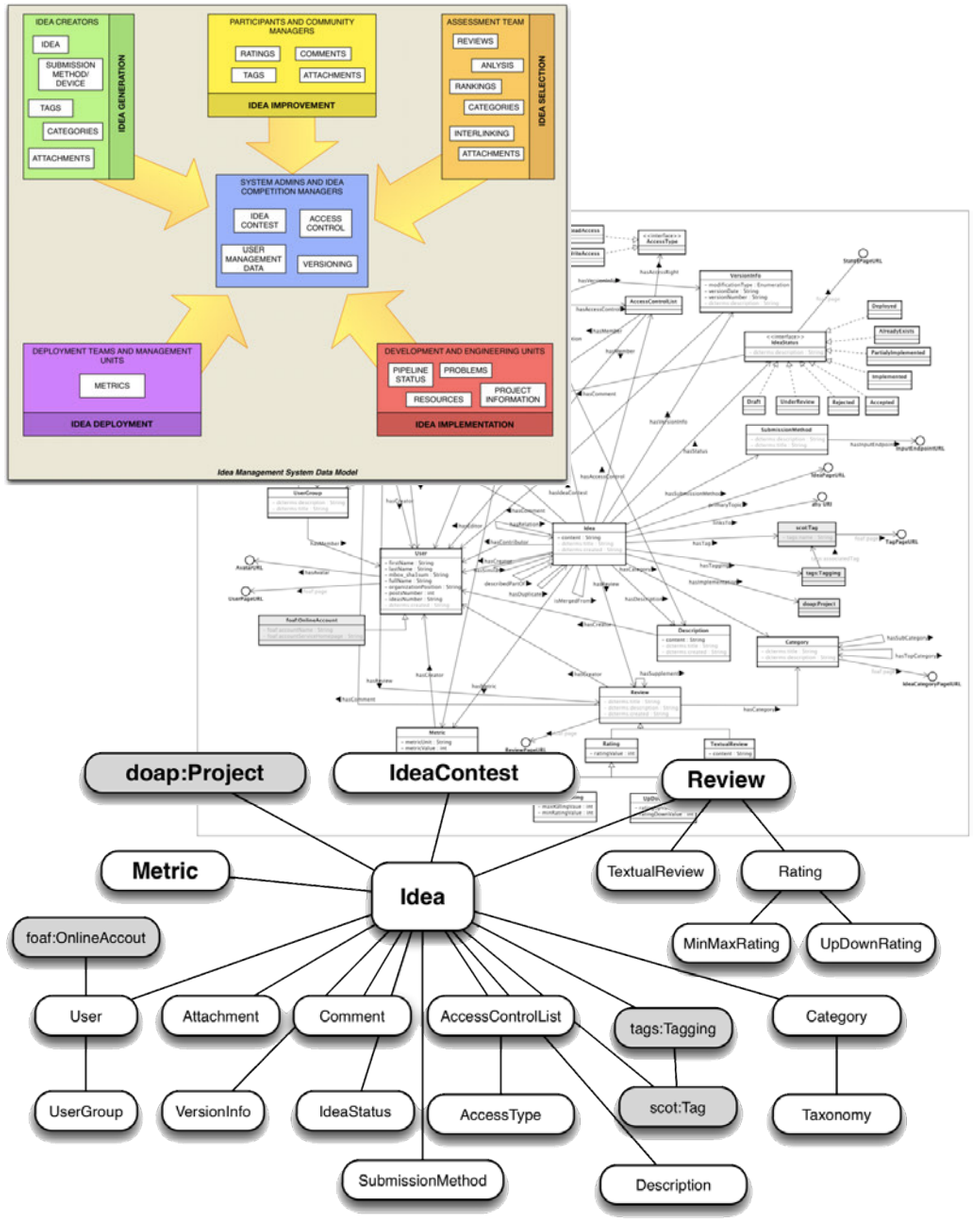
# Idea Life Cycle: Idea Metadata

## [Solution Part I: Generic Model for Idea Management Systems]



# Gi2MO Ontology: Life Cycle Formalization

## [Solution Part I: Generic Model for Idea Management Systems]




gsi UPM Grupo de Sistemas Inteligentes

GI2MO

### GI2MO Ontology Specification

V0.3 - 09 June 2010

This version: http://purl.org/gi2mo/0.3/ns (RDF/XML, HTML)
Latest version: http://purl.org/gi2mo/ns
Previous version: http://purl.org/gi2mo/0.2/ns
Editors: Adam Westerski
Authors: Adam Westerski
Contributors: See acknowledgements



#### Abstract

Generic Idea and Innovation Management Ontology (**GI2MO**) is a standardised data schema (also referred as "ontology" or "vocabulary") designed to annotate and describe resources gathered inside Idea Management facilities. The following document contains the description of ontology and instructions how to connect it with descriptions of other resources.

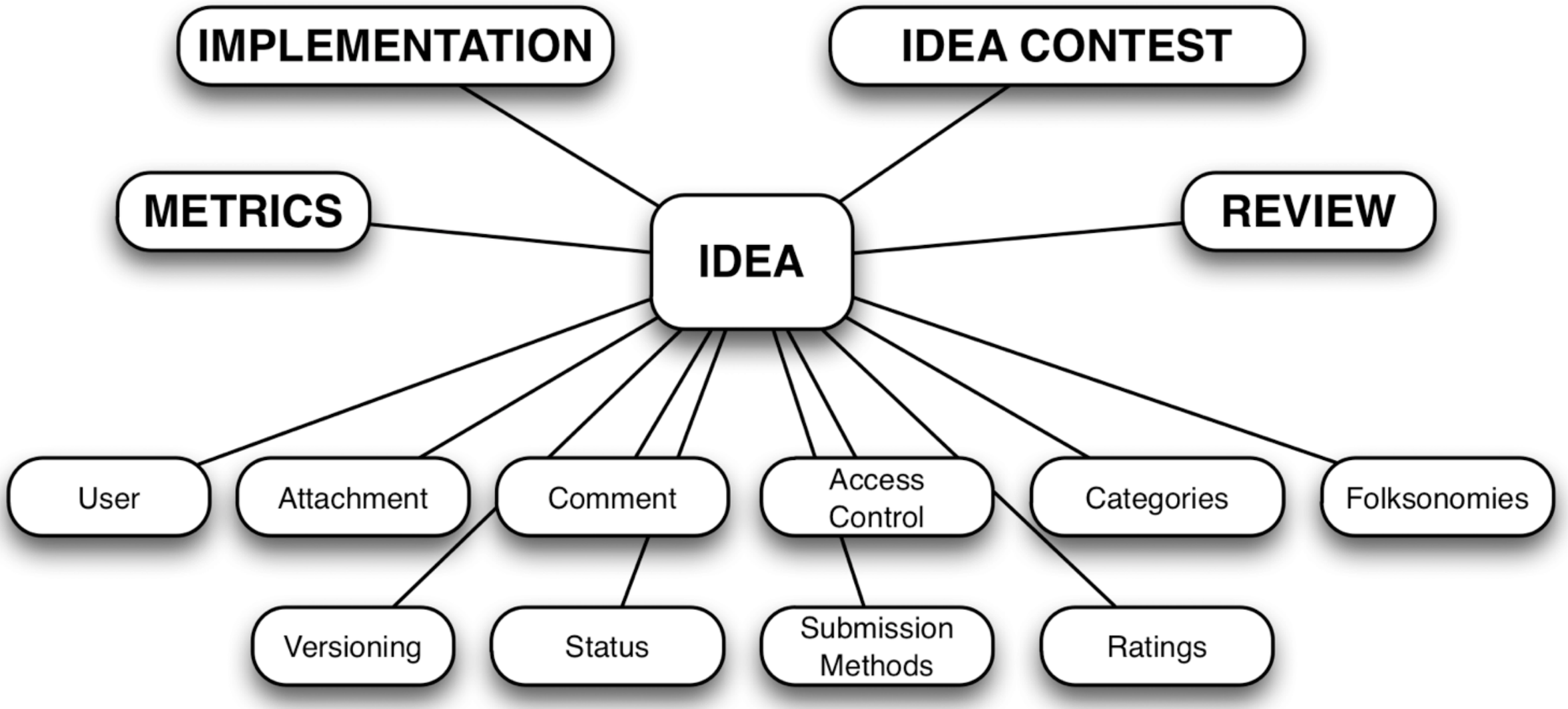
#### Table of Contents



#### 1 Introduction

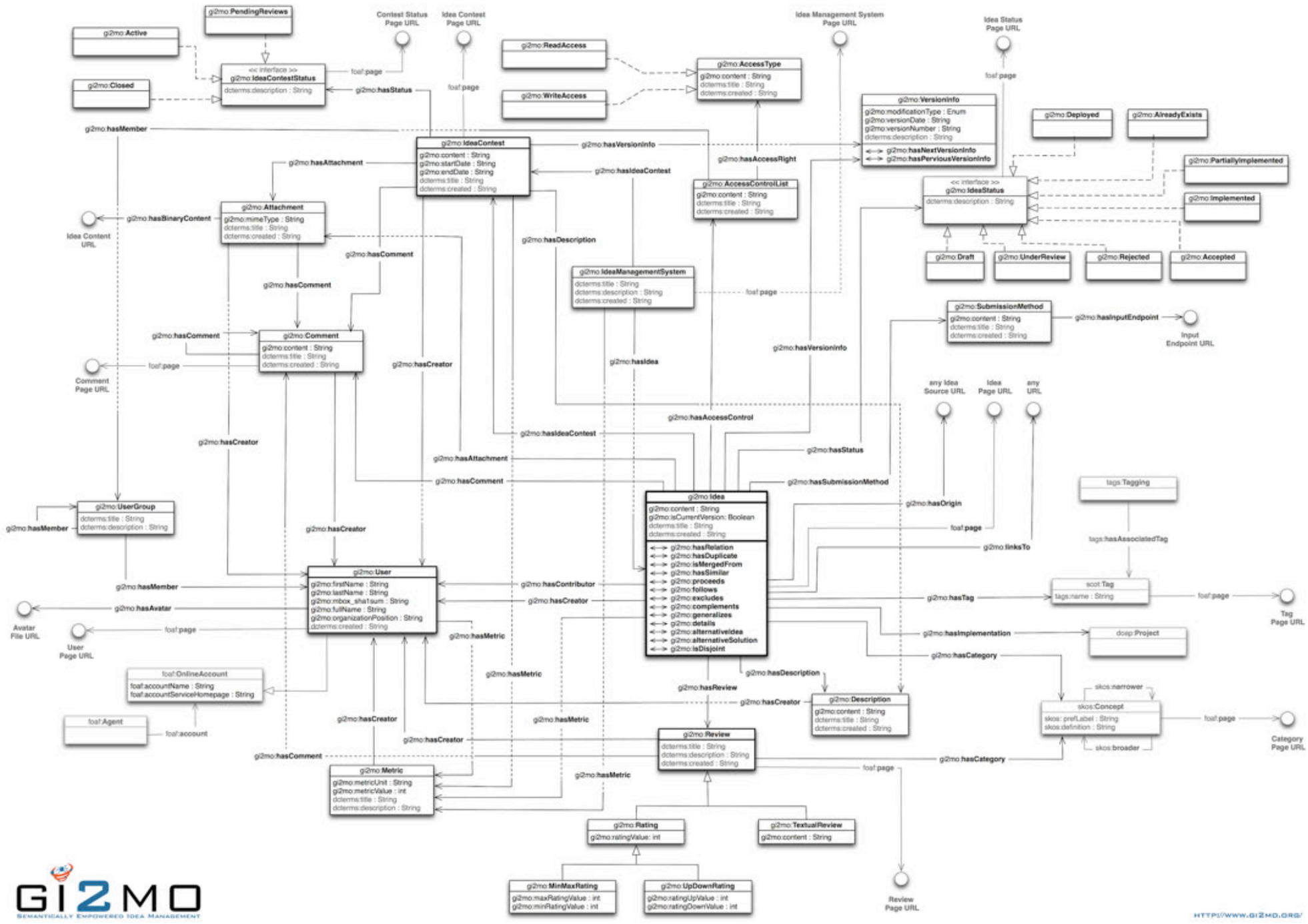
# Gi2MO Ontology: Overview

## [Solution Part I: Generic Model for Idea Management Systems]



# Gi2MO Ontology: Overview II

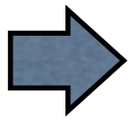
## [Solution Part I: Generic Model for Idea Management Systems]



# Gi2MO Ontology: Experiments

## [Solution Part I: Generic Model for Idea Management Systems]

HTML scraping ideas in the wild



Gi2MO Mappings

| System Name | ideas/ comments/ users | Area | Characteristics |
|---|---|---|---|
| Dell IdeaStorm | 15k / 90k / 2k | Computers, teleco hardware, related services | indefinite/ existing products/ focus sessions |
| myStarbucks Ideas | 8k / 80k / 3k | Coffee, related products and services | indefinite/ products & services |
| Cisco i-Prize | 1k / 4k / 1k | Networking and communications equipment | fixed time / abstract ideas on future products/ money incentives |
| Adobe Acrobat Ideas | 500 / 2k / 600 | Computer software | indefinite/ single product |

**Results:** the model covers on average 87% of metadata

# Progress beyond SoA

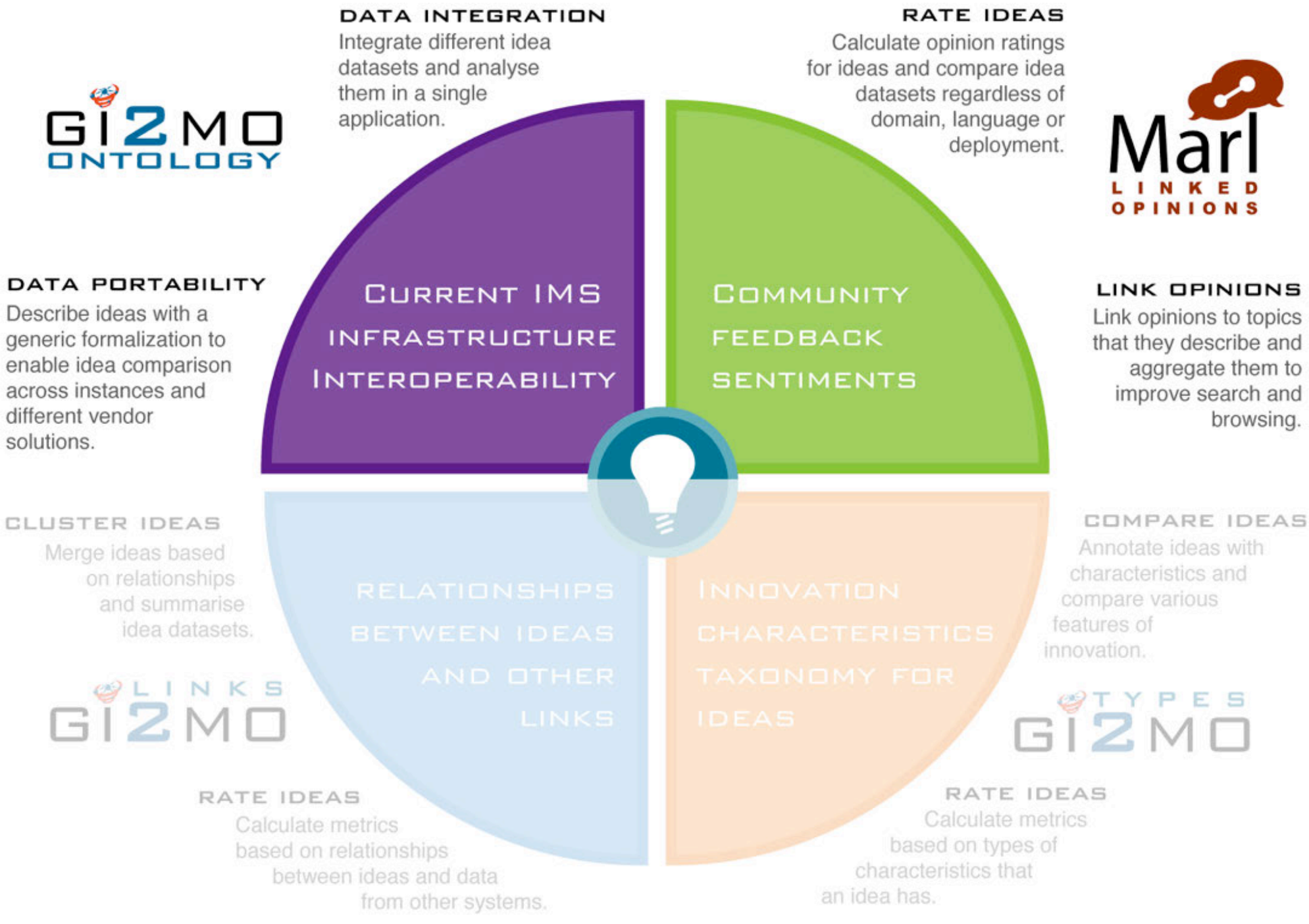
## [Solution Part I: Generic Model for Idea Management Systems]

- Riedl et al., 2009, **An idea ontology for innovation management**, International Journal on Semantic Web and Information Systems
- Bullinger, A. C., 2008, **Innovation and Ontologies**, Gabler

**Thesis approach:** data interoperability as a goal, wider scope, different modeling approach

# Community Opinions in Idea Management

## [Solution - Part II]



# Community Opinions in Idea Management

[Solution - Part II]

**Methodology:**

1. Analyse opinions posted in variety of industrial systems & **available indicators** in related research areas
2. Capture common elements in a form of **ontology**
3. **Evaluate** the use of ontology in a number of case studies
4. Evaluate the value of identified data properties **in Idea Management Systems**

# Measuring Opinions: Sentiment Analysis

## [Solution Part II: Community Opinions in Idea Management Systems]

Comment:
I like this is idea and I totally support it!

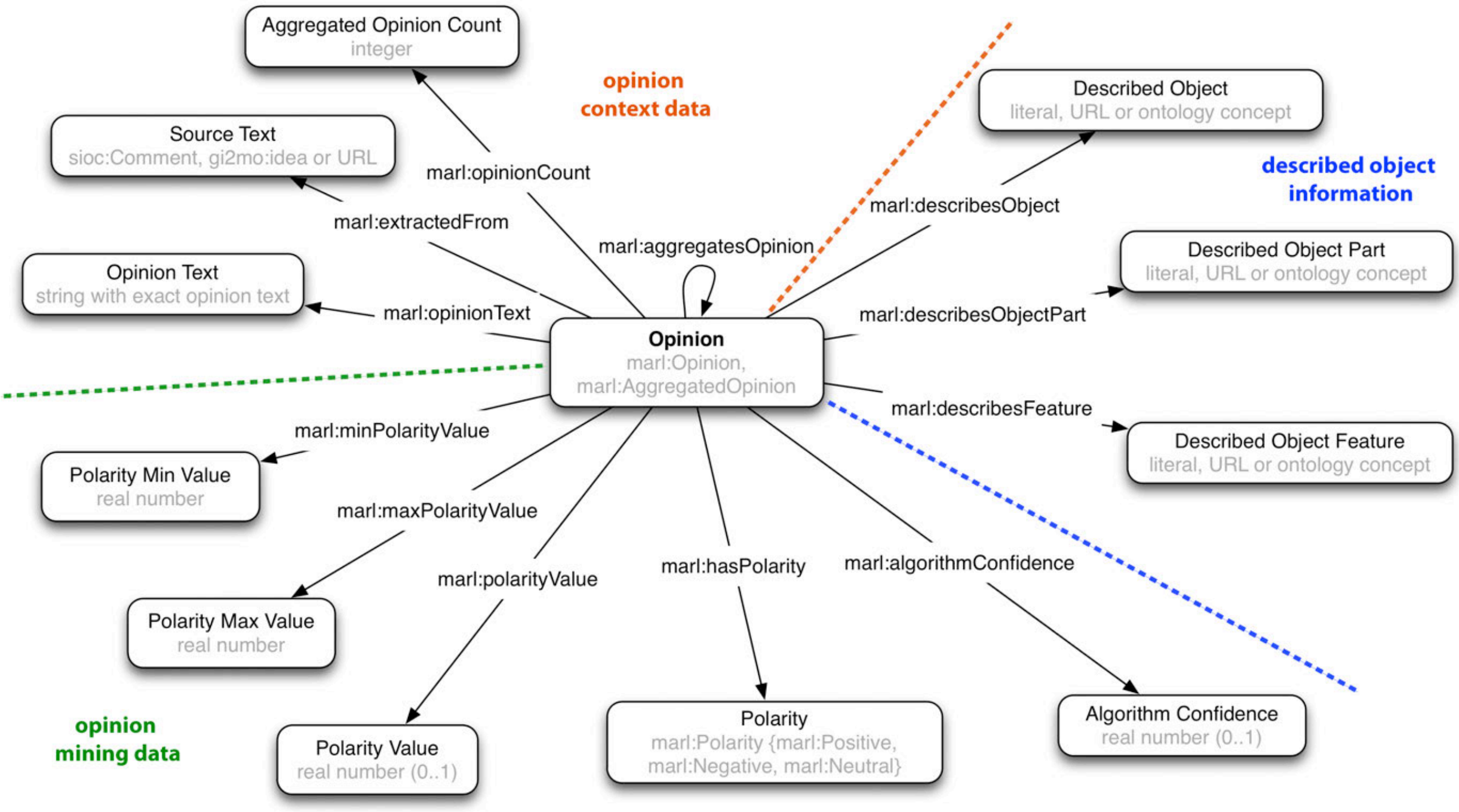
Polarity: positive
Polarity rating: 1.4

Comment:
I think this idea is terrible and putting a bigger screen will make the product worse

Object: Screen
Feature: Size
Polarity: negative
Polarity rating: -2.1

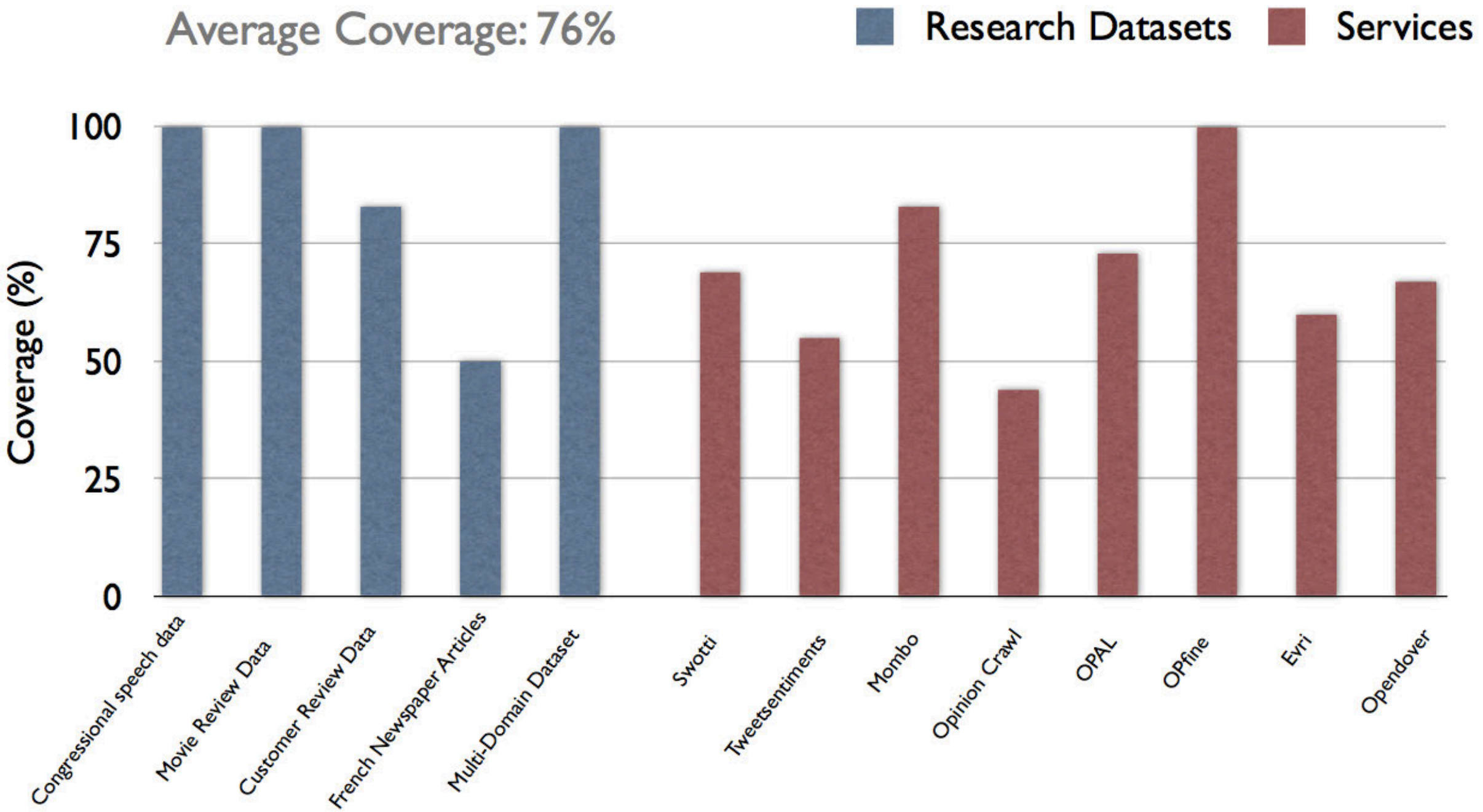
# MARL: Describing and Linking Data

## [Solution Part II: Community Opinions in Idea Management Systems]



# Evaluation part I: Ontology Coverage

## [Solution Part II: Community Opinions in Idea Management Systems]



# Evaluation part II: Opinion Analysis in IMS

[Solution Part II: Community Opinions in Idea Management Systems]

**Can structured OM data be useful in the context of IMS ?**

- **H1:** sentiments in idea comments are an indicator if idea is accepted or not (by IMS managers)
  - **EX1:** measure correlation between opinion rating and idea adoption (1 = implemented , 0 = not)
- **H2:** analyzing comments with opinion mining delivers a new assessment tool in comparison to current metrics
  - **EX2:** measure correlation between opinion rating and IMS metrics: comment count, solution count, max. /min. / avg. solution rating, idea age

# Evaluation part II: Results

## [Solution Part II: Community Opinions in Idea Management Systems]

| **Metric** | **Correlation with idea adoption** |
|---|---|
| Comment count | 0.03 |
| Solution count | 0.04 |
| Max. solution rating | 0.3 |
| Min. solution rating | 0.24 |
| Avg. solution rating | 0.37 |
| Idea age | 0.12 |
| **Opinion rating** | **0.04** |

do decision makers take all community comments into account ?

# Evaluation part II: Results (II)

## [Solution Part II: Community Opinions in Idea Management Systems]

| **Metric** | **Correlation with opinion rating** |
|---|---|
| Comment count | 0.28 |
| Solution count | -0.08 |
| Max. solution rating | 0.25 |
| Min. solution rating | 0.38 |
| Avg. solution rating | 0.41 |
| Idea age | 0.19 |

does opinion rating duplicate other metrics ?

# Progress beyond SoA

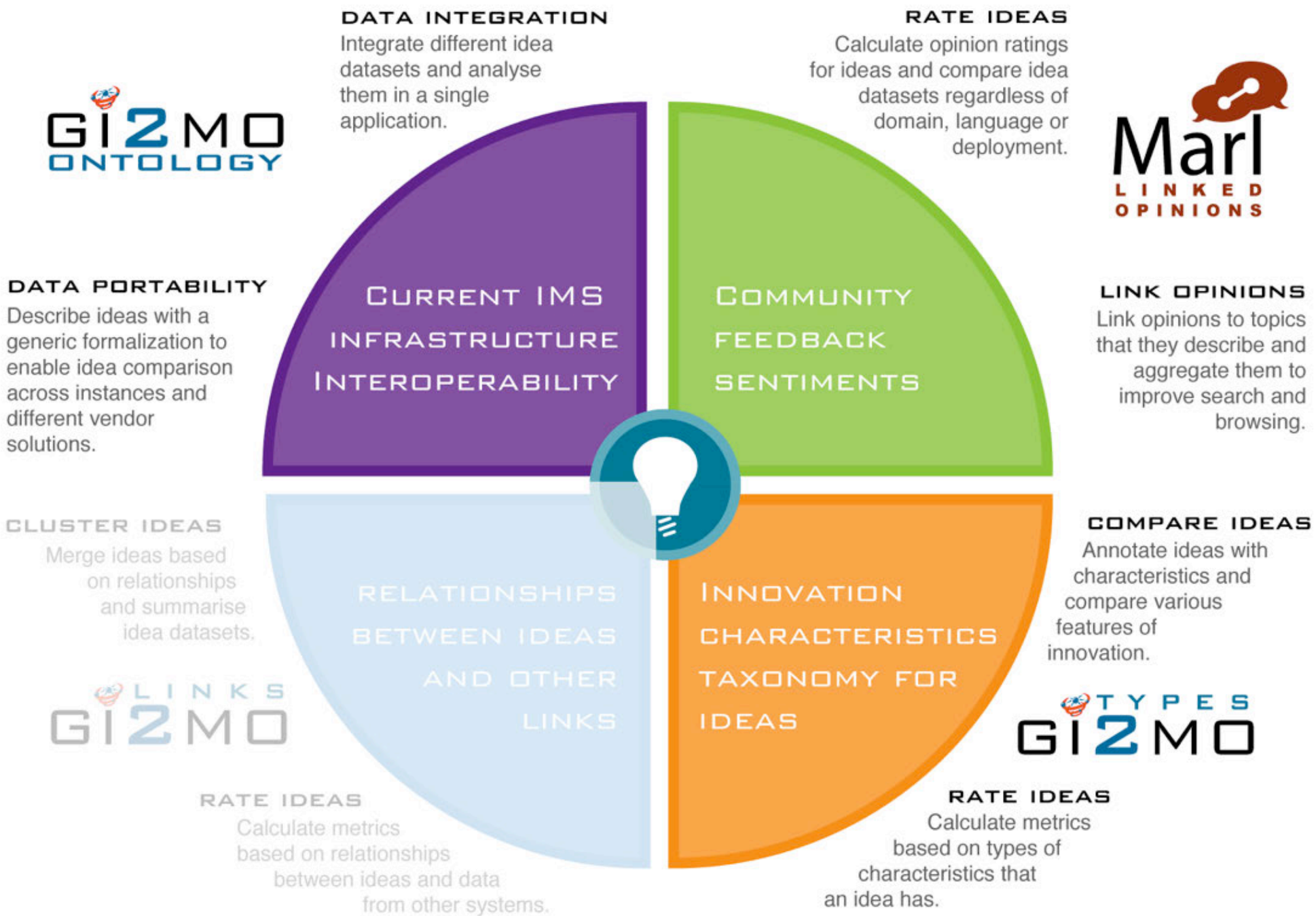
## [Solution Part II: Community Opinions in Idea Management Systems]

- opinion modeling:
  - Softic, S., and Hausenblas, M., 2008, **Towards opinion mining through tracing discussions on the web**, SDoW, 7th International Semantic Web Conference (ISWC)
  - industry approaches: hReview, Schema.org ...
- community activity measures:
  - targeted at particular areas: movie reviews, product reviews, business intelligence
  - IMS: Bothos et al, Hrastinski et al ...

**Thesis approach:** opinions are not the same as reviews; add modeling of polarity; use OM to generate new metric for IMS

# Idea Characteristics Model

## [Solution - Part III]



# Idea Characteristics Model

[Solution - Part III]

**Methodology:**

1. Analysis of the **theoretical models** from Innovation Management area
2. Confront **theory vs. practice** of IMSes
3. **Taxonomy** for domain independent idea characteristics in IMS
4. **Annotation** of ideas with the taxonomy
5. **Generate metrics** based on annotations to assess content of IMS and compare ideas

# Building the taxonomy (I)

## [Solution Part III: Idea Characteristics Model]

### Theoretical Innovation Models (examples)

- product vs. process vs. market vs.  input innovation
  Schumpeter 1934; Drejer 2004
- radical vs. incremental innovation
  Abernathy 1978; Porter 1986; Garcia 2002
- modular vs. architectural innovation
  Henderson and Clark 1990
- new markets vs. old markets
  Abernathy and Clark 1985

**Approach:** Innovation Management theory vs. Idea Management System practice (Gi2MO Ontology & Idea Life Cycle research)

# Building the taxonomy (II)

## [Solution Part III: Idea Characteristics Model]

**Idea**

**Trigger**

- Observation Type
  - Faulty Experience
    - Potential Cause
  - Lack of Feature
  - Potential Opportunity
- Creativity Origin
  - Event
    - Idea Contest
    - Object Release
      - New Object
      - Object Update
      - Promotional Event
    - Object Purchase
    - Object Comparison
    - Object Observation
    - Other Event
  - Object Interaction
- Associated Object
  - Object Of Innovation
  - Other Object
    - Object Relation
      - Competitive
      - Complementary
      - Suplementary
      - Unrelated
    - Offering Placement
      - Own Offering
      - Other Party Ofering

**Innovation**

- Dependence
  - Proceeds
  - Follows
  - Encapsulates
  - Duplicates
  - Excludes
  - Is Part Of
- Target Audience
  - New Audience
  - Existing Audience
- Originality
  - New
  - Incremental
    - Additive
    - Subtractive
    - Replacement
  - None
  - Relative To
    - Current State
      - Organisation
        - Structure
          - Group
          - Department
          - Branch
        - Product Line
        - Product Type
      - Market
        - Local
        - Global
    - Innovation Proposals

**Object**

- History Relationship
  - Evolutionary
  - Regressive
- Type
  - Service
  - Process
  - Product
    - Product Line
    - Product Type
    - Specific Product
- Offering Placement
  - New
  - Existing
    - Additive
    - Subtractive
    - Replacement
- Structure
  - Complete
  - Element
  - Characteristic
- Relationships
  - Part Of
  - Complementary With
  - Composed Of

**Proposal Type**

- Solution
- Request
- Issue Report
- Suggestion

# Idea Annotation

## [Solution Part III: Idea Characteristics Model]

**Idea Title:** More buttons on the tablet.

**Idea Summary:**
I would like to propose adding more physical buttons on the 9 inch tablet that is currently available in your offer. When using the tablet I feel that the single 'home' button is not enough for many activities that the tablet is advertised for making the experience bad.

For example, for reading ebooks, it would be very useful to have "back" and "forward" buttons for scrolling pages of the book. I own a e-paper reader and I think those buttons could be also used for different activities (for example web browsing or games).

**Trigger**

- Observation Type: Faulty Experience
- Creativity Origin: Object Interaction
- Associated Object | Other Object | Object Relation: Competitive
- Associated Object | Other Object | Offering Placement: Other Party

**Innovation**

- Target Audience: Existing Audience
- Originality: Incremental | Additive
- Relative to: Current State | Organisation: Product Line

**Proposal Type**

- Request

**Object**

- History Relationship: Evolutionary
- Structure: Element
- Type | Product: Specific Product
- Offering Placement | Existing: Additive

# Generating the metrics

## [Solution Part III: Idea Characteristics Model]

**Innovation**

- Dependence
  - Proceeds
  - Follows
  - Encapsulates
  - Duplicates
  - Excludes
  - Is Part Of
- Target Audience
  - New Audience
  - Existing Audience
- Originality
  - New
  - Incremental
    - Additive
    - Subtractive
    - Replacement
  - None
  - Relative To
    - Current State
      - Organisation
        - Structure
          - Group
          - Department
          - Branch
        - Product Line
        - Product Type
      - Market
        - Local
        - Global
    - Innovation Proposals

→ Idea Dependability

→ Idea Adaptiveness = { 1 = Existing audience; 0 = New audience }

→ Idea Originality

→ Idea Originality Scope

# Experiments & Results

## [Solution Part III: Idea Characteristics Model]

- Manual and Automatic Annotation Experiments:
  - manual:
    - 10 ideas x 10 annotators - full agreement in 34% cases
    - 2x (200 ideas x 1 annotator) - 70% annotations the same (alpha = 0.613 = substantial agreement)
  - automatic: 400 ideas - single annotator vs. machine learning algorithm - 0.46 f-measure, various experiments to improve, final result: trigger and object only applicable with f-measure above 0.6

# Experiments & Results

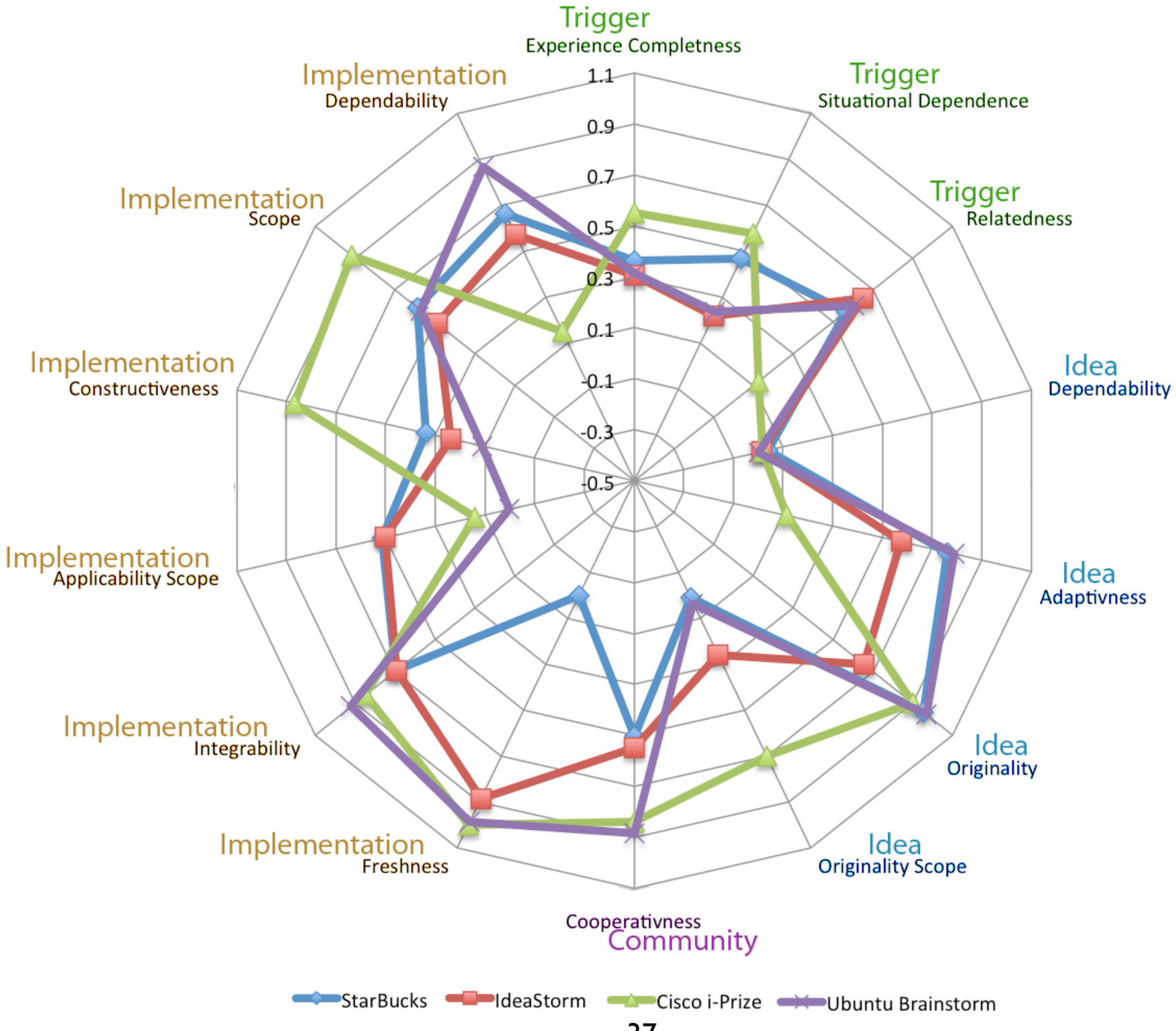
## [Solution Part III: Idea Characteristics Model]

- Usage for Idea Comparison and Idea Assessment
  - (200 ideas x 1 annotator) x 4 datasets

- Experiments:
  - correlation between Gi2MO Types metrics and idea adoption - same as for legacy metrics
  - differences in metrics vs. expected characteristics as described in case studies - new approach

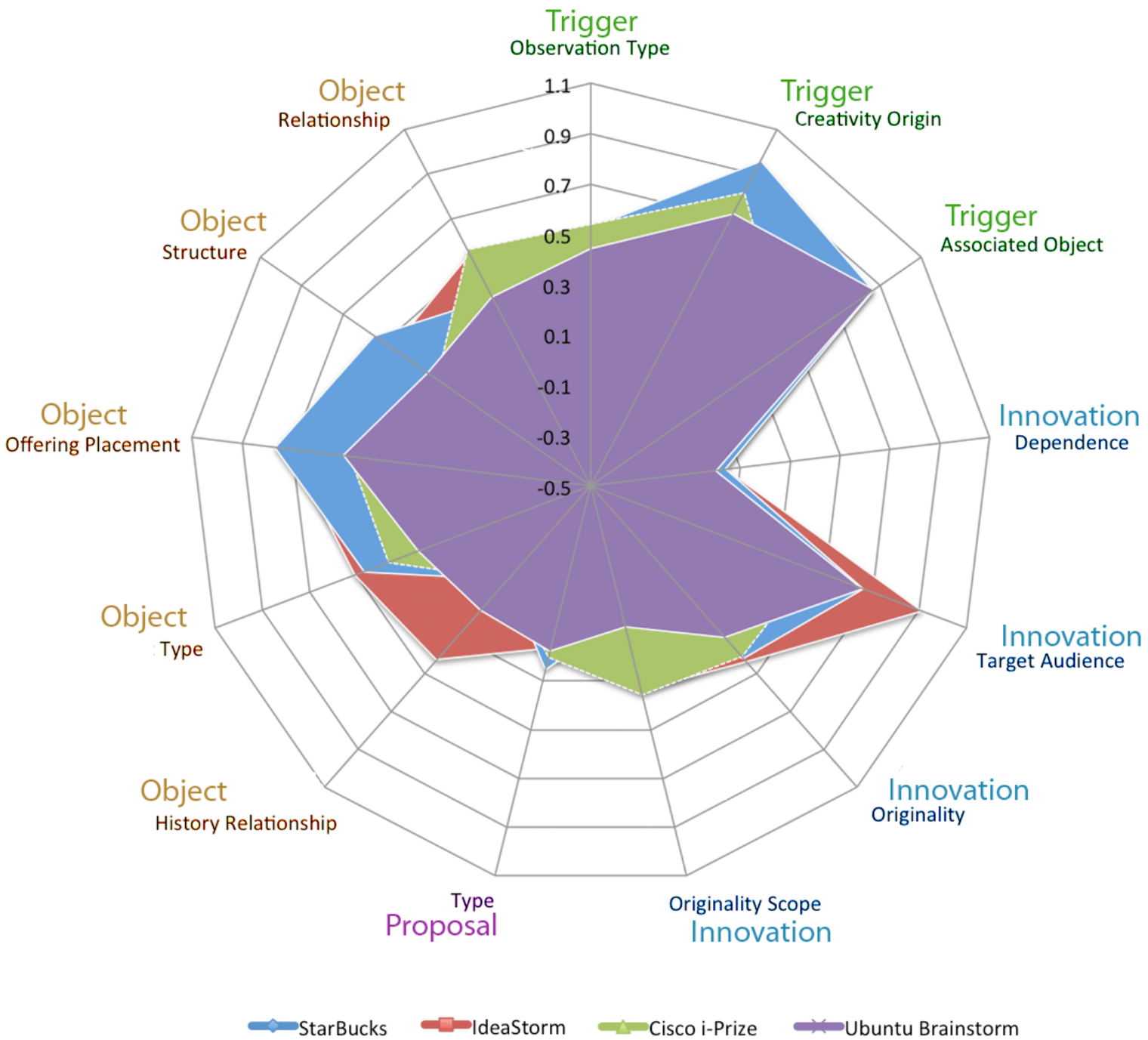
# Experiments: Dataset comparison

## [Solution Part III: Idea Characteristics Model]



# Experiments: Idea Similarity

## [Solution Part III: Idea Characteristics Model]




Use of information entropy as term diversity measure

Masisi 2008; Ghosh 2011; Huang 2008

number of term combinations in a taxonomy branch

term combination

$$E(tb_x) = -\sum_{i=1}^{n} p(i) \log_2 p(i)$$

frequency of annotations

# Progress beyond SoA

## [Solution Part III: Idea Characteristics Model]

- Schumpeter, J., 1934, **The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle**, Harvard University Press

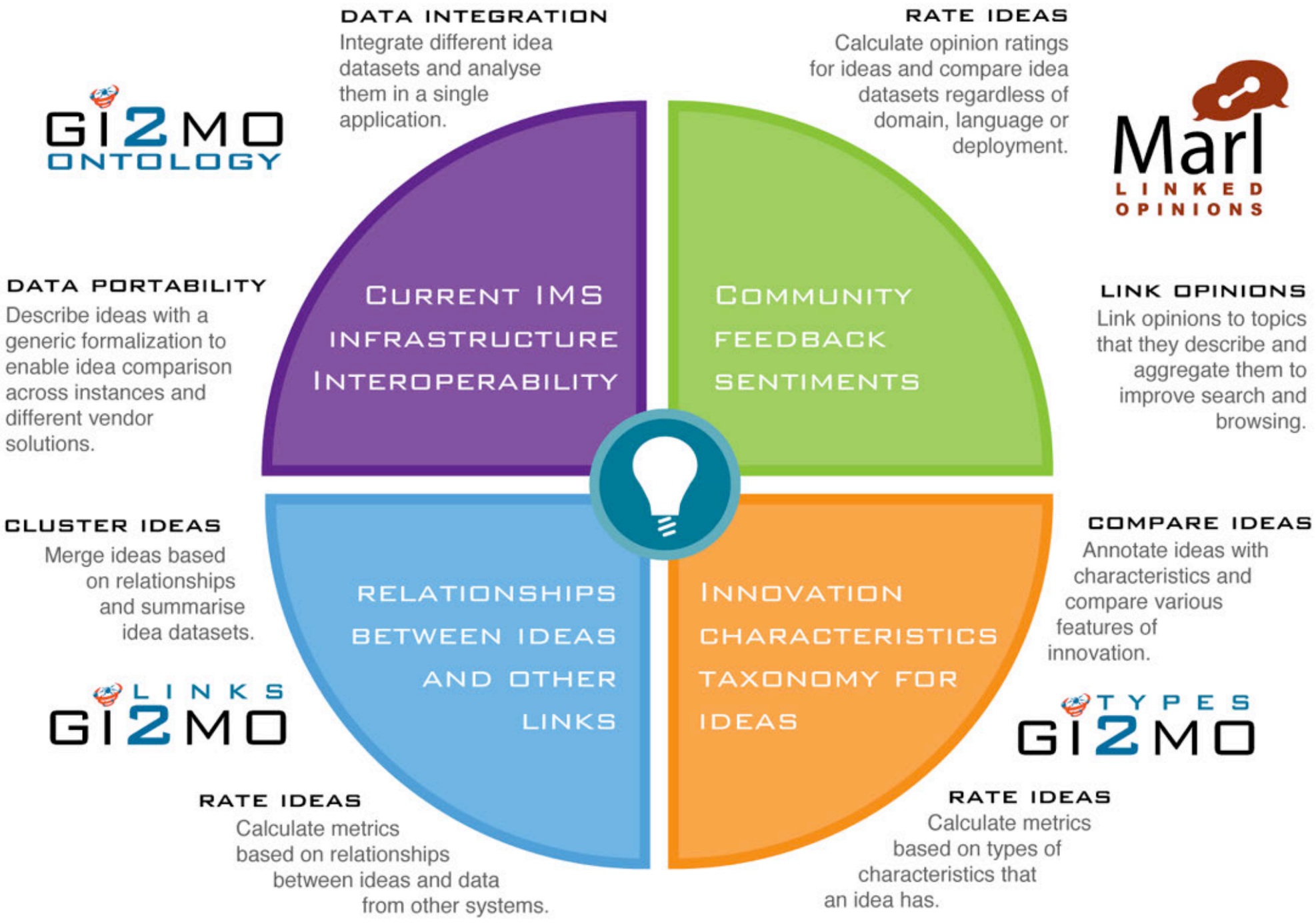
- Henderson, R. M., and Clark, K. B., 1990, **Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms**. Administrative Science Quarterly

- Abernathy, and W. J., Clark, K. B., 1985, **Innovation: Mapping the winds of creative destruction**, Research Policy

**Thesis approach:** apply the well known models for Idea Management Systems to allow cross domain system comparison

# Idea Relationship Model

## [Solution - Part IV]



# Idea Relationship Model

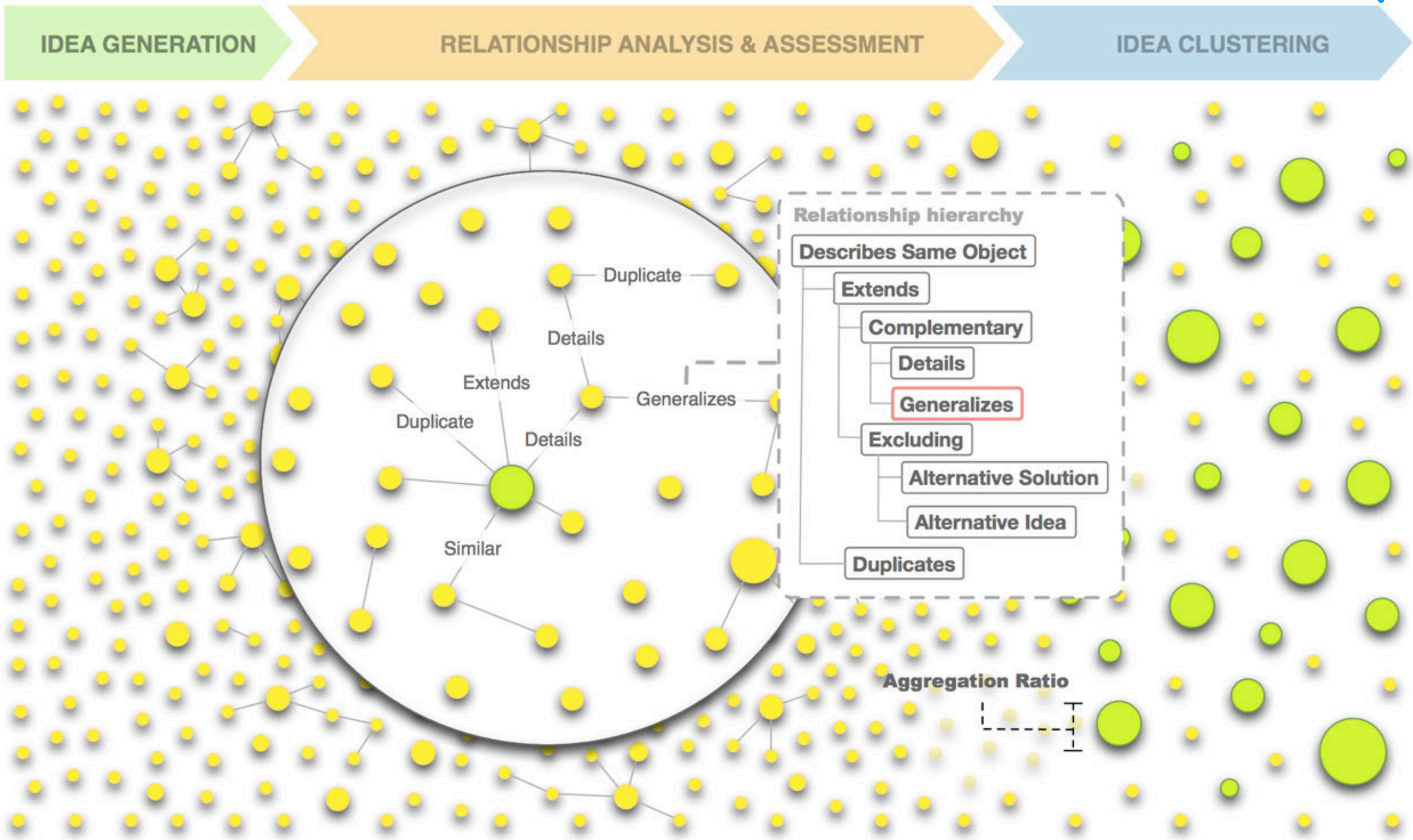
[Solution - Part IV]

**Methodology:**

1. Analyse state of the art on **relationship modeling** in various domains (Semantic Web domain ontologies, e-Learning metadata etc.)
2. Propose a **relationship hierarchy** for Idea Management Systems
3. Use of hierarchy to **annotate ideas** in available datasets from case studies
4. **Evaluation**: amount of relationships detected, clustering capabilities etc.

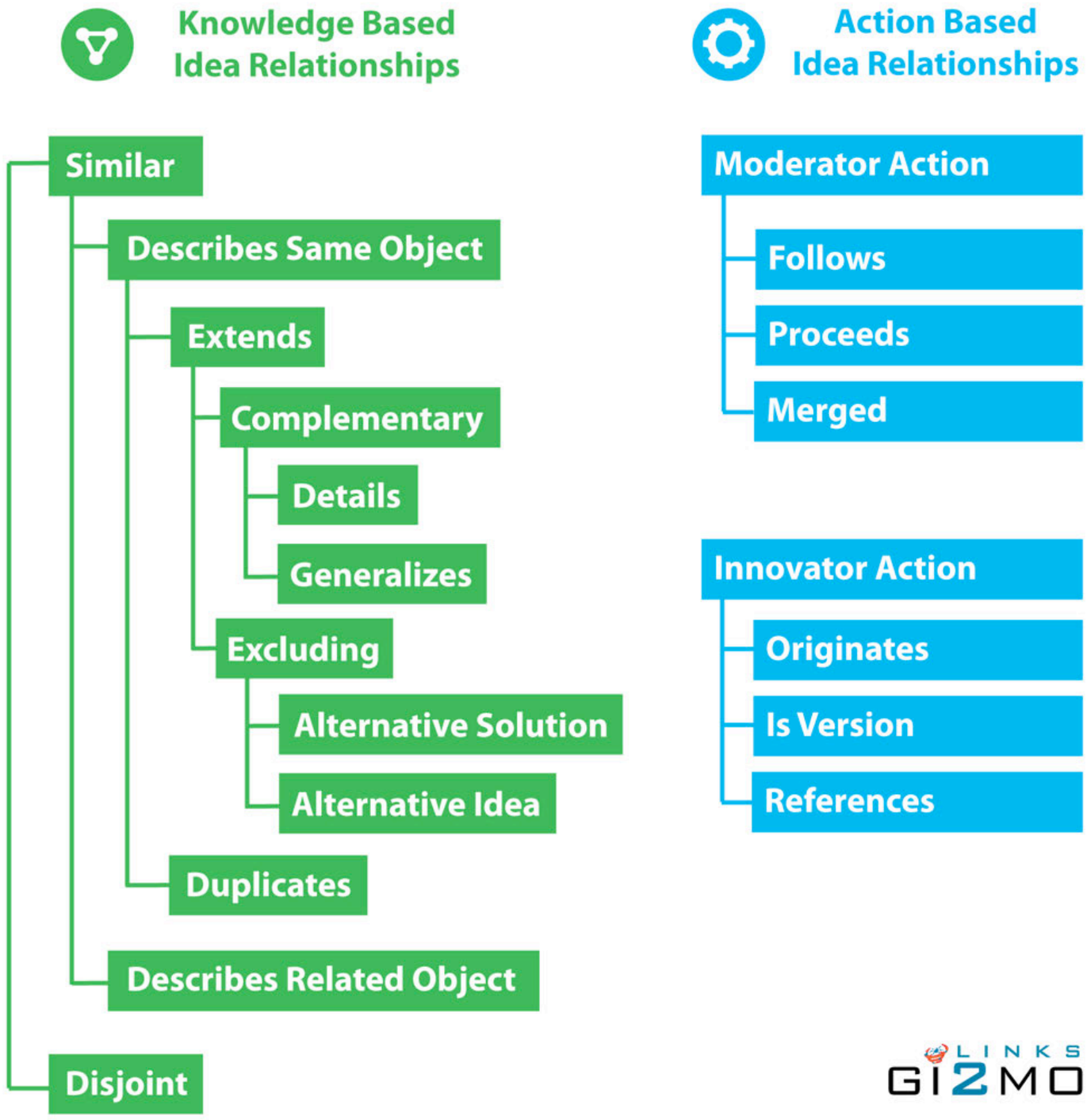
# Idea Relationships: IMS Scope

## [Solution Part IV: Idea Relationship Model]



# Idea Relationships: Hierarchy Proposal

## [Solution Part IV: Idea Relationship Model]



# Idea Relationships: Evaluation Dataset

## [Solution Part IV: Idea Relationship Model]

**Data**

- dataset size: 21k Ubuntu Brainstorm ideas
- manual annotation: 200 ideas - single annotator

**Annotation experiment**

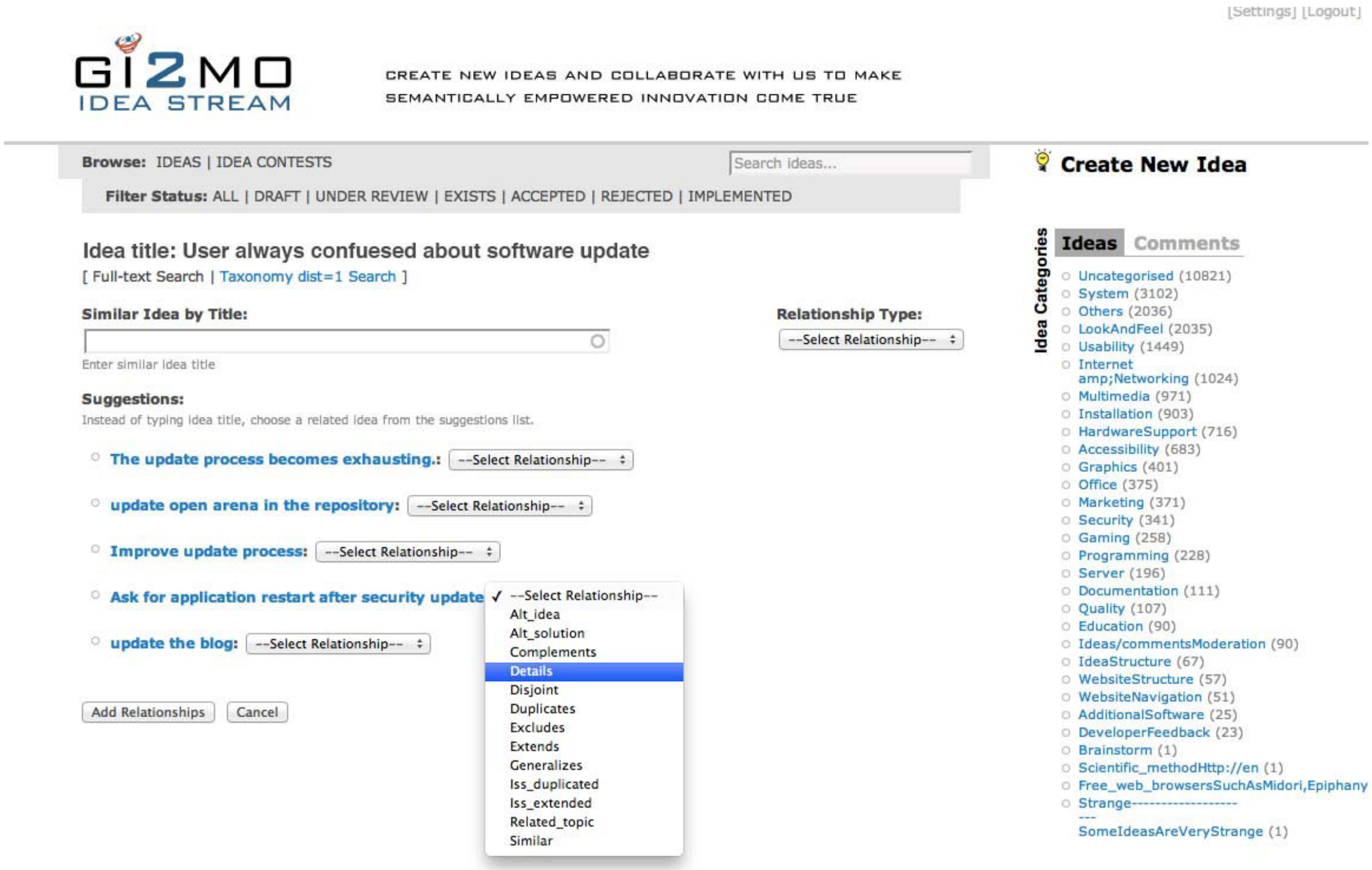
- each idea was given 5 suggestions of similar ideas
- suggestions based on keyword similarity between idea texts
- annotator to choose the type or relationships (or its lack)

**Output**

- 200x5 relationship annotations
- “duplicate” annotations from the original Ubuntu dataset

# Idea Relationships: Dataset Annotation

## [Solution Part IV: Idea Relationship Model]



# Idea Relationships: Evaluation

## [Solution Part IV: Idea Relationship Model]

**3 hypotheses & 3 experiments:**

- **H1:** semantics of relationships are more complex than "duplicate"
  - relationship amount comparison - 76.7% increase
- **H2:** wide range of relationships can be used to summarize datasets more than the current techniques
  - idea aggregation - 95% increase
- **H3:** apart of idea topic there are characteristics that impact how annotators perceive idea similarity
  - relationships vs. idea characteristics - 8 out of 14 metrics had small correlation with similar/dissimilar

# Progress beyond SoA

## [Solution Part III: Idea Characteristics Model]

- Kornish, L. J., and Ulrich, K. T., 2011, **Opportunity spaces in innovation: Empirical analysis of large samples of ideas**, Management Science

- industry: BrightIdea, IdeaCentral, SpigitEngage etc.

**Thesis approach:** define a hierarchy of relationships, extend the current state where only "duplicate" is used

# Contributions Overview

[Conclusions and future work]

- **Generic Model for Idea Management Systems**
  - **Idea Life Cycle**: proof for IMS data being highly interconnected and mutually dependent
  - **Gi2MO Ontology**: a single formalization based on Semantic Web methodologies that covers majority of IMS data and contributes to interoperability
- **Community Opinions in Idea Management Systems**
  - **MARL**: a model for structuring opinion data in IMS and Web
  - **opinion rating** - a new tool for analysis of community activity that brings new unique information into the idea assessment process

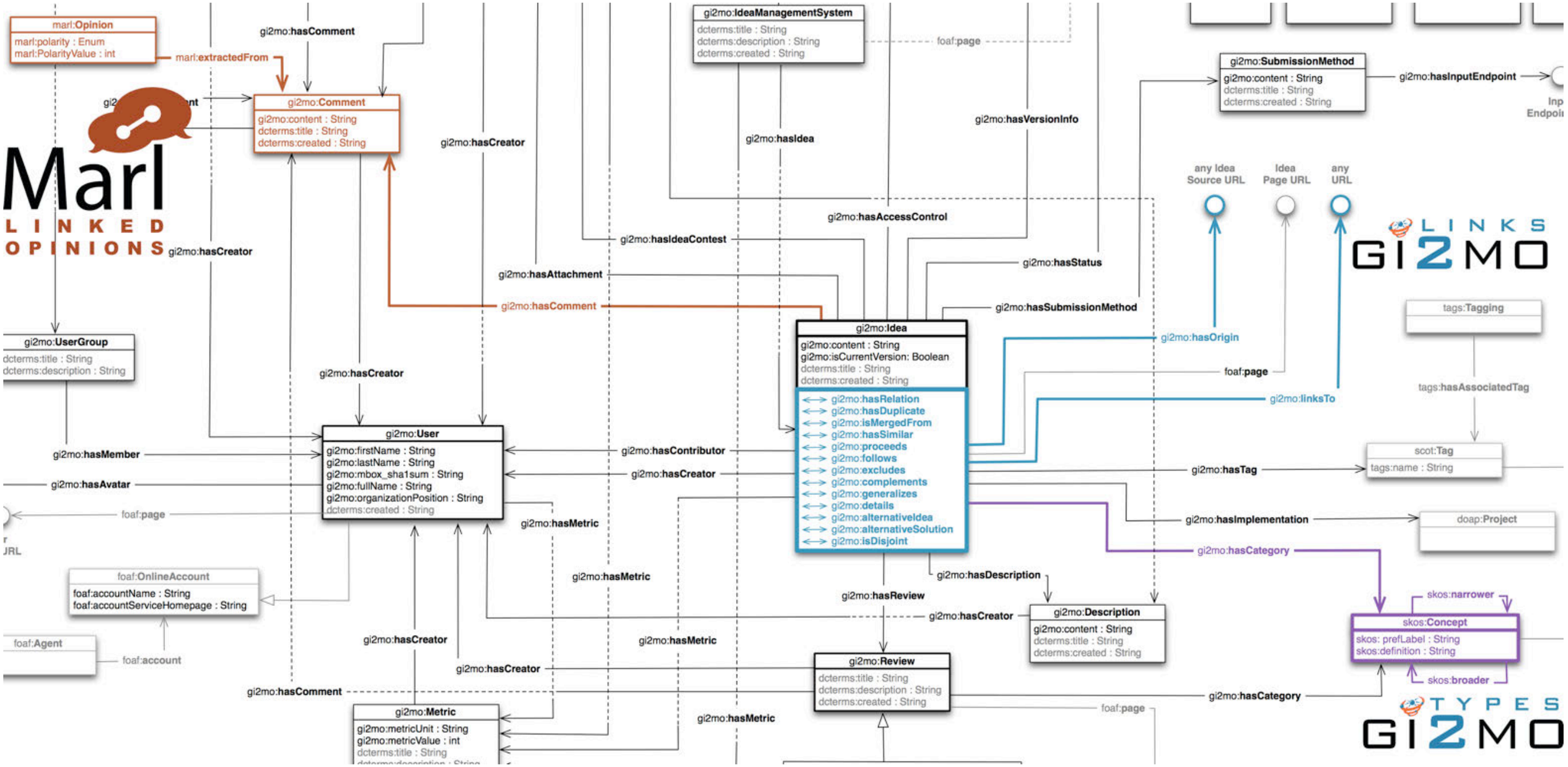
# Contributions Overview

[Conclusions and future work]

- **Idea Characteristics Model**
  - **Gi2MO Types**: a taxonomy for domain independent characteristics derived from innovation management theories
  - **Types metrics**: a new solution for idea assessment based on analysis of characteristics of IMS instances or data subsets
- **Idea Relationships Model**
  - **Gi2MO Links**: a hierarchy of relationships, proven as hugely omitted topic by contemporary systems
  - **Gi2MO Links vs. Gi2MO Types:** non-domain characteristics can be used to determine general similarity or dissimilarity

# Contributions Overview: Data Model Synergy

## [Conclusions and future work]



# Future work

[Conclusions and future work]

- **Established lines of research to continue in the future**
  - automatic idea annotation experiments
  - use of structured opinion data in other domains
  - ...
- **New possibilities opened based on work done**
  - use of relationships for idea assessment
  - interoperability between IMS and other systems
  - ...

# Results: Publications

## [Publications and results]

- **Generic Idea Management System Model** [2010-2012]
- [Journal] The Road from Community Ideas to Organisational Innovation: A Life Cycle Survey of Idea Management Systems Journal Web-Based Communities 2011
- [Conference] A Model for Integration and Interlinking of Idea Management Systems, MTSR 2010

- **Community Opinions in Idea Management Systems** [2011]
- [Workshop] Linked Opinions: Describing Sentiments on the Structured Web of Data SDoW2011 @ ISWC 2011
- [Workshop] Mining Sentiments in Idea Management Systems as a Tool for Rating Ideas LSDeliberation @ COOP 2012

- **Idea Characteristics Model** [2011-2012]
- [Journal] Taxonomy usability study Decision Support Systems Journal 2013 (JCR Q1 Impact Factor: 1667 / CORE- A*)

- **Idea Relationships Model** [2010 & 2012]
- [Workshop] Exploiting Structured Linked Data in Enterprise Knowledge Management Systems VORTE2011 @ EDOC 2011
- [Poster] Gi2MO: Interoperability, Linking and Filtering in Idea Management Systems ESWC 2011
- [Conference, Poster, Demo] Idea Relationship Analysis in Open Innovation Crowdsourcing Systems CollaborateCom 2012

- **Co-authors publications on Gi2MO extensions** [2011-2012]
- [Workshop] Building Consensus via a Semantic Web Collaborative Space WWW 2012
- [Conference] Application of Semantic Search in Idea Management Systems ICITST 2012

# Results: Software

## [Publications and results]

Gi2MO — Semantically Empowered Idea Management

Gi2MO Apps

- Gi2MO Ontology | IMS Interlinking, Data Portability →
  - **RDFme** (PHP/Drupal) [2010-]
  - **Idea Browser** (Flash) [2010/2011]
  - **Gi2MO RDF2HTML** (Javascript, PHP) [2010]
  - **Gi2MO IdeaStream** (PHP/Drupal) [2011-]
  - **Gi2MO Stats** (Objective-C/iPhone) [2012]
  - **Gi2MO Reader** (HTML5/Windows 8) [2013]
- Enterprise Linked Data Model | Enterprise Data Interlinking →
  - **Idea Analyst** (Flash) [2010]
  - **Google Wave Plugin** (Python/Wave) [2010]
  - **IdeaStream Recommender** (PHP/Drupal) [2011]
- Marl Opinion Ontology | Opinion Analysis →
  - **OPAL** (PHP/Drupal) [2010]
- Gi2MO Types, Gi2MO Links | Idea Assessment →
  - **IdeaStream Analytics** (PHP/Drupal) [2011-]
  - **IdeaStream Similarity** (PHP/Drupal) [2012-]
  - **Relationship Vis** (JavaScript/PHP) [2013]

# Results: Referencing & Use

[Publications and results]

## References to research contributions of the thesis

- Mondragon Univeristy | Use of Idea Management in technological clusters of enterprises
  ↳ Larrinaga et al., 2011

- IMC Technologies | Use of Idea Management for question answering in large scale deliberation spaces
  ↳ Anadiotis et al. 2012

- INRIA | Research on Green Services and citizen participation
  ↳ Leitzellman et al., 2011; Negri, 2012

- DERI | Research on structuring information from brainstorming processess
  ↳ Lorenzo et al., 2011

## Use of software from Gi2MO Project

- large enterprises | e.g. Saab group for gathering ideas from employees
- small-medium companies | e.g. Ericpol consulting for collecting feedback about ongoing projects
- research laboratories | e.g. INRIA, support for experiments in e-government domain
- university associations | e.g. ETSIT UPM Fundatel design competitions for students

# Results: Referencing & Use II

## [Publications and results]

## Contribution to Funded Projects

- RESULTA | Research on improving communication in the enterprise and in consortiums of consulting companies
  - ↳ Gi2MO Ontology, MARL Linked Opinions

- THOFU | Research on hotel services for the future and use of new technologies to improve hotel business and tourism
  - ↳ Gi2MO Types, Gi2MO Links

- EuroSentiment (ongoing) | Research on accessibility, sharing and interoperability of multilingual resources
  - ↳ MARL Linked Opinions

## Contributions to thesis research via collaborations

- ATOS Origin, Spain | In-the-house Idea Management solution & case study of Idea Management use in the enterprise
- Athena Research Institute, Greece | Solution for automatic annotation based on machine learning

# Semantic Technologies in Idea Management Systems:

A Model for Interoperability, Linking and Filtering

## Thanks for attention!

## Questions ?

PhD candidate: Adam Westerski
westerski.adam@gmail.com

Supervisor: Carlos A. Iglesias
cif@gsi.dit.upm.es