
Visualization and Analysis of Multidisciplinary Research Networks

A Case Study of UNIST's College of Information & Biotechnology

Sumin Park

20211113

2024.06.07



Motivation

increase in interdisciplinary studies -> topics are not confined to the department -> difficult to find **all labs** in fields of interest



UNIST freshman student interested in
HCI and Human factor

BME	BCI lab
	DNCE lab
	Color lab
	...
DES	Emotion lab
	DECS lab
	IPD lab
	...
CSE	iVHCI lab
	TACT lab
	...

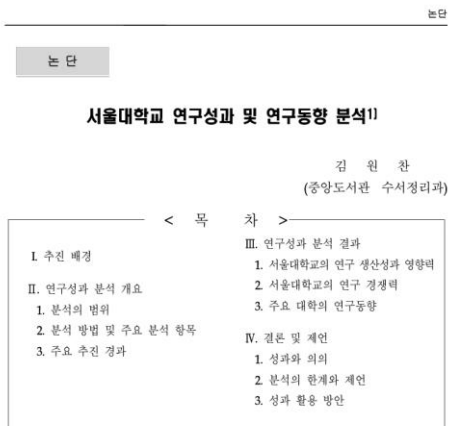
Goal

- Graphical representation and analysis of each lab's research interest and overall research trend at the UNIST College of Information & Biotechnology (IB).
- Presentation of analysis methods in the field of study



Related works

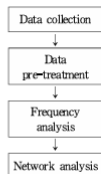
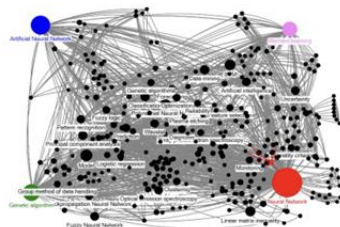
1. How to analyze research trends



2. Using graph method

Analysis of major research trends in artificial intelligence through analysis of thesis data

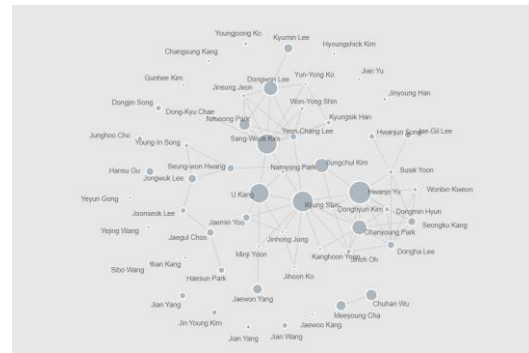
Myoung-Sug Chung, Seong-Hyeon Park, Byeong-Hoon Chae, Joo-Yeoun Lee
New industry convergence technology R&D center, Ajou university



Data analysis process

[Fig. 4] An Analysis of the Keyword Network for 20 Years of Global Catalog

3. Research Trends Analysis Site



<https://alinlab.kaist.ac.kr/KCSS>



Data

Communities **Researchers & Labs** **Titles** **Issue Date**

Lim, Chiehyeon
임치현
Department of Industrial Engineering (산업공학과)
Service Engineering & Knowledge Discovery Lab. Website

Research Interests Service Engineering; Knowledge Discovery; 서비스 엔지니어링; 지식 디스커버리

Lab Description We solve service problems in industries and develop knowledge discovery methods.
1. Service Engineering: Service intelligence development- Recommendation system development: Service quality and efficiency analysis- Service improvement and optimization- Service system modeling and design
2. Knowledge Discovery: Text mining: Bio and health data mining: Behavioral data mining: Clustering method development: Featuring method development

Determining Directions of Service Quality Management using Online Review Mining with Interpretable Machine Learning
Shin, Jongkyung, Joang, Junegak, Lim, Chiehyeon
Issue Date 2024-04 View 12

MMP Net: A feedforward neural network without intermediate outputs
Cho, Hyeon, Kim, Ryeongsan, Yoon, Kihyeon
Issue Date 2023-09 View 9

Chemistry-informed machine learning
Kim, Yareon, Lim, Chiehyeon, Lee, Janghye
Issue Date 2023-06 View 11

From technology enablers to circular in Industry 4.0
Kim, Minjun, Lim, Chiehyeon, Hsuan, Julian
Issue Date 2023-06 View 8

Author(s) Shin, Jongkyung, Joang, Junegak, Lim, Chiehyeon
Issued Date 2024-04
DOI 10.1016/j.ijhm.2023.103684
URI https://scholarworks.unist.ac.kr/handle/201301/81359
Citation INTERNATIONAL JOURNAL OF HOSPITALITY MANAGEMENT, v.118, pp.103684
Abstract Determining the importance values of service features is necessary to prioritize the points in service quality management and improvement. Existing studies have used linearly additive relationship models to estimate service feature importance, such as linear and logistic regression. This traditional approach is interpretable but often limited in terms of model fitness and prediction performance. Meanwhile, modern advanced machine learning models provide high fitness and performance but often lack interpretability. Thus, to achieve both reliable prediction and interpretation, we propose a systematic framework for estimating the importance of service features using online review mining with interpretable machine learning. An interpretable machine learning-based method is proposed to estimate the importance values of features by applying the shapley additive global importance metric to the highest performance prediction model. We validate the superiority of our framework over existing methods through a case study on the global importance estimation of hotel service features in Singapore. To facilitate additional applications, we offer the implementation code of our work at https://github.com/JK-SHIN-PG/OnReviewServImprovement.

Publisher Pergamon Press
ISSN 0278-4319
Keyword (Author) Customer needs, Customer reviews, Explainable artificial intelligence, Feature importance, Interpretable machine learning, Service management

92 laboratory (8 departments)

->Node

1. Lab URL
2. Professor name
3. Lab name
4. Field [IE, BME, EE,,]
5. research interest
6. Lab description
7. Title and Keywords of Recent 20 Articles (1388 papers)

->Feature

[Web Crawling]





Process

Preprocessing

Noise
(special characters,
mixed case usage)

Vectorizer

TF-IDF
BERT Model

Similarity

Representation

2d / 3d

Clustering

K-means
DBSCAN

+ **searching part**



Process

Model Selection

TF-IDF

Language model

Term Frequency
- Inverse Document Frequency

BERT model

BERT : textual embedding
(sci_BERT)
Strengths in **understanding the context** of word use

K-means

Clustering model

When data is well separated into
spherical cluster.
Based on **distance between
cluster centers** and data point

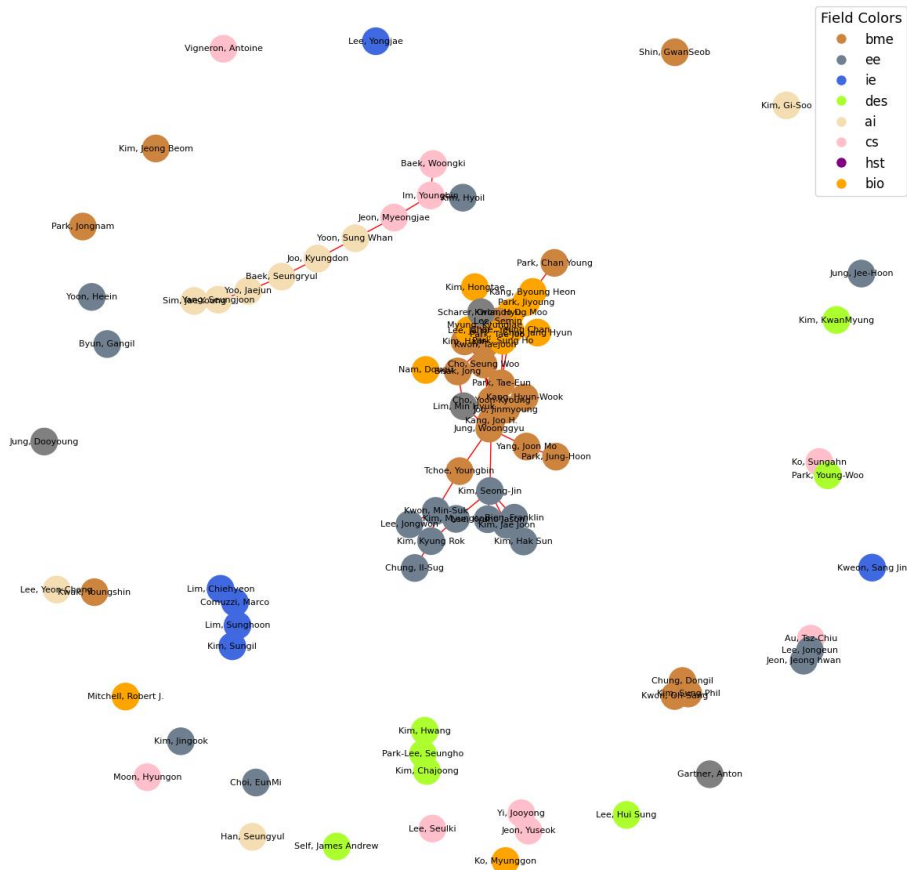
DBSCAN

when shape of cluster is arbitrary
May not work when **difference between
clusters are not large**

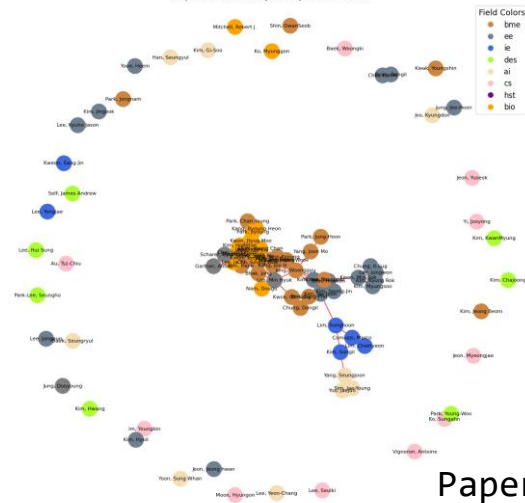


Result

Combined Graph Visualization of Keywords and Lab Descriptions by Field with SciBERT

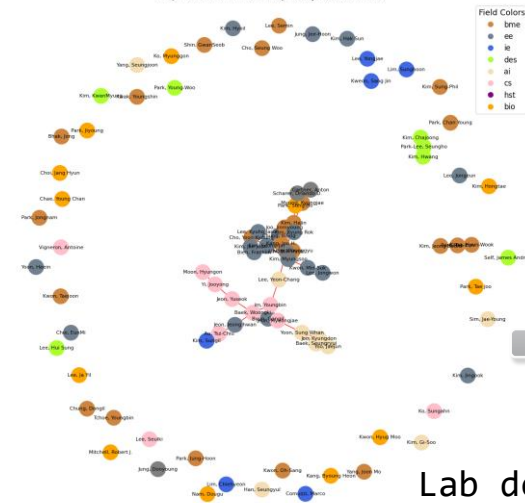


Graph Visualization of Keywords by Field with SciBERT



Paper keywords

Graph Visualization of Lab Descriptions by Field with SciBERT

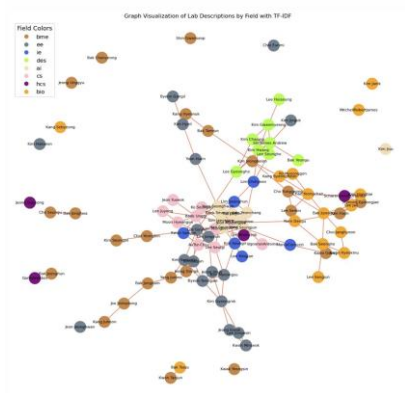


Lab description

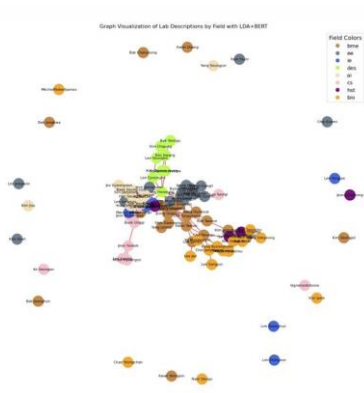
Results

Sci_BERT

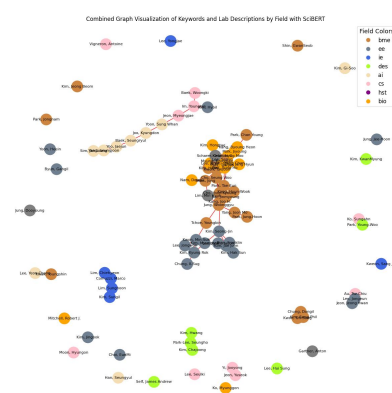
Bidirectional Encoder Representations from Transformers



TF-IDF
Silhouette Score:
0.0148



BERT
Silhouette Score:
0.1094



SciBERT
Silhouette Score:
0.0597

SciBERT (Scientific BERT):

Pretrained model (data of Biomedical & Computer science paper)



Result

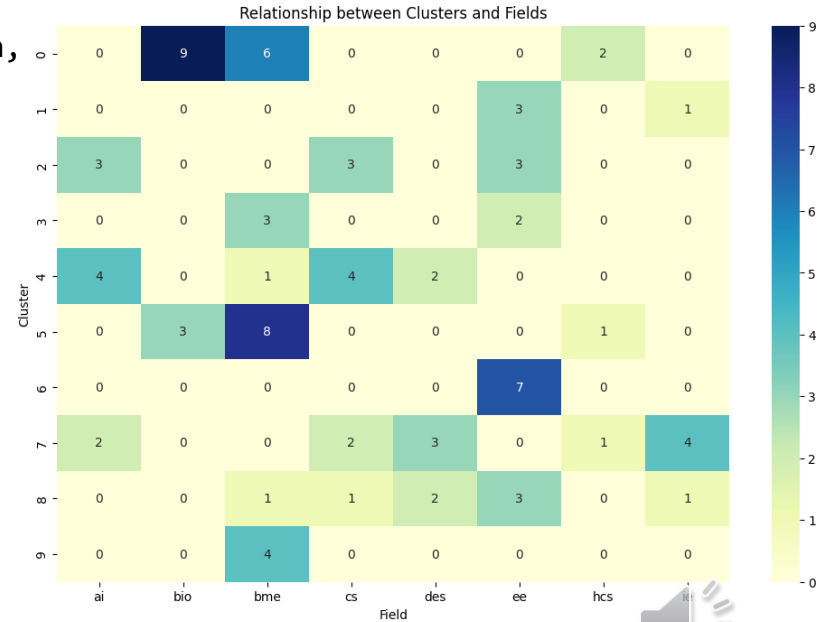


```
search_keyword = 'hci'
>>
Most similar cluster for 'hci':
Cluster 8 with similarity 0.7826
```

Cluster 8 keywords : 'interaction, mobility, future, electronic, humanrobot, driving, theory, robots, robotics, self',

UNIST freshman student interested in
HCI and Human factor

Researcher Name	Keywords	Field
Park, Young-Woo	MENTALHEALTHCARE, INTERVENTIONS, METHODOLOGY, DEPENDENT MOVEMENT, SURFACE	des
Kim, KwanMyung	ELECTROMYOGRAPHY, CHIN TUCK, ELASTIC MEMRISTOR, 2D	des
Kim, Myungsoo	material, hexagonal boron nitride, ...	ee
Moon, Hyungon	, , , , domainspecific architecture, instruc...	cs
Byun, Gangil	FRONT, DESIGN, PLANA R, BROADBAND, WAVE PLATE PHASE	ee



Conclusion

Represent the researchers' keywords and lab descriptions to identify research trends

Method: SciBERT embeddings and K-means clustering

Key Findings

SciBERT Embeddings: Specialized in scientific texts

Accurately reflect lab characteristics

K-means Clustering:

Forms clear and interpretable clusters

More effective than DBSCAN in forming meaningful clusters



Conclusion

1. **Help to Choose Department & Laboratory for UNIST Students**
2. **Suggest Analysis Method for Multidisciplinary Research:**
3. **Provide Motif for Multidisciplinary Research:**
4. **Reference in Department Classification Process:**



Discussion

limitation

1. Data limitation
2. Limitation of clustering algorithms
3. Limitation of embedding model
4. Subjectivity in cluster interpretation

Future work

1. Comparison with other university or Global research trend
2. Dynamic Graph Representation with this data (change of research area and lab – detect trend)
3. Experiment with Additional Embedding Models:
4. Automatic Cluster Labeling:

