

---

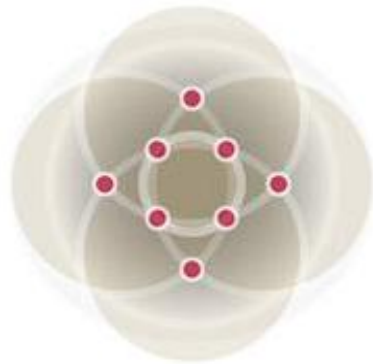
# FROM PATENTS TO TECHNOLOGIES – A NEW LEVEL OF ANALYSIS?

Patricia Helmich, Dr. Rainer Frietsch, Dr. Peter Neuhäusler

Fraunhofer ISI, Karlsruhe

13. September 2016, Valencia

---



6th Annual

# Global TechMining Conference

---

2016.09.13 • València, Spain

---

Source: <http://www.gtmconference.org/images/gtmheader.jpg>

---

# Structure of the Presentation

---

---

1. Introduction
2. The Challenge
3. Our approach
4. Future Work

# Introduction

---

---



Source : <https://albert-schweitzer-stiftung.de/wp-content/uploads/patent-stempel-adventtr-i.jpg>, <http://genericpharmafinance.com/wp-content/uploads/sites/10/2014/03/epo-european-patent-organisation-logo.jpg>, <http://www.pi-news.net/wp/uploads/2009/07/es-ward-licht.gif>

# Introduction

---

---



**In innovation research, patents are used as an indicator for the output of application-oriented research and development of a country that allows assessments of its current and future technological competitiveness.**

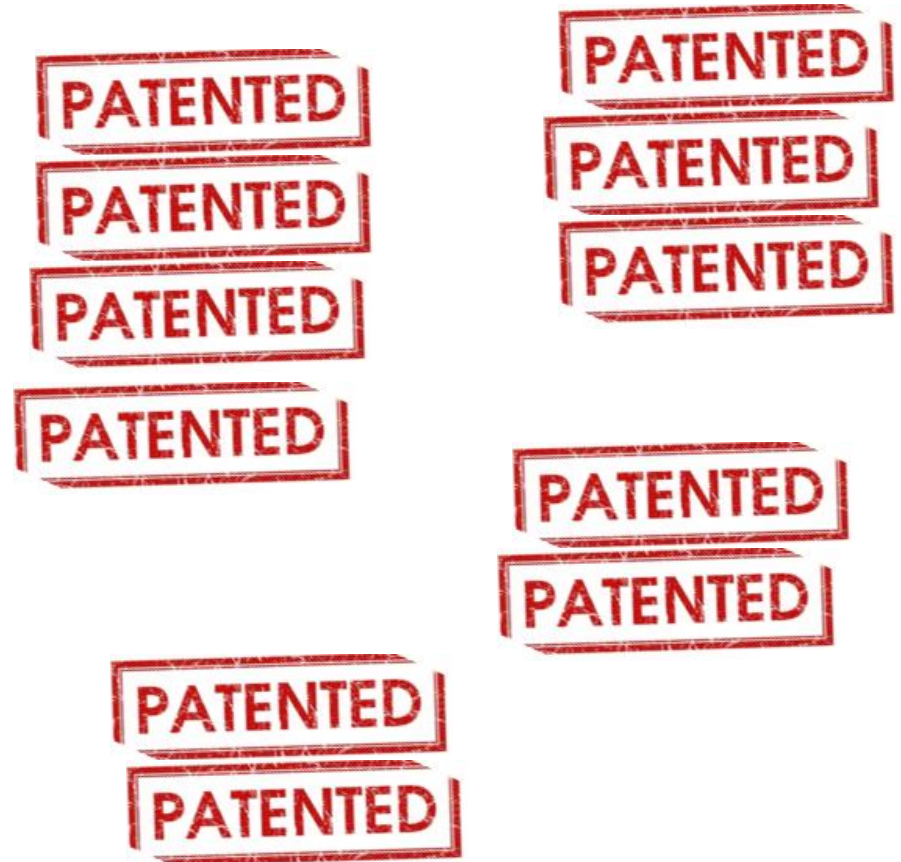
Source : [https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp\\_141325435.jpg](https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp_141325435.jpg)

# Introduction

---

---

- Patents can be analyzed at different levels, e.g.
  - Countries
  - Technology fields
  - Inventors/applicants
  - Single filings



Source : [https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp\\_141325435.jpg](https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp_141325435.jpg)

# Introduction

---

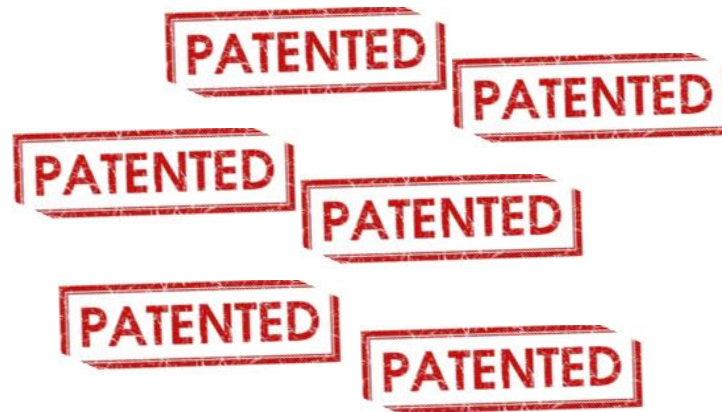
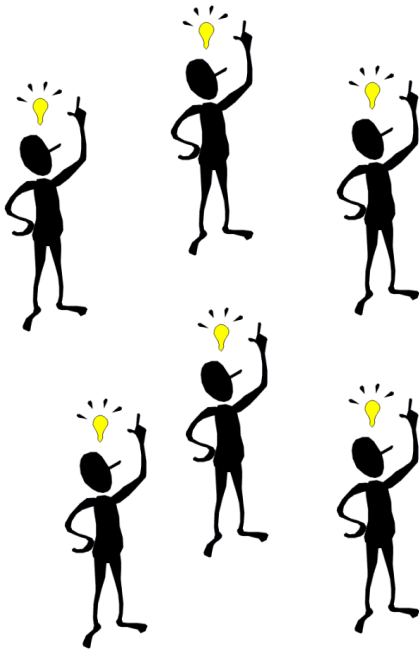
---



**PATENTED**

Source : [https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp\\_141325435.jpg](https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp_141325435.jpg), <http://www.pi-news.net/wp/uploads/2009/07/es-ward-licht.gif>

## Technology (e.g. MP3)



Source : [https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp\\_141325435.jpg](https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp_141325435.jpg), <http://www.pi-news.net/wp/uploads/2009/07/es-ward-licht.gif>



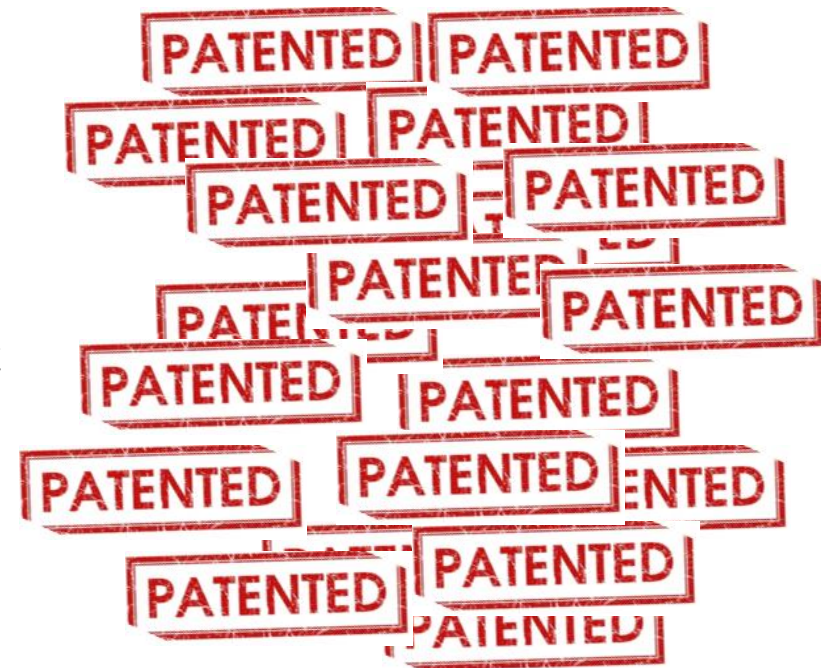
# Introduction

---

---

- Technologies generally not protected by only a single patent, but by a whole bunch of technologically related patents (including core patents)
- New level of analysis:
  - Identification of technologies developed by a company
  - Comparison of technologies developed by different companies/countries
  - Novel statistical evidence on the development stage of a technology

## Technology



Source : [https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp\\_141325435.jpg](https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp_141325435.jpg)

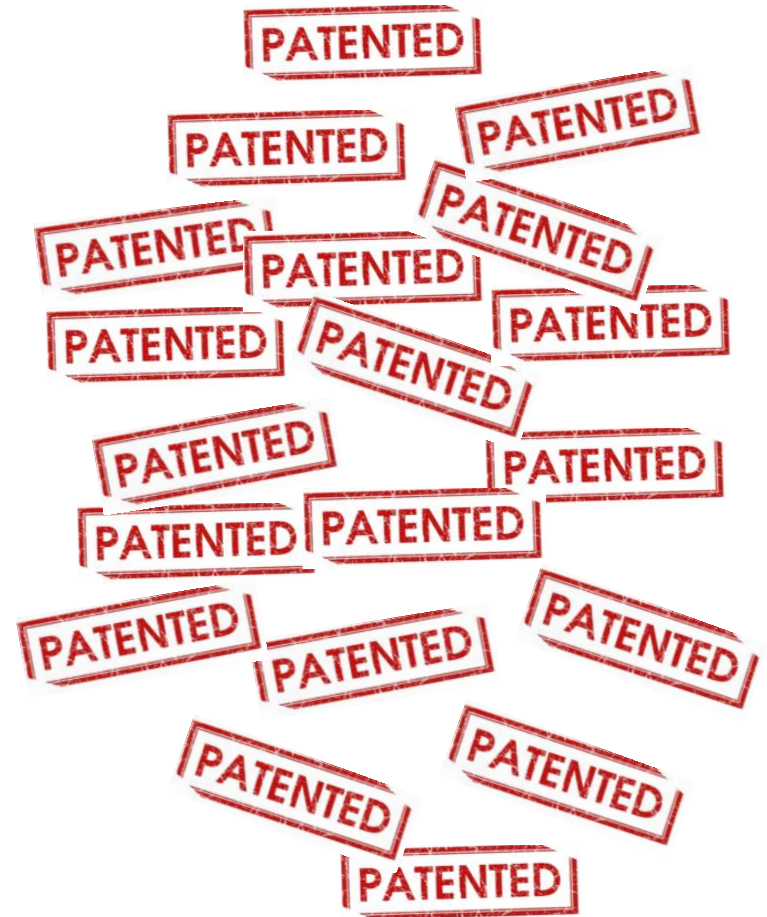


# The challenge

---

---

No classification systems available that allow the analysis of technologies.



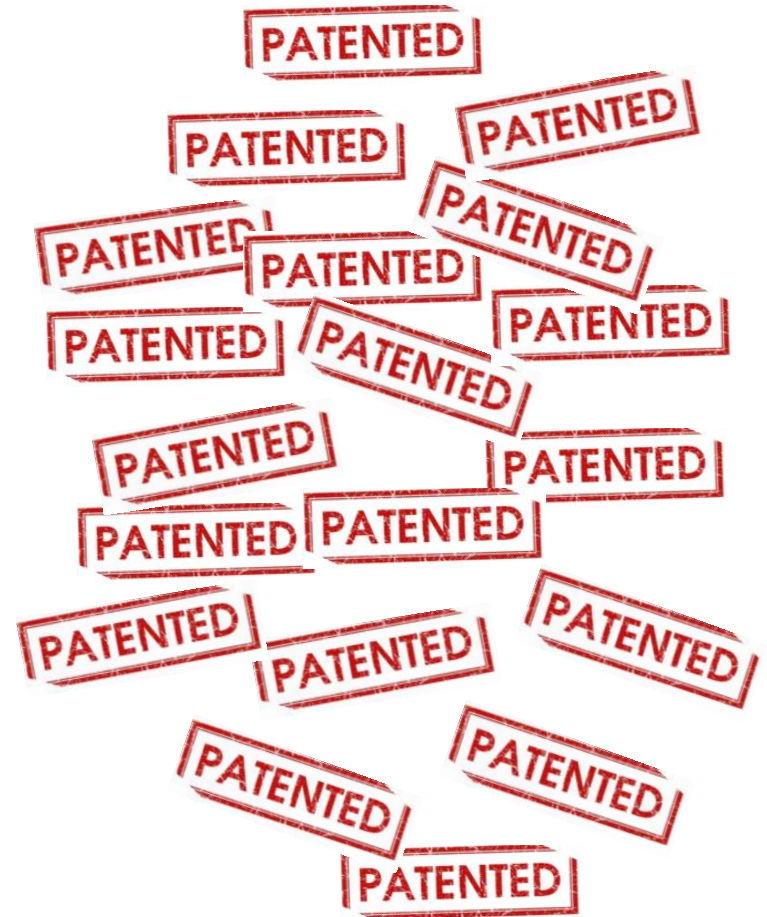
Source : [https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp\\_141325435.jpg](https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp_141325435.jpg)

# Our approach

---

## Criteria

Combination of certain indicators available within the patent system



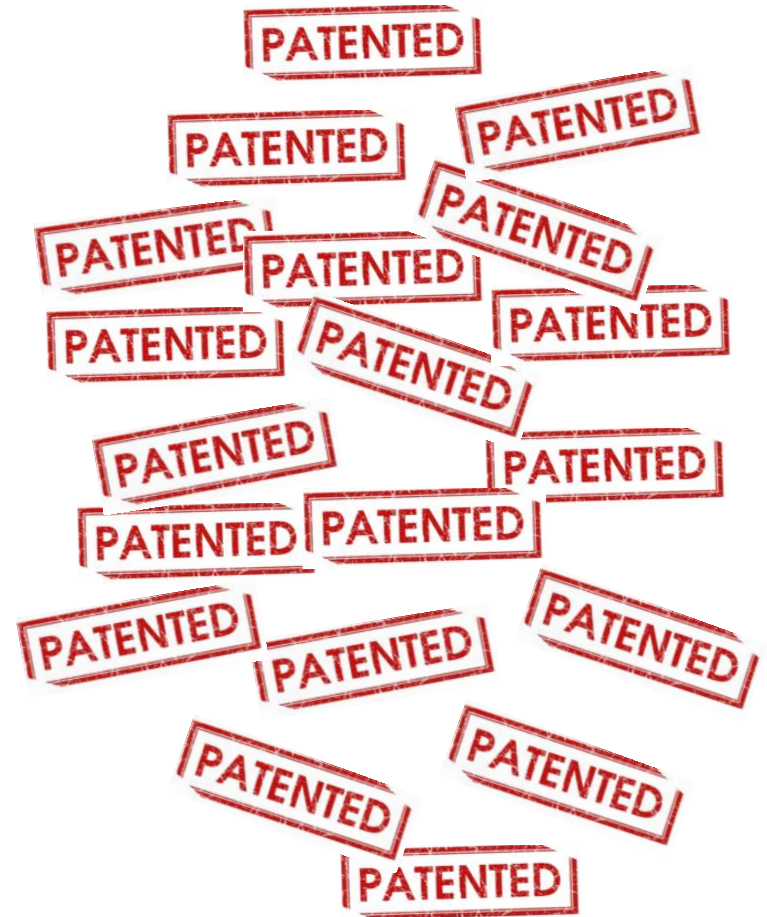
Source : [https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp\\_141325435.jpg](https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp_141325435.jpg)

# Our approach

---

## Criteria

1. Same company
2. Same inventor
3. Technologically related
4. Continuation of an earlier filing
5. Same technology field, i.e. WIPO35 list (Schmoch 2008)
6. Same IPC (International Patent Classification) 1-digit/4-digit class
7. Similarity in title and abstract
8. Related through citations (1st order, 2nd order)
9. Bibliographic coupling
10. Co-citation analysis

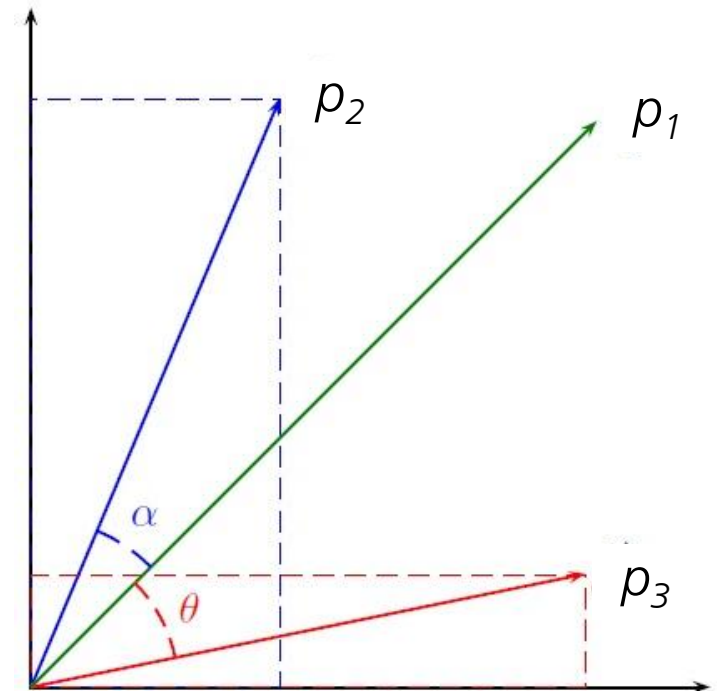


Source : [https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp\\_141325435.jpg](https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp_141325435.jpg)

# Our approach

## Comparison of Abstract and Title

- Vector space model with word dimensions
- Representation of abstract and title as a vector
- Documents similarity computation based on angles between vectors (cosine similarity)



Source : [https://en.wikipedia.org/wiki/Vector\\_space\\_model#/media/File:Vector\\_space\\_model.jpg](https://en.wikipedia.org/wiki/Vector_space_model#/media/File:Vector_space_model.jpg)

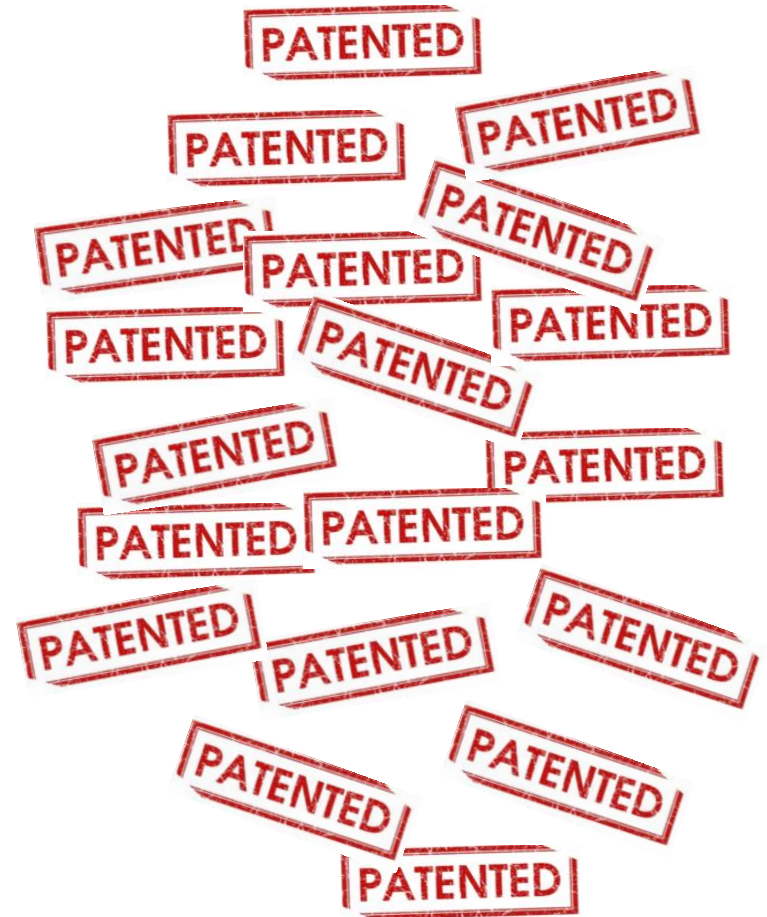
# Our approach

---

## Criteria

## Weighting Scheme

Weighting scheme for criteria



Source : [https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp\\_141325435.jpg](https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp_141325435.jpg)



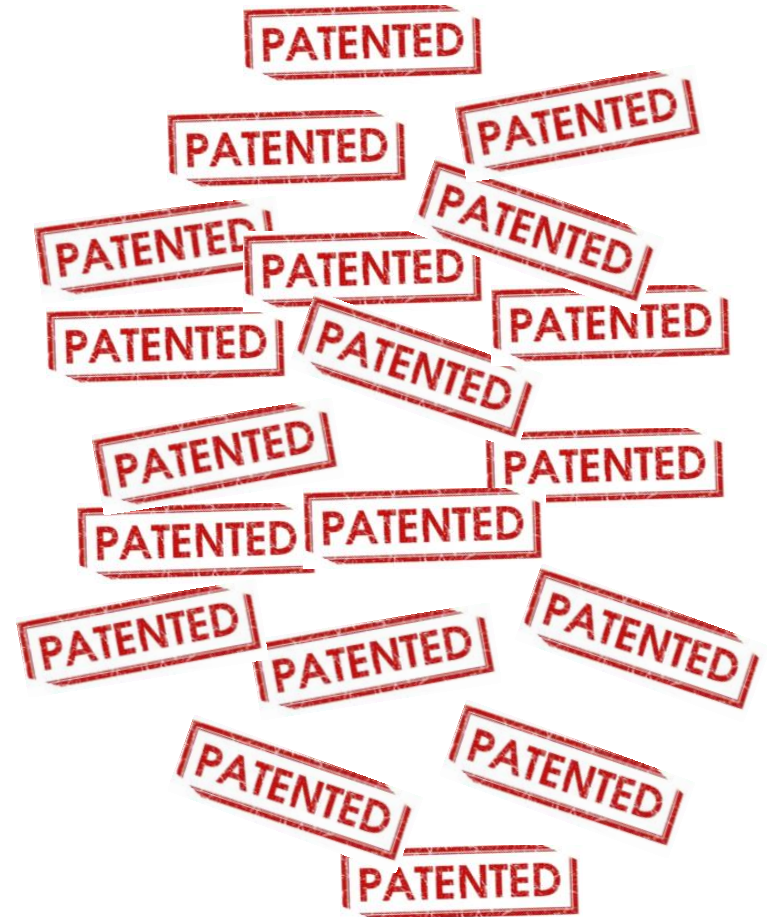
# Our approach

---

---

## Weighting Scheme

- Possible weighting scheme (initial):
  - Technical relation
  - Continuation
  - Direct citation
  - Abstract/Title
  - Indirect citation
  - Bibliographic coupling
  - Co-citation analysis
  - IPC4
  - WIPO class
  - IPC1



Source : [https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp\\_141325435.jpg](https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp_141325435.jpg)

# Our approach

## Weighting Scheme

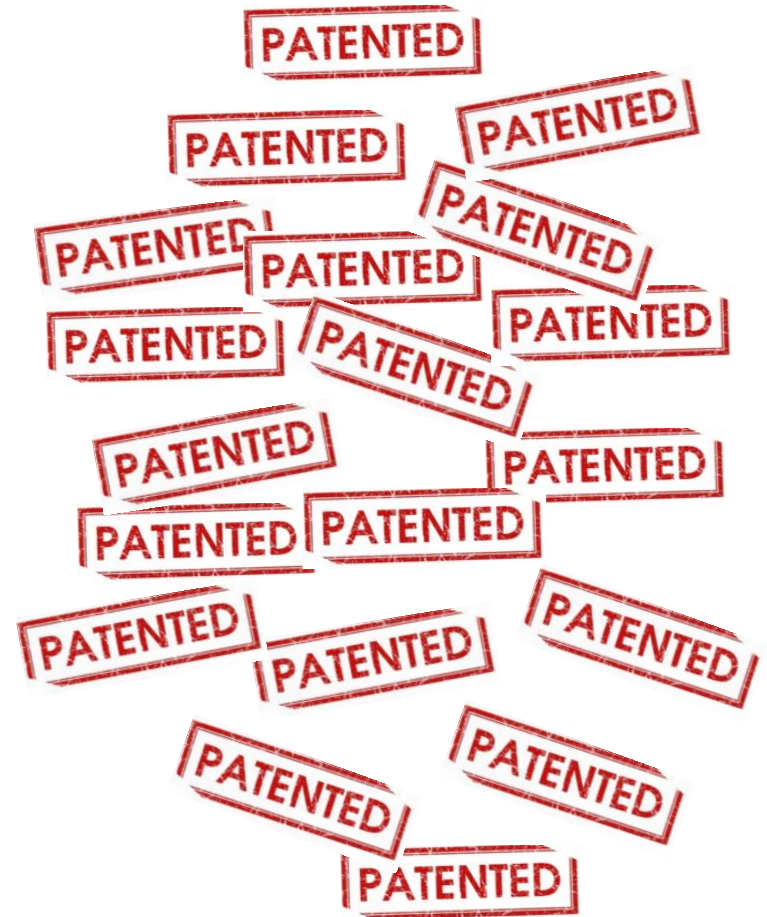
- Combination of weighted criteria into similarity measure
- Similarity between two patents  $p_1$  and  $p_2$ :

$$\text{sim}(p_1, p_2) = \frac{\sum_i w_i f_i(p_1, p_2)}{\sum_i w_i}$$

$f_i$  Similarity between  $p_{1i}$  and  $p_{2i}$  of criterion  $i$  for both patents

$w_i$  Weighting of the criterion

- Individual weighting scores for each criterion determined based on training data set



Source : [https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp\\_141325435.jpg](https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp_141325435.jpg)



# Our approach

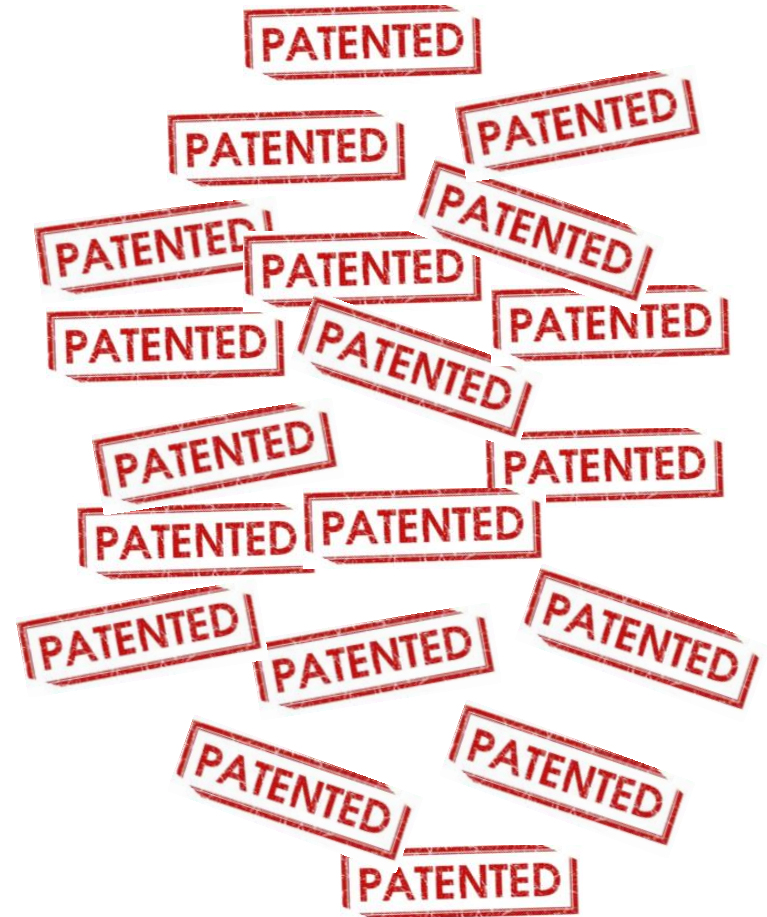
---

**Criteria**

**Weighting Scheme**

**Threshold**

Probability threshold for patent grouping



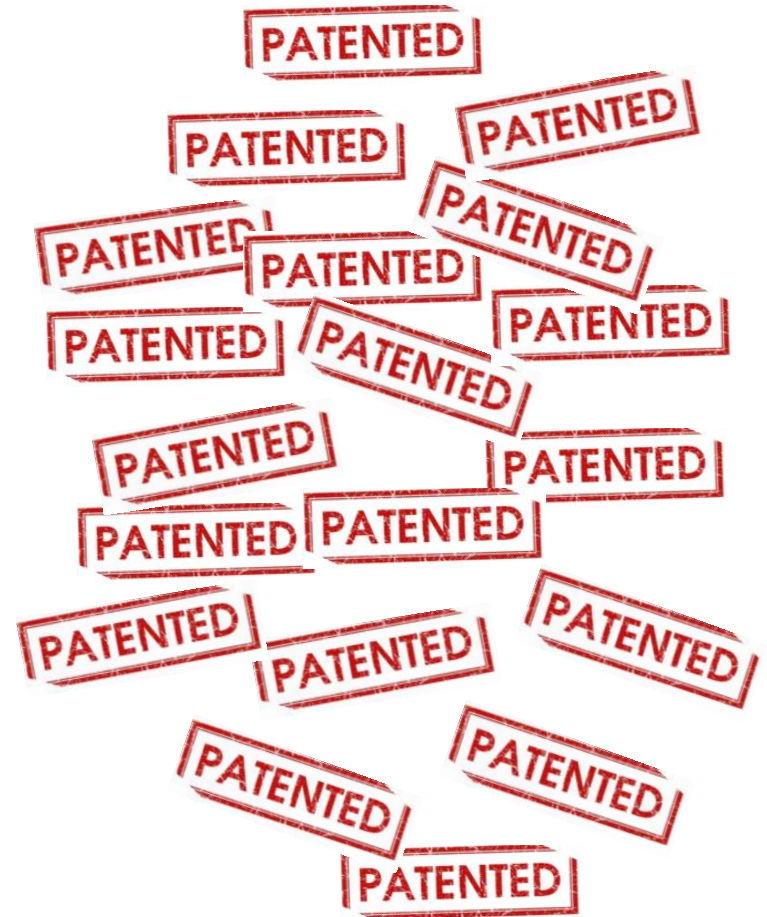
Source : [https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp\\_141325435.jpg](https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp_141325435.jpg)

# Our approach

---

## Threshold

- Threshold for patent similarity computation
- Only patents with a similarity above the threshold are grouped together



Source : [https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp\\_141325435.jpg](https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp_141325435.jpg)

# Our approach

---

Criteria

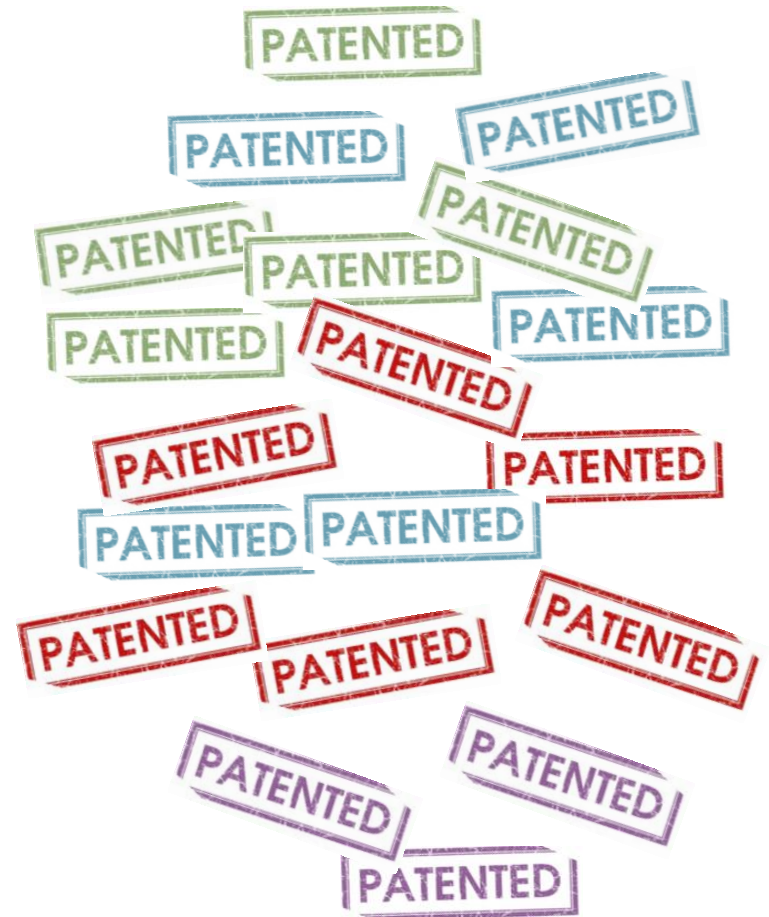
Weighting Scheme

Threshold



**Clustering**

Computation of probability that certain patents can be grouped together in order to form a „technology“



Source : [https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp\\_141325435.jpg](https://www.agra-net.com/agra/incoming/article438325.ece/ALTERNATES/w620/patented+stamp_141325435.jpg)

# Our approach

## Clustering

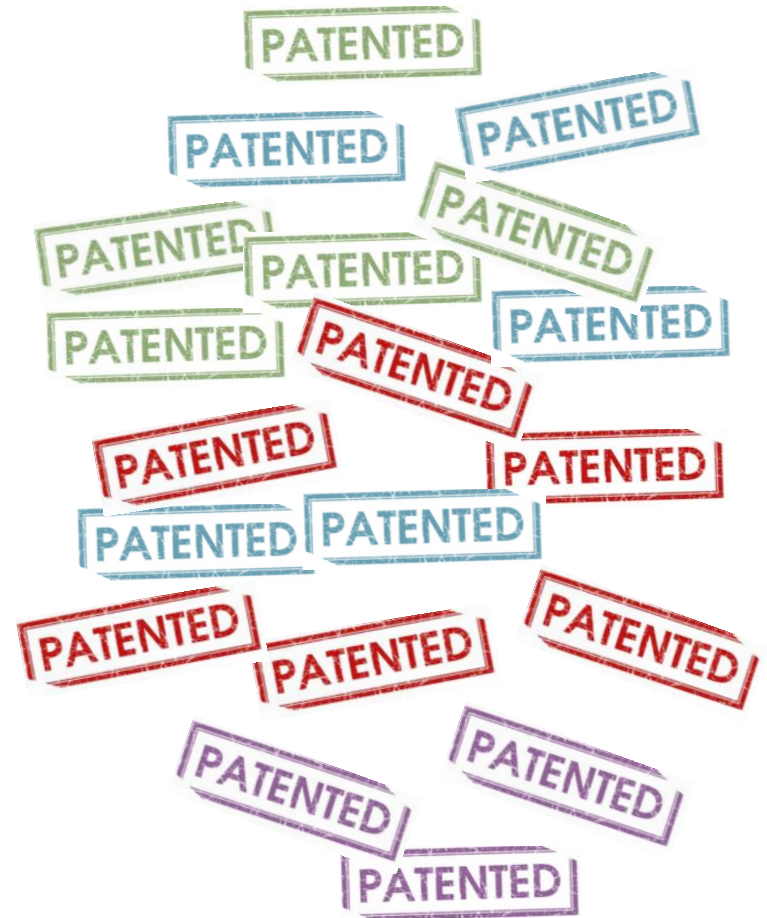
- Determination of similarities between patents

$$\text{sim}(p_1, p_2) = \frac{\sum_i w_i f_i(p_1, p_2)}{\sum_i w_i}$$

$f_i$  Similarity between  $p_{1i}$  and  $p_{2i}$  of criterion  $i$  for both patents

$w_i$  Weighting of the criterion

- Each two patents with a similarity higher than the given threshold are grouped together (Soft clustering)



---

# Future Work

---

---

- Create a training data and a test data set
  - Determine optimized weighting scores
  - Determine optimized threshold
- Evaluation of our system
- Revision of criteria/weighting
- Soft clustering vs. Hard clustering

# References

---

---

Blind, K. & Edler, J. & Frietsch, R. & Schmoch, U. (2006): Motives to patent: Empirical evidence from Germany, *Research Policy*, 35, 655-672.

Manning, C. D. & Raghavan, P. & Schütze, H. (2008). Introduction to Information Retrieval, *Cambridge University Press*.

Neuhäusler, P. (2012): The use of patents and informal appropriation mechanisms - Differences between sectors and among companies, *Technovation*, 32, 681-693.

Schmoch, U. (2008): Concept of a Technology Classification for Country Comparisons. *Final Report to the World Intellectual Property Office (WIPO)*, Karlsruhe: Fraunhofer ISI.

Shapiro, C. (2001): Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard-Setting, Available at SSRN: <http://ssrn.com/abstract=273550> or <http://dx.doi.org/10.2139/ssrn.273550>

---

---

# Thank you for your attention!