

**ĐẠI HỌC QUỐC GIA TP.HCM  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**BÁO CÁO HỌC THUẬT LUẬN ÁN - LẦN 1**

**PHÂN TÍCH MẠNG XÃ HỘI THEO CHỦ ĐỀ VÀ ỨNG DỤNG  
VÀO CÔNG TÁC TƯ VẤN HỌC TẬP CHO SINH VIÊN  
(PHÂN TÍCH MẠNG XÃ HỘI DỰA THEO MÔ HÌNH CHỦ ĐỀ  
VÀ ỨNG DỤNG)**

**Cán bộ hướng dẫn khoa học: PGS.TS. Đỗ Phúc  
Nghiên cứu sinh. Hồ Trung Thành**

# Nội dung

---

Dẫn nhập

Tổng quan về phân tích mạng xã hội

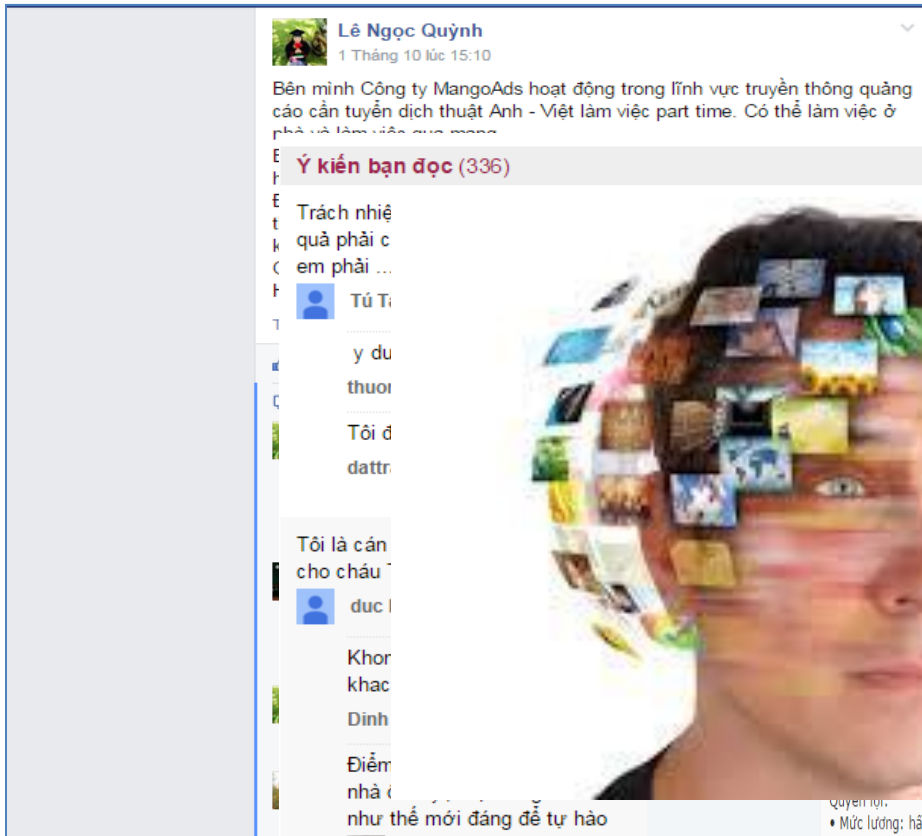
Giới thiệu mô hình chủ đề

Tổng quan nghiên cứu của luận án

Các nghiên cứu liên quan và bài toán đặt ra

Tài liệu tham khảo

# Dẫn nhập



**Lê Ngọc Quỳnh**  
1 Tháng 10 lúc 15:10

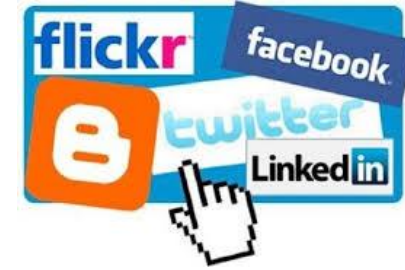
Bên mình Công ty MangoAds hoạt động trong lĩnh vực truyền thông quảng cáo cần tuyển dịch thuật Anh - Việt làm việc part time. Có thể làm việc ở nhà và làm việc của mình

**Ý kiến bạn đọc (336)** Mới nhất | Quan tâm nhất

Trách nhiệm  
quả phải c  
em phải ...

Tú T:  
y du  
thuoi  
Tôi đ  
dattr

Tôi là cán  
cho châu  
duc l  
Khor  
khac  
Dinh  
Điểm  
nhà c  
như thế mới đáng để tự hào



#1

anh của công ty, công ty TNHH truyền hình cáp Saigontourist

chăm sóc khách hàng và phát triển kinh doanh.  
ia các đơn vị trong công ty.

rver,...).

**Zzz.CÔNG DÂN GƯƠNG M**

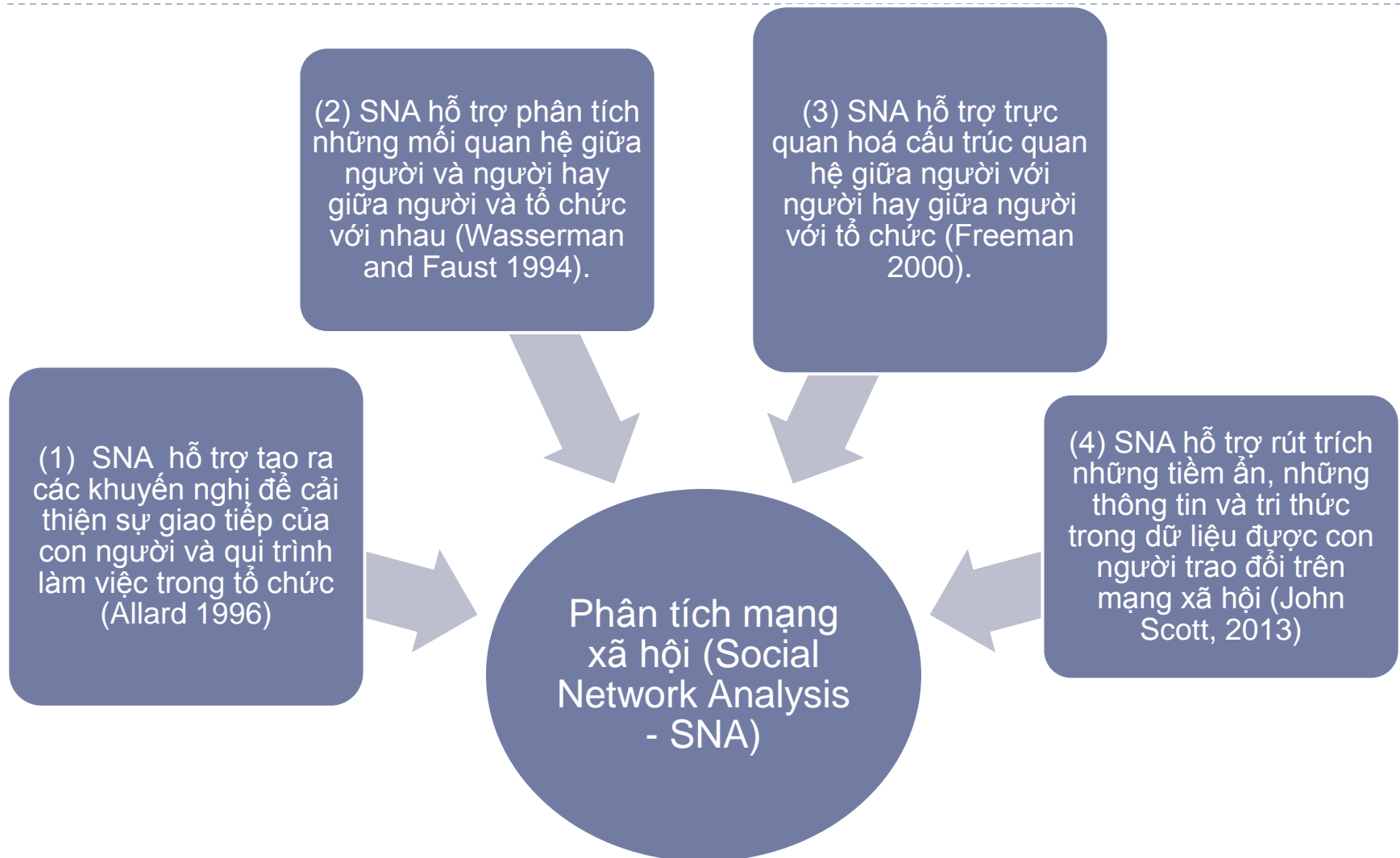
Đại Học Y Dược Huế phải chịu hoàn t  
hưởng đến tương lai của cả 1 người.

**Tùng Billy** - 10:57 09/10

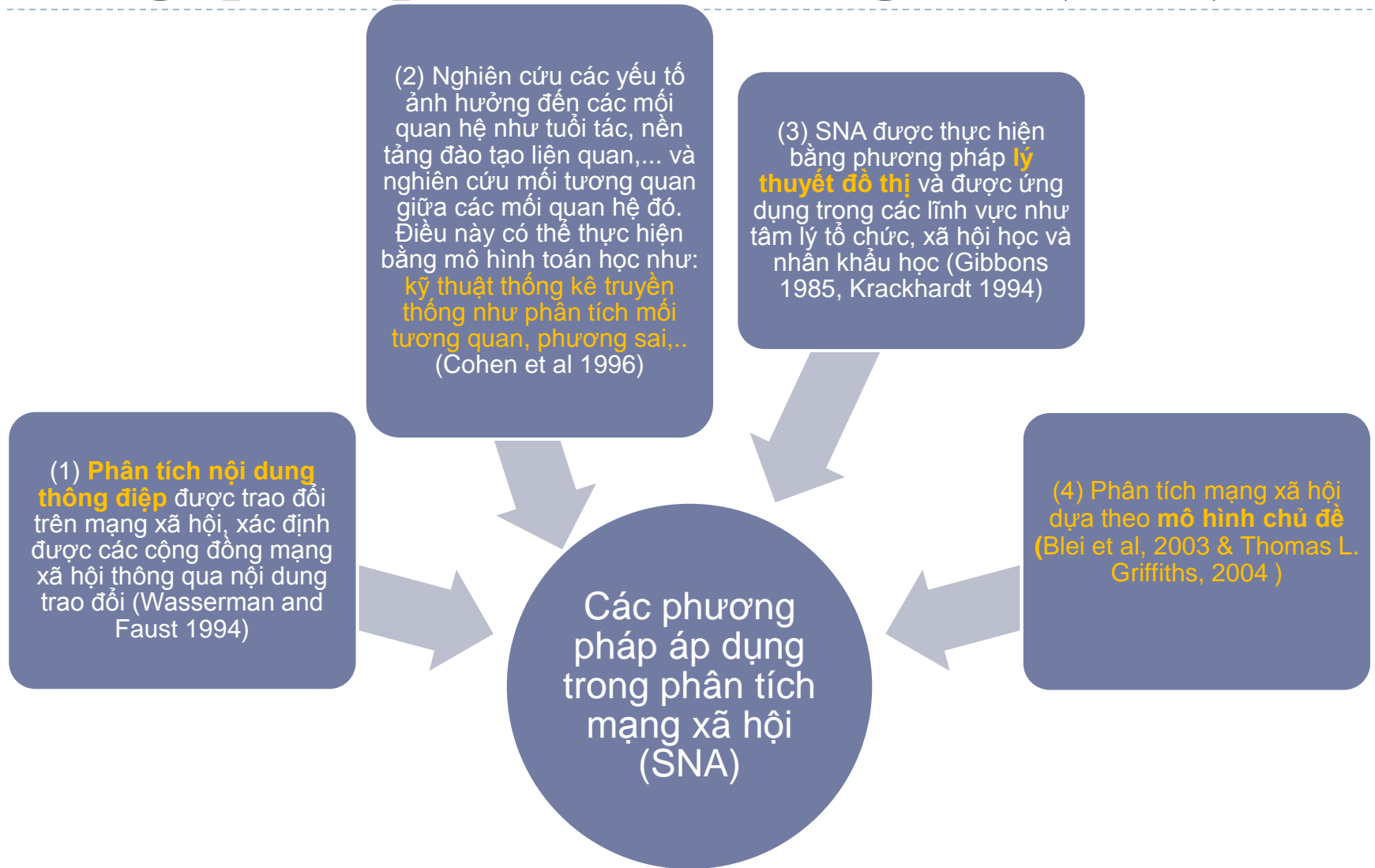
- Quyền lợi:
- Mức lương: hấp dẫn.
- Môi trường làm việc năng động.
- Được làm việc trực tiếp với đội ngũ kỹ sư lập trình nhiều kinh nghiệm.
- Được hưởng các chế độ BHXH và BHYT theo quy định của nhà nước.
- Được hưởng các chế độ phúc lợi theo quy định của công ty như: Bảo hiểm Sức khỏe, Thưởng Quý, Lễ, Tết,....

Hồ sơ gửi trực tiếp qua địa chỉ email: anh.vtt@sctv.com.vn

# Tổng quan phân tích mạng XH (SNA)



# Tổng quan phân tích mạng XH (SNA)



# Chủ đề

## Chủ đề là gì?

Theo từ điển Cambridge, chủ đề là một vấn đề được thảo luận, viết hay nghiên cứu.

Theo từ điển Oxford, chủ đề là một vấn đề được trình bày trong văn bản, bài luận hay trong cuộc hội thoại

## Chủ đề tiềm ẩn là gì?

Là chủ đề chưa biết (chưa được gán nhãn) trong quá trình tạo lập văn bản của người dùng.

## Chủ đề quan tâm là gì?

Là chủ đề được người dùng quan tâm trao đổi liên quan đến một lĩnh vực cụ thể nào đó.

# Mô hình chủ đề

---

Mô hình chủ đề được Deerwester cùng cộng sự đề xuất năm 1990, sau đó các nghiên cứu của Hofmann, 1999 và Blei et al, 2003.

Mô hình chủ đề cho phép kiểm tra và khai thác tập tài liệu văn bản dựa trên việc tìm kiếm và thống kê các từ có liên quan đến chủ đề trong mỗi tài liệu, và khám phá ra những chủ đề tiềm ẩn trong tài liệu văn bản đó.

Một số tiếp cận hiện nay trong việc mô hình nội dung tài liệu dựa trên ý tưởng tính phân bố xác suất của mỗi từ đặc trưng trong tài liệu. Phân bố này xem văn bản là hỗn hợp nhiều chủ đề, mỗi chủ đề là sự kết hợp của nhiều từ kèm phân bố xác suất riêng cho từng từ trong chủ đề.

Trong cách tiếp cận phân tích mạng xã hội theo chủ đề, các nghiên cứu chủ yếu tập trung xây dựng mô hình toán học, ước lượng tham số mô hình dựa trên nền tảng mạng xác suất Bayes.

# Các nghiên cứu đầu tiên về Mô hình chủ đề

---

Mô hình LSI -  
Latent Semantic  
Indexing  
(Deerwester et al,  
1990)

Mô hình PLSI -  
Probabilistic  
Latent Semantic  
Indexing (Thomas  
Hofmann, 1999)

Mô hình LDA -  
Latent Dirichlet  
Allocation (Blei et  
al, 2003)



# Mô hình chủ đề - LDA (Latent Dirichlet Allocation)

---

## Vấn đề “Sinh văn bản” (Bài toán thuận)

- Khi tạo lập thông điệp, người tạo lập (người viết văn bản) xác định trước chủ đề, sau đó xây dựng văn bản bằng cách chọn các từ xoay quanh chủ đề đã xác định

## Vấn đề “Khám phá chủ đề” (Bài toán ngược)

- Có văn bản, cần tìm các chủ đề mà người viết đã dựa trên đó để hình thành văn bản. Nghĩa là cần khám phá chủ đề tiềm ẩn trong nội dung thông điệp được người dùng trao đổi.

# Mô hình chủ đề - LDA (Latent Dirichlet Allocation)

Mô hình xác suất theo mạng Bayes 3 cấp: tài liệu, chủ đề và từ

- Mỗi tài liệu (document) được mô tả dưới dạng kết hợp ngẫu nhiên của một tập các chủ đề.
- Mỗi chủ đề (topic) là một phân bố rời rạc của một tập các từ vựng (words).

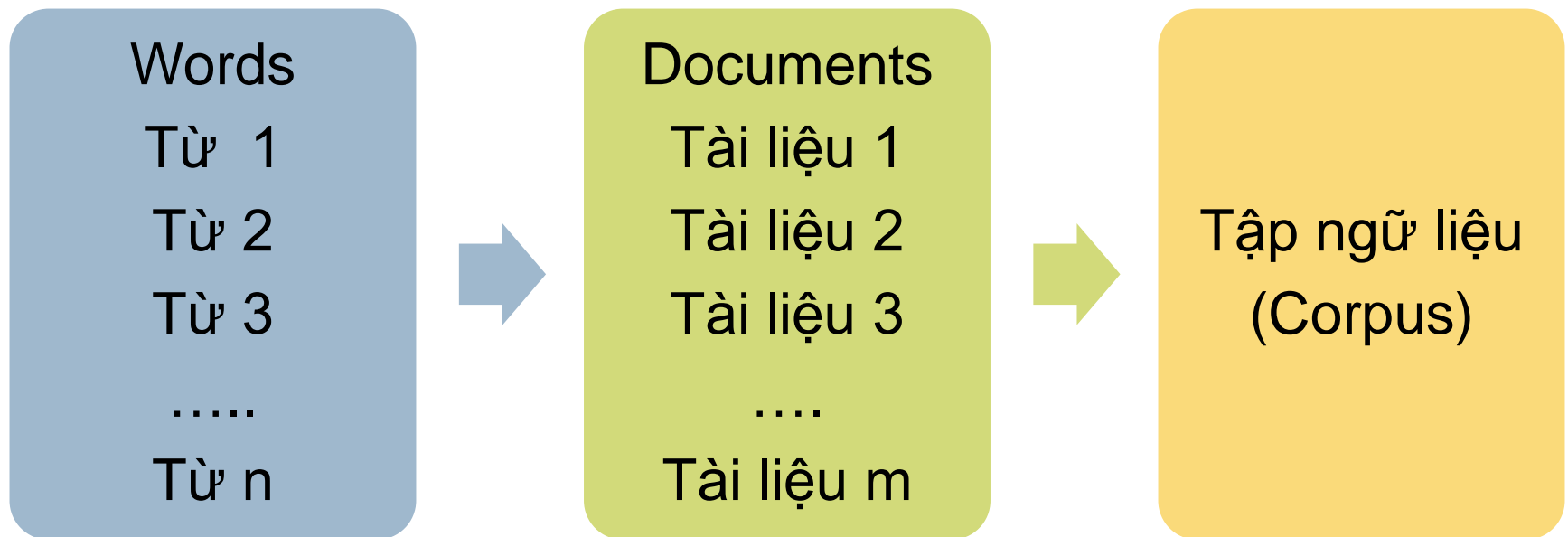
Mô hình sinh tài liệu

- Dựa trên việc rút trích tập từ đặc trưng để sinh tài liệu
- Khám phá chủ đề tiềm ẩn

Mô hình LDA là mô hình nền tảng và kết hợp vào nhiều mô hình phức tạp hơn.

# Mô hình chủ đề - LDA (sinh văn bản)

---



# Mô hình chủ đề - LDA (khám phá chủ đề)

Nếu **sinh viên** đang yêu thích **ngành nghiệp** trong lĩnh vực **kinh tế** hoặc yêu thích công nghệ thông tin, thì **chương trình đào tạo** của Khoa thực sự phù hợp với bạn. Khi **tốt nghiệp đại học** sau 4 năm, với bằng tốt nghiệp đại học là **cử nhân** kinh tế, ngành Hệ thống thông tin quản lý, chắc chắn bạn sẽ có rất nhiều **cơ hội việc làm**, **lương** cao và **khả năng** phát triển trong nhiều lĩnh vực **ngành nghiệp** khác nhau và có cơ hội **học tập** nâng cao hoặc có thể tham gia giảng dạy hay **du học** **nước ngoài** theo những chương trình **hợp tác quốc tế**, **liên kết** các trường đại học trên **thế giới** của Trường ....

Chủ đề 01	
Từ	Xác suất
Chương trình	0.92
Đào tạo	0.78
Đại học	0.71
Cử nhân	0.61
Giảng dạy	0.59
Sinh viên	0.52
.....	.....

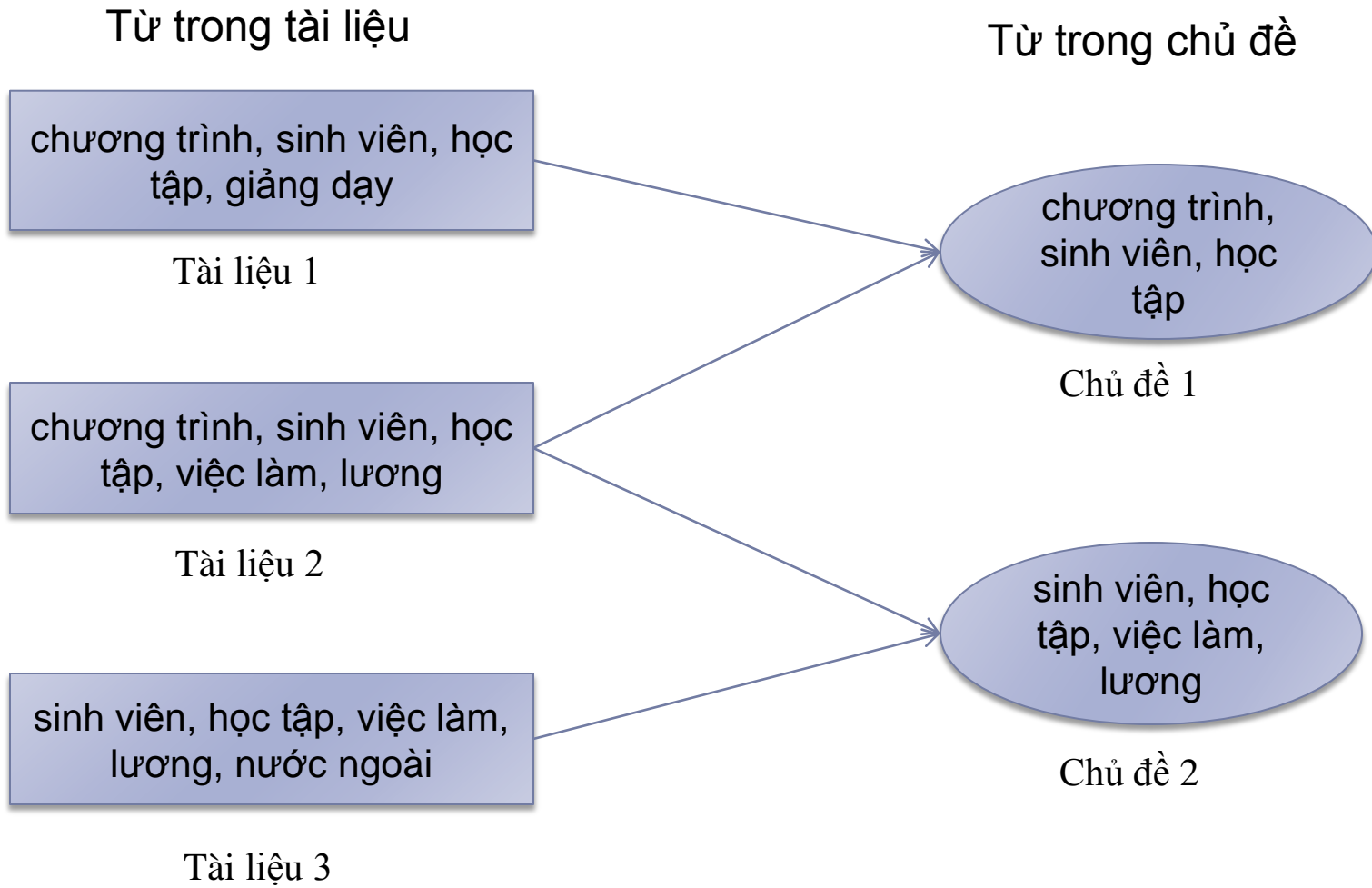
Chủ đề 02	
Từ	Xác suất
Nghề nghiệp	0.91
Việc làm	0.87
Lương	0.72
Cơ hội	0.68
Khả năng	0.54
.....	.....

Chủ đề 03	
Từ	Xác suất
Du học	0.83
Nước ngoài	0.72
Thế giới	0.69
Quốc tế	0.43
Hợp tác	0.41
.....	.....



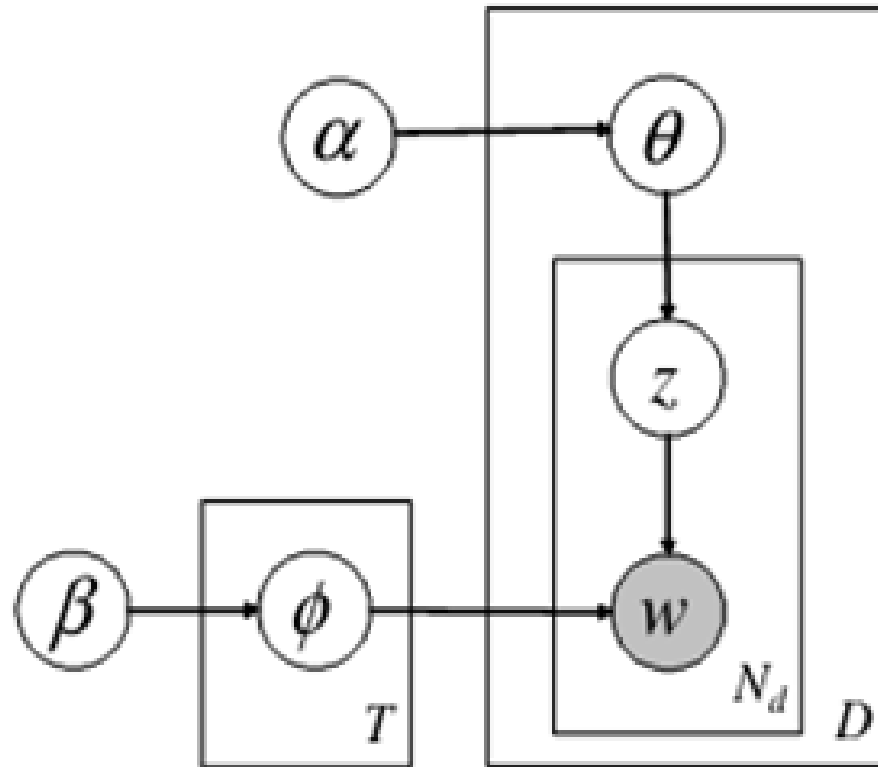
# Mô hình chủ đề - LDA (khám phá chủ đề)

---



# Mô hình chủ đề - LDA

---



# Tổng quan nghiên cứu của luận án - Mục tiêu nghiên cứu

---

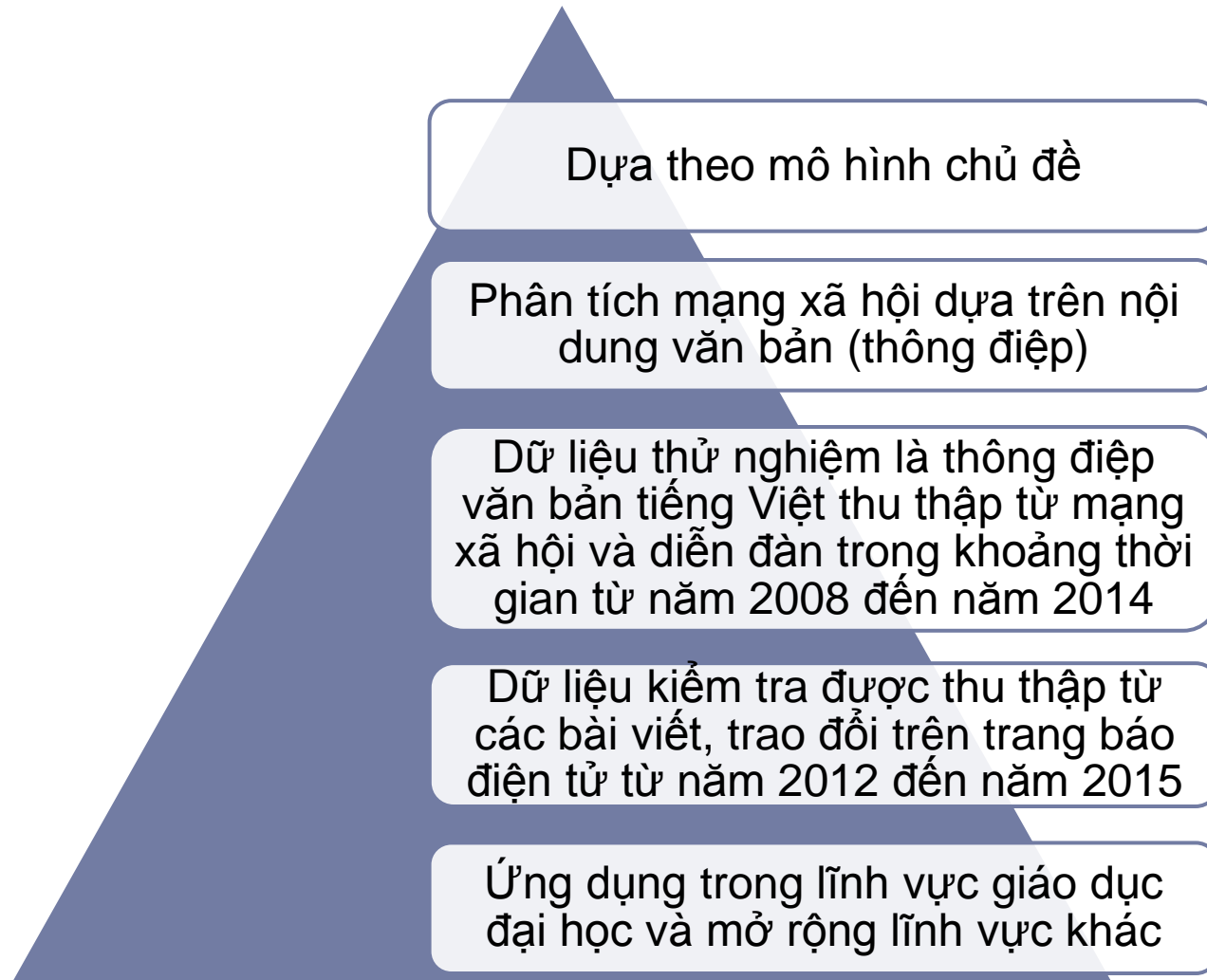
Bài toán 1. Đề xuất mô hình khám phá chủ đề tiềm ẩn trong văn bản và gán nhãn chủ đề dựa theo mô hình chủ đề.

Bài toán 2. Đề xuất mô hình khám phá chủ đề quan tâm của người dùng dựa theo mô hình chủ đề có yếu tố thời gian.

Bài toán 3. Đề xuất mô hình khám phá cộng đồng người dùng trên mạng xã hội dựa theo mô hình chủ đề có yếu tố thời gian và khảo sát sự thay đổi chủ đề quan tâm của cộng đồng mạng xã hội.

# Tổng quan nghiên cứu của luận án - Phạm vi nghiên cứu

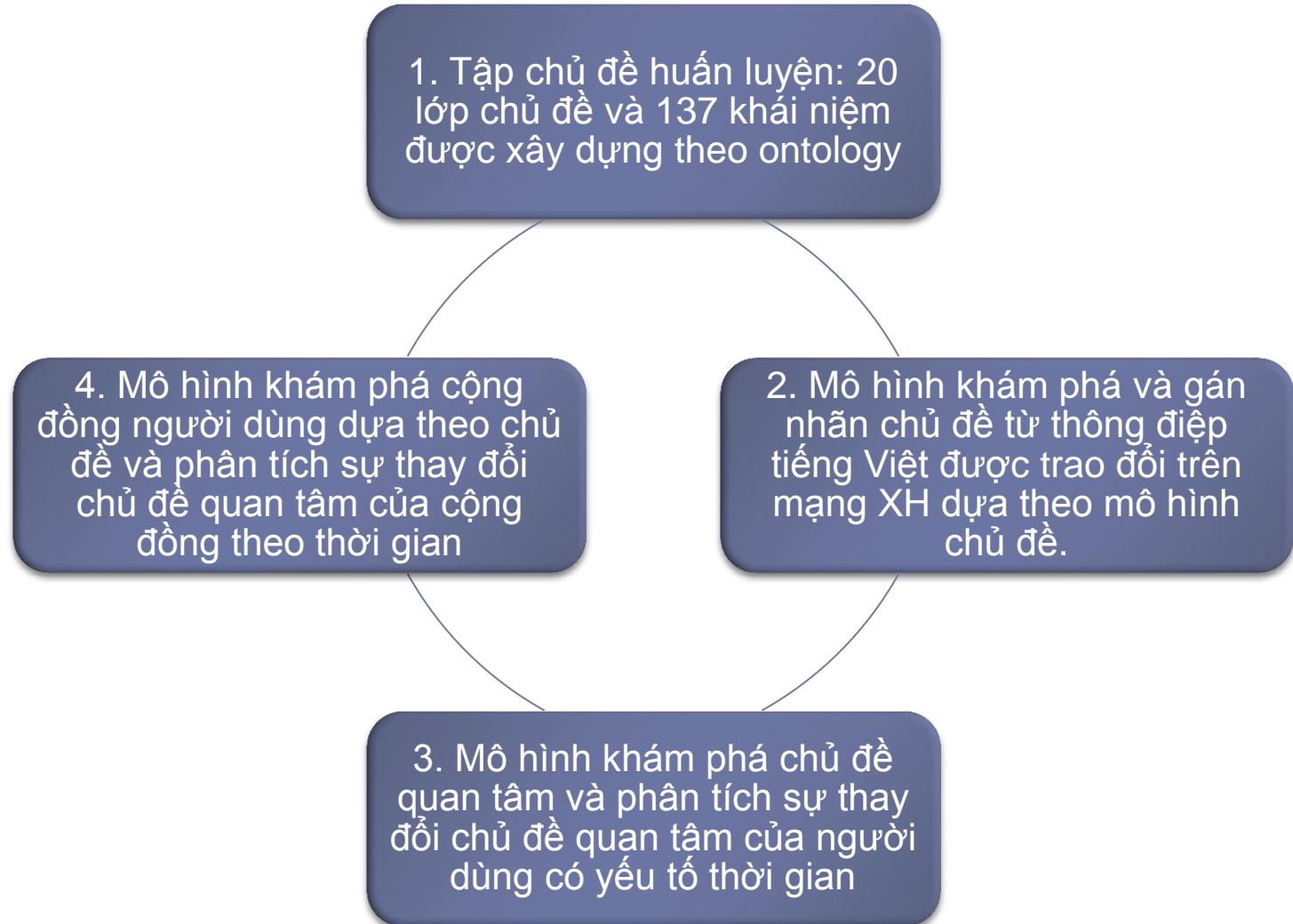
---





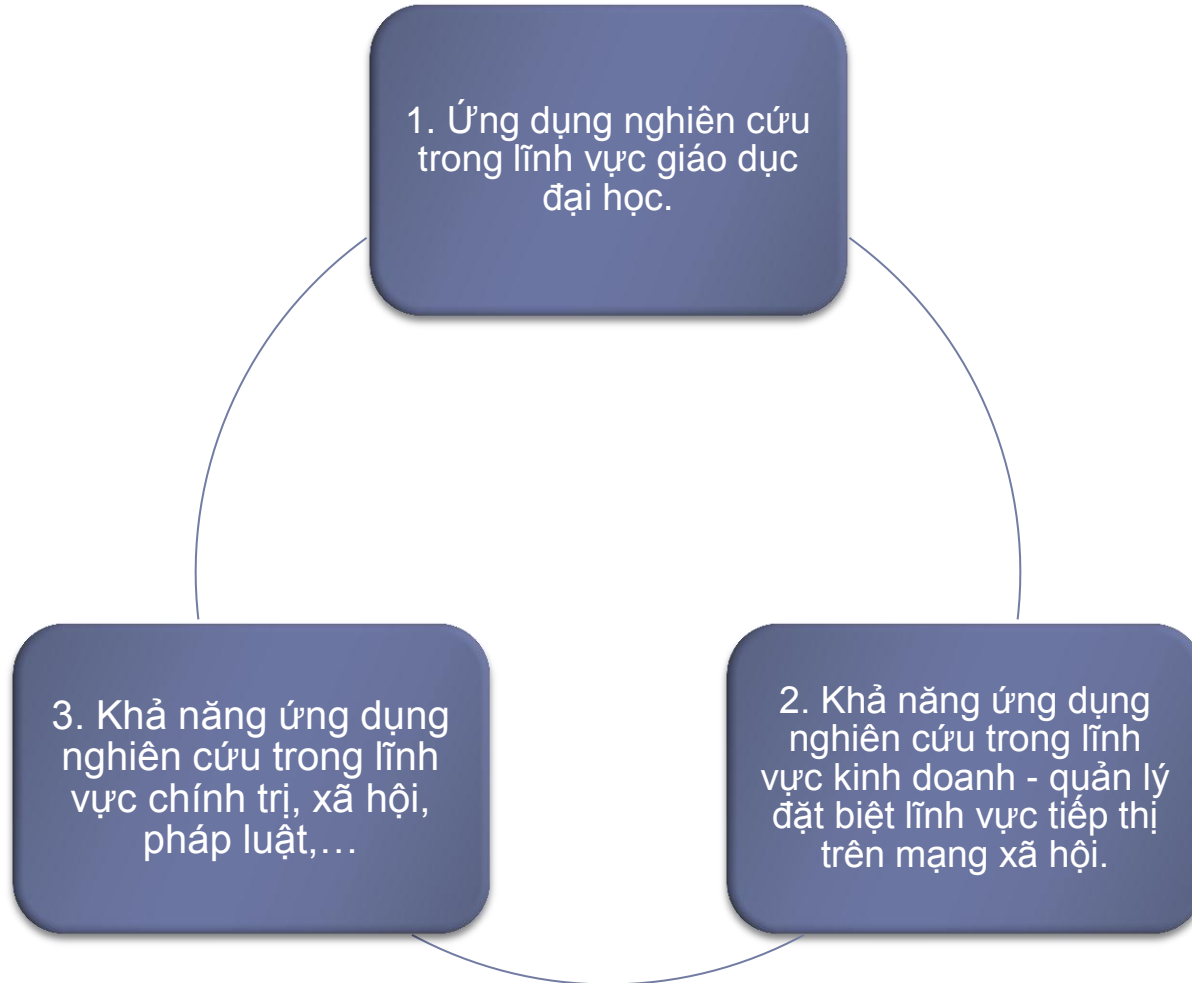
# Ý nghĩa nghiên cứu - Về khoa học

---



# Ý nghĩa nghiên cứu - Về thực tiễn

---

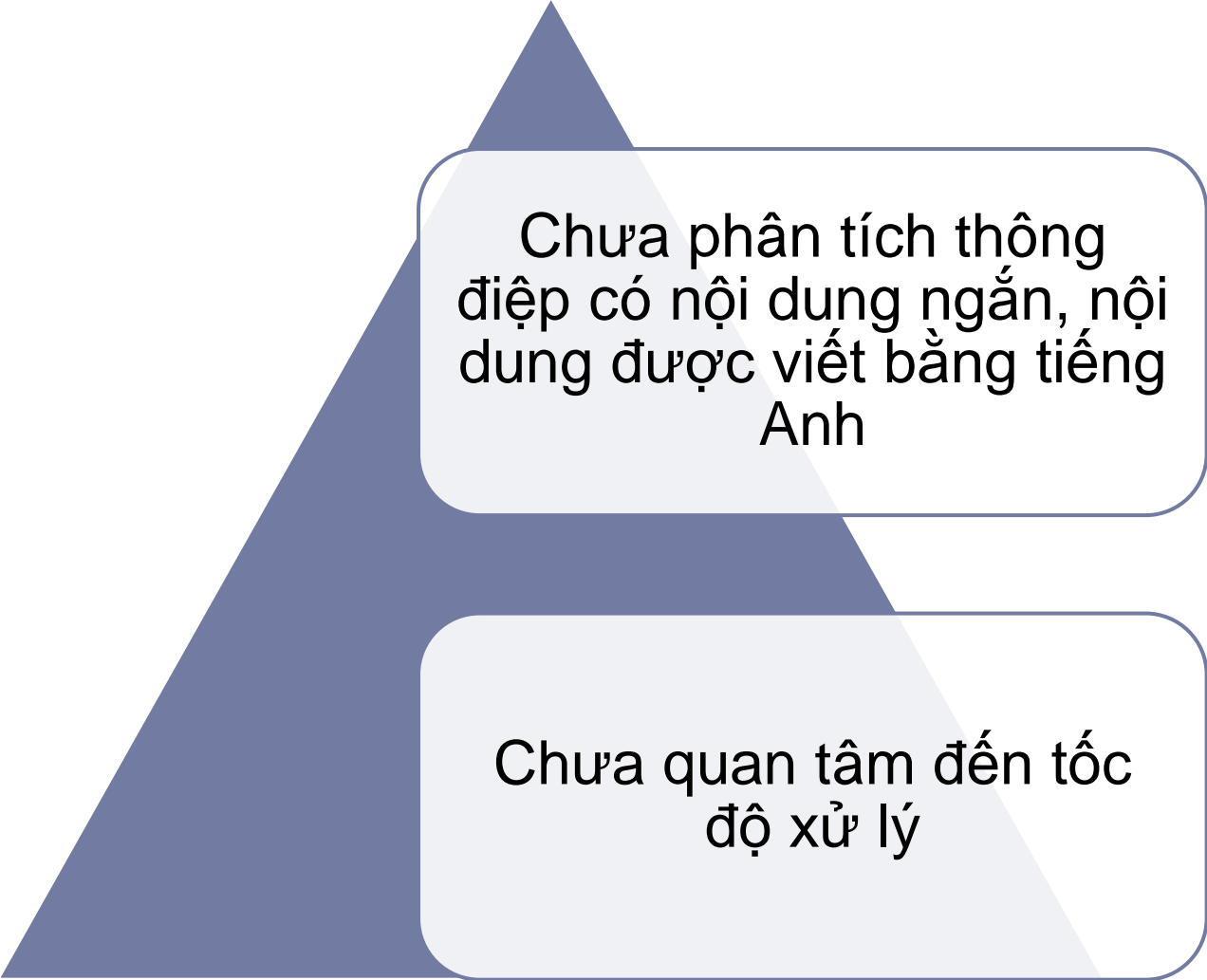


# Tổng quan nghiên cứu của luận án – Thách thức đặt ra



# Tổng quan nghiên cứu của luận án - Hạn chế trong nghiên cứu

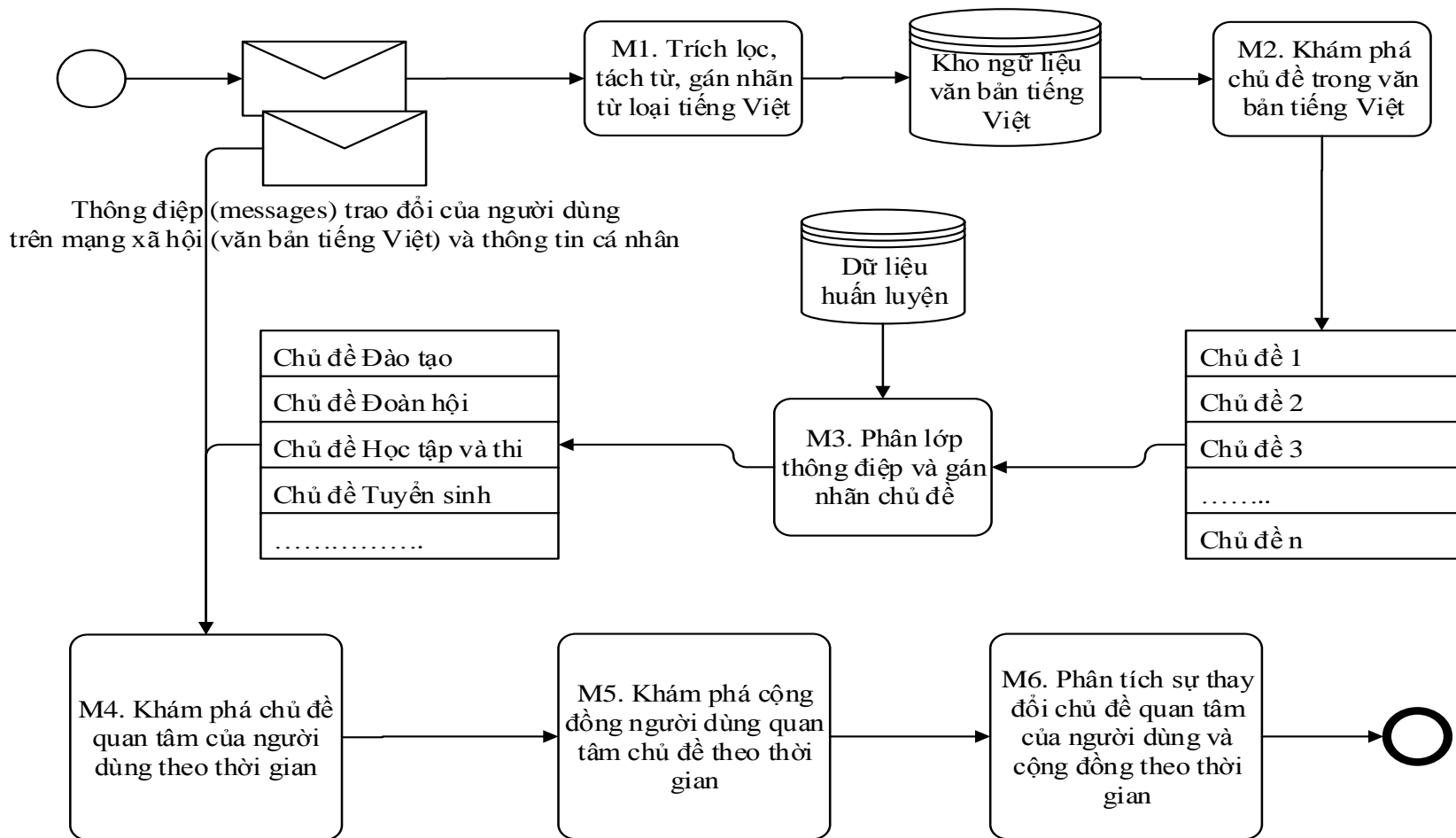
---



Chưa phân tích thông điệp có nội dung ngắn, nội dung được viết bằng tiếng Anh

Chưa quan tâm đến tốc độ xử lý

# Mô hình tổng thể thực hiện nghiên cứu



# Phương pháp thực hiện – 3 bài toán chính

---

1. Khám phá chủ đề từ thông điệp trao đổi trên MXH và gán nhãn chủ đề

2. Khám phá chủ đề quan tâm của người dùng có yếu tố thời gian

3. Khám phá cộng đồng người dùng theo chủ đề có yếu tố thời gian

# 1. Khám phá chủ đề và gán nhãn chủ đề - Phát biểu bài toán 1

---

Chủ đề tiềm ẩn trong thông điệp được trao đổi. Hiểu được chủ đề sẽ hiểu được nội dung trao đổi của thông điệp?



Khảo sát mô hình chủ đề, các phương pháp phân tích mạng xã hội dựa theo mô hình chủ đề



Khám phá chủ đề tiềm ẩn và gán nhãn chủ đề để “hiểu” người dùng trao đổi chủ đề gì trên mạng xã hội



Đề xuất mô hình kết hợp khám phá chủ đề, phân lớp văn bản và gán nhãn chủ đề

# 1. Khám phá chủ đề và gán nhãn chủ đề - Các nghiên cứu liên quan – Khám phá chủ đề

---

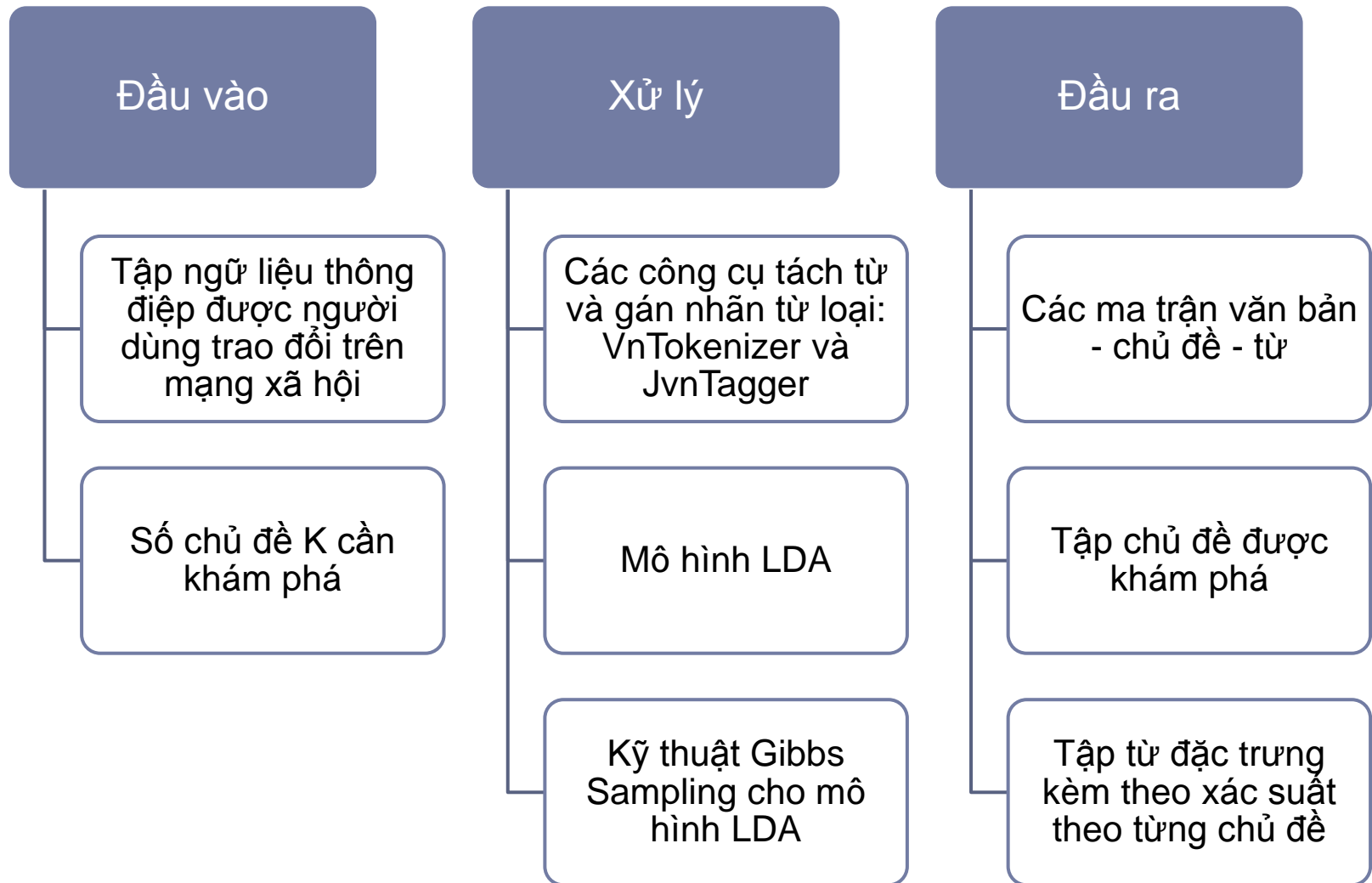
Mô hình LSI -  
Latent Semantic  
Indexing  
(Deerwester et al,  
1990)

Mô hình PLSI -  
Probabilistic  
Latent Semantic  
Indexing (Thomas  
Hofmann, 1999)

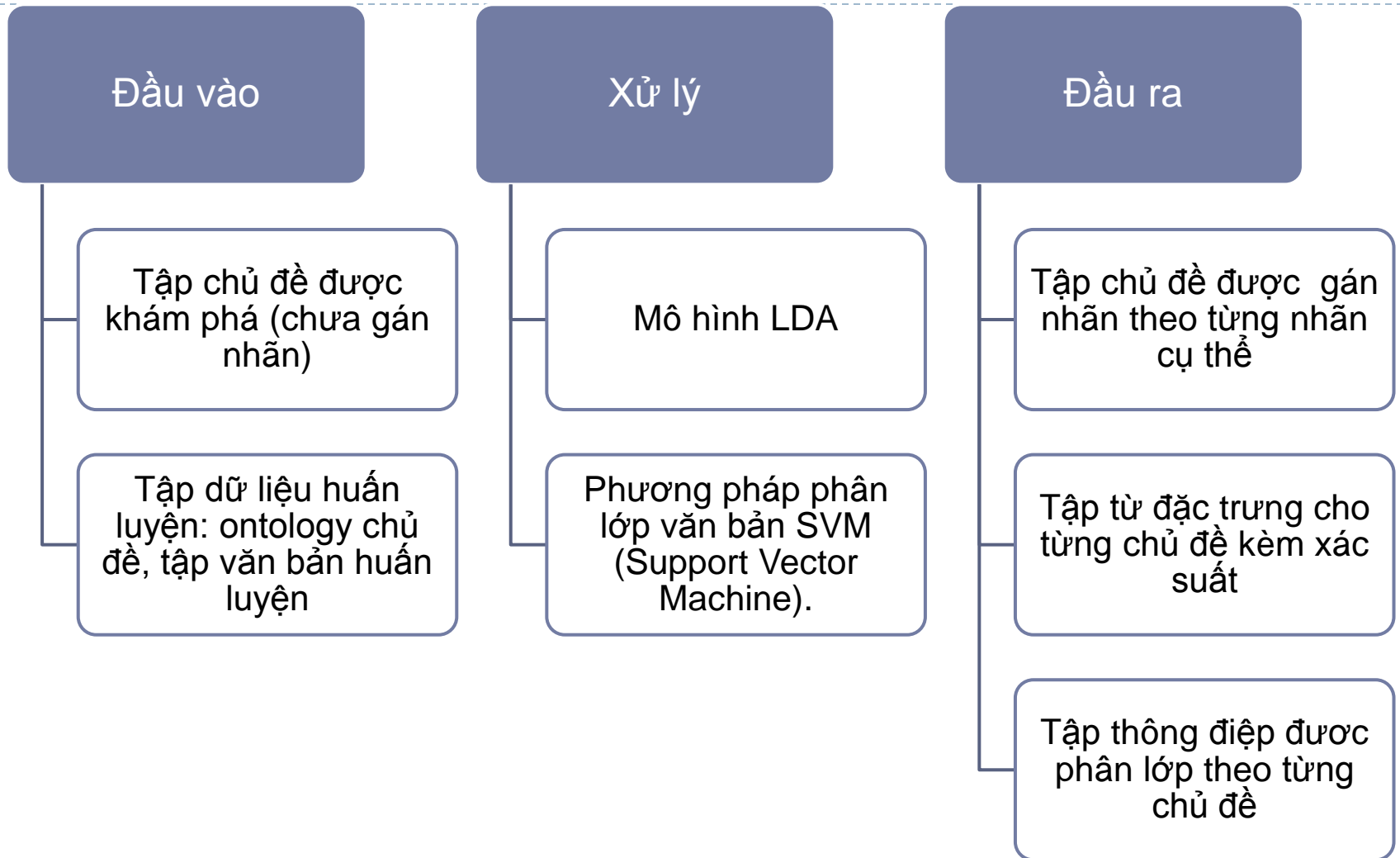
Mô hình LDA -  
Latent Dirichlet  
Allocation (Blei et  
al, 2003)



# 1. Khám phá chủ đề và gán nhãn chủ đề - Phát biểu bài toán (Khám phá chủ đề)



# 1. Khám phá chủ đề và gán nhãn chủ đề - Phát biểu bài toán (Gán nhãn chủ đề)



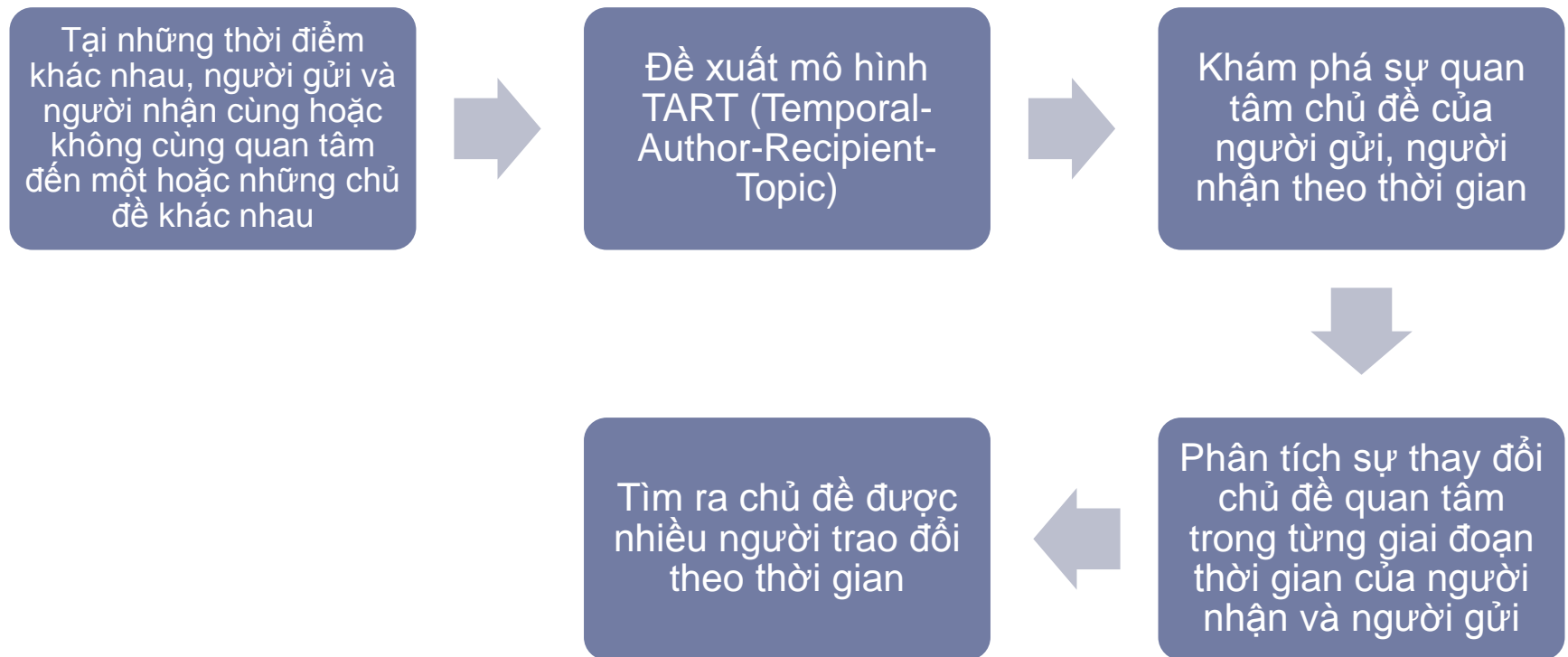
# 1. Khám phá chủ đề và gán nhãn chủ đề - Công bố

---

- [1] **Hồ Trung Thành, Đỗ Phúc** (2014), *Ontology tiếng Việt trong lĩnh vực giáo dục đại học*, Tạp chí Khoa học Công nghệ - Viện Hàn lâm Khoa học Công nghệ Việt Nam, Tập 52, số 1B, pp. 89-100, ISSN: 0866-708x.
- [2] **Hồ Trung Thành, Đỗ Phúc** (2014), *Mô hình tích hợp khám phá và gán nhãn chủ đề tiếp cận theo mô hình chủ đề*, Tạp chí Khoa học Công nghệ ĐHQG-HCM, số K4, tập 17, ISSN: 1859-0128.
- [3] **Muon Nguyen, Thanh Ho, Phuc Do** (2013), *Social Networks Analysis Based on Topic Modeling*, The 10th IEEE RIVF International Conference on Computing and Communication Technologies, Hanoi, pp. 119-123, ISBN: 978-1-4799-1350-3.

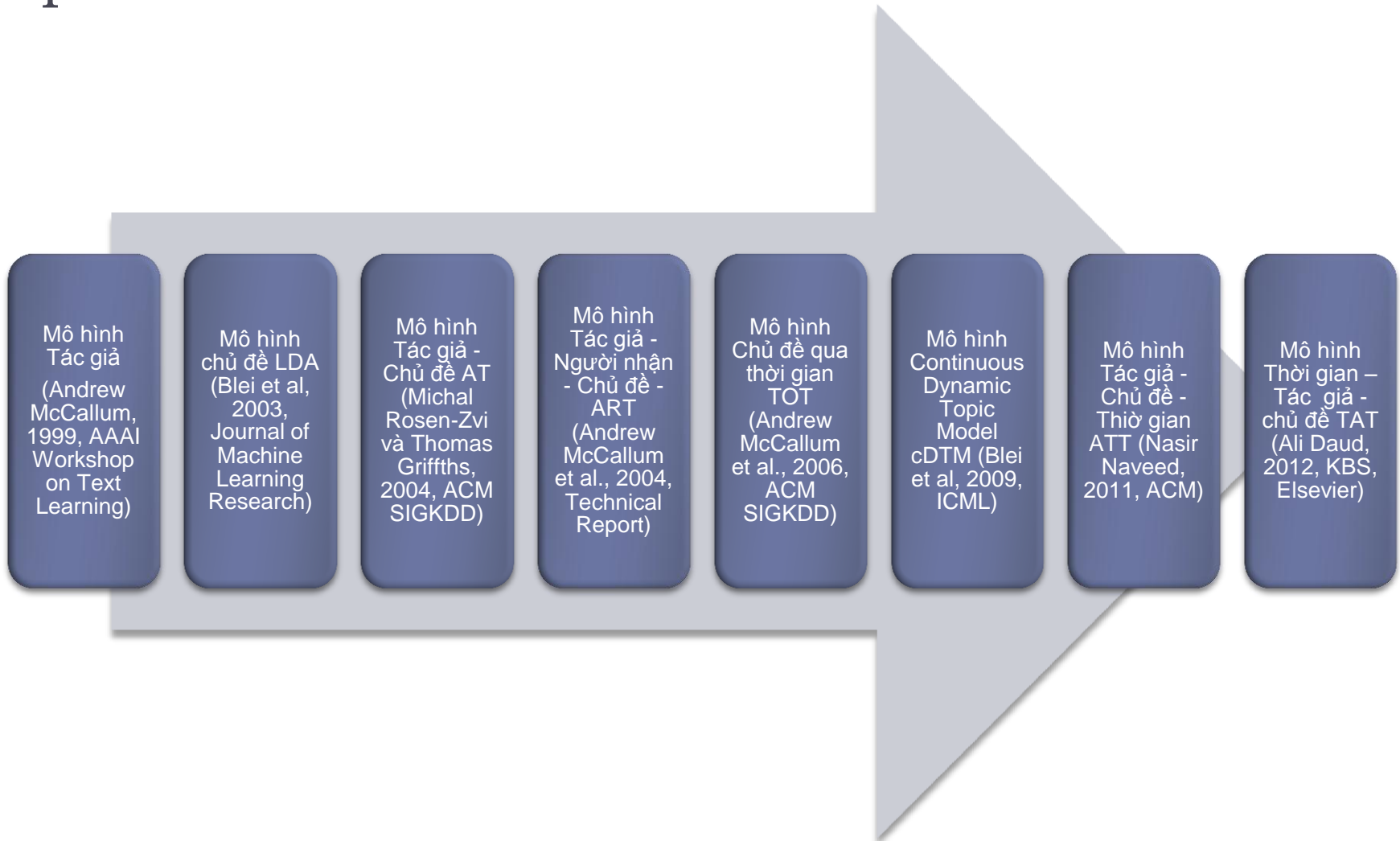
## 2. Khám phá chủ đề quan tâm của người dùng có yếu tố thời gian – Phát biểu bài toán 2

---

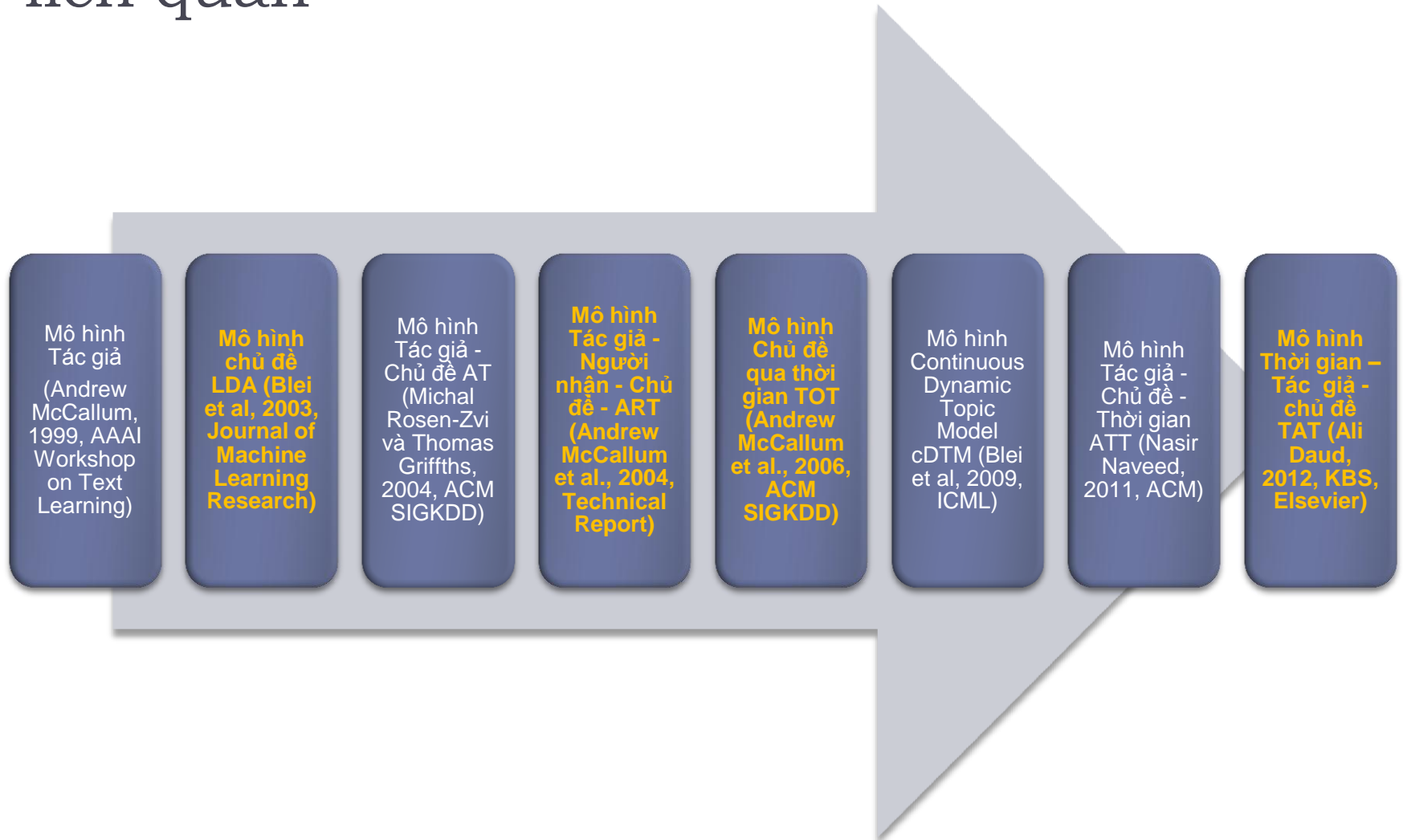


## 2. Khám phá chủ đề quan tâm của người dùng có yếu tố thời gian - Các nghiên cứu liên quan

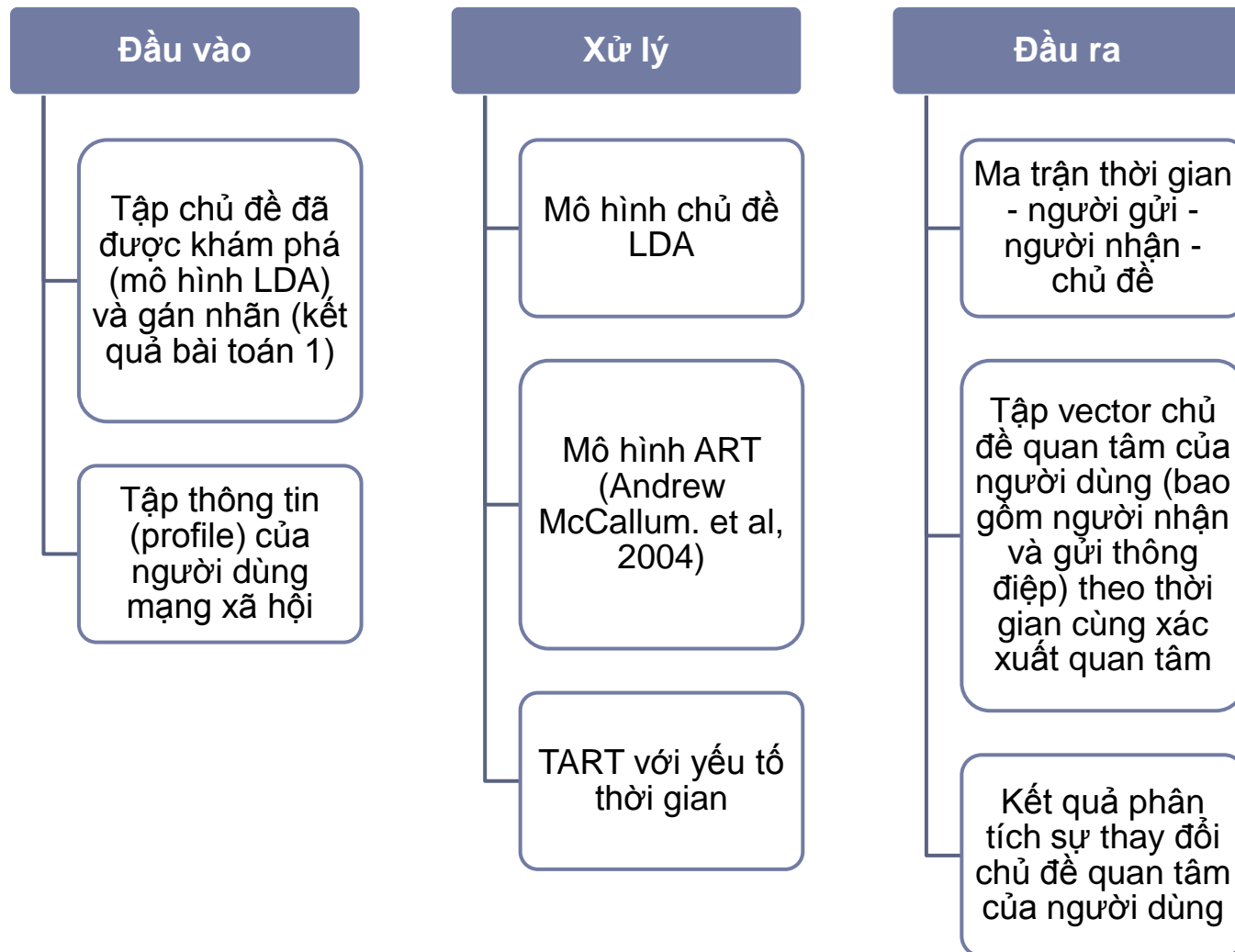
---



## 2. Khám phá chủ đề quan tâm của người dùng có yếu tố thời gian - Các nghiên cứu liên quan

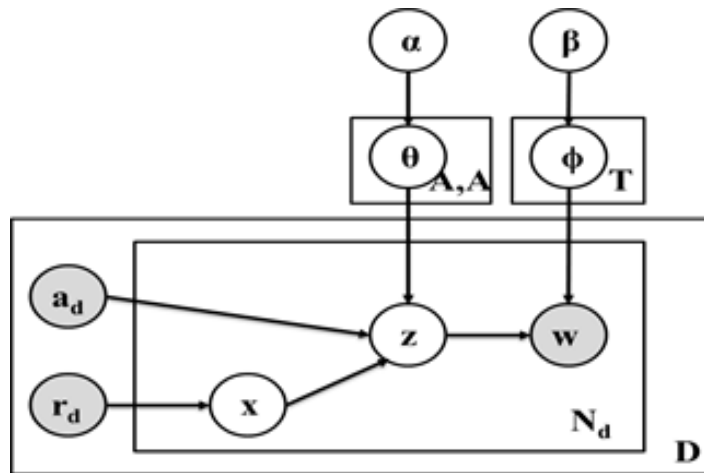


## 2. Khám phá chủ đề quan tâm của người dùng có yếu tố thời gian – Phát biểu bài toán

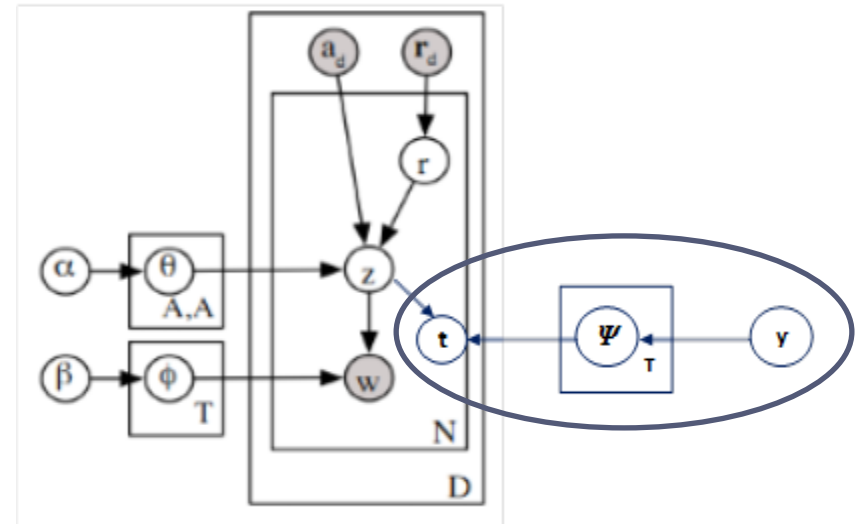


## 2. Khám phá chủ đề quan tâm của người dùng có yếu tố thời gian – Mô hình đề xuất

- ▶ Mô hình ART (Author – Recipient – Topic)



- ▶ Mô hình TART (Temporal - Author – Recipient – Topic)





## 2. Phân tích chủ đề quan tâm của người dùng có yếu tố thời gian – Công bố

---

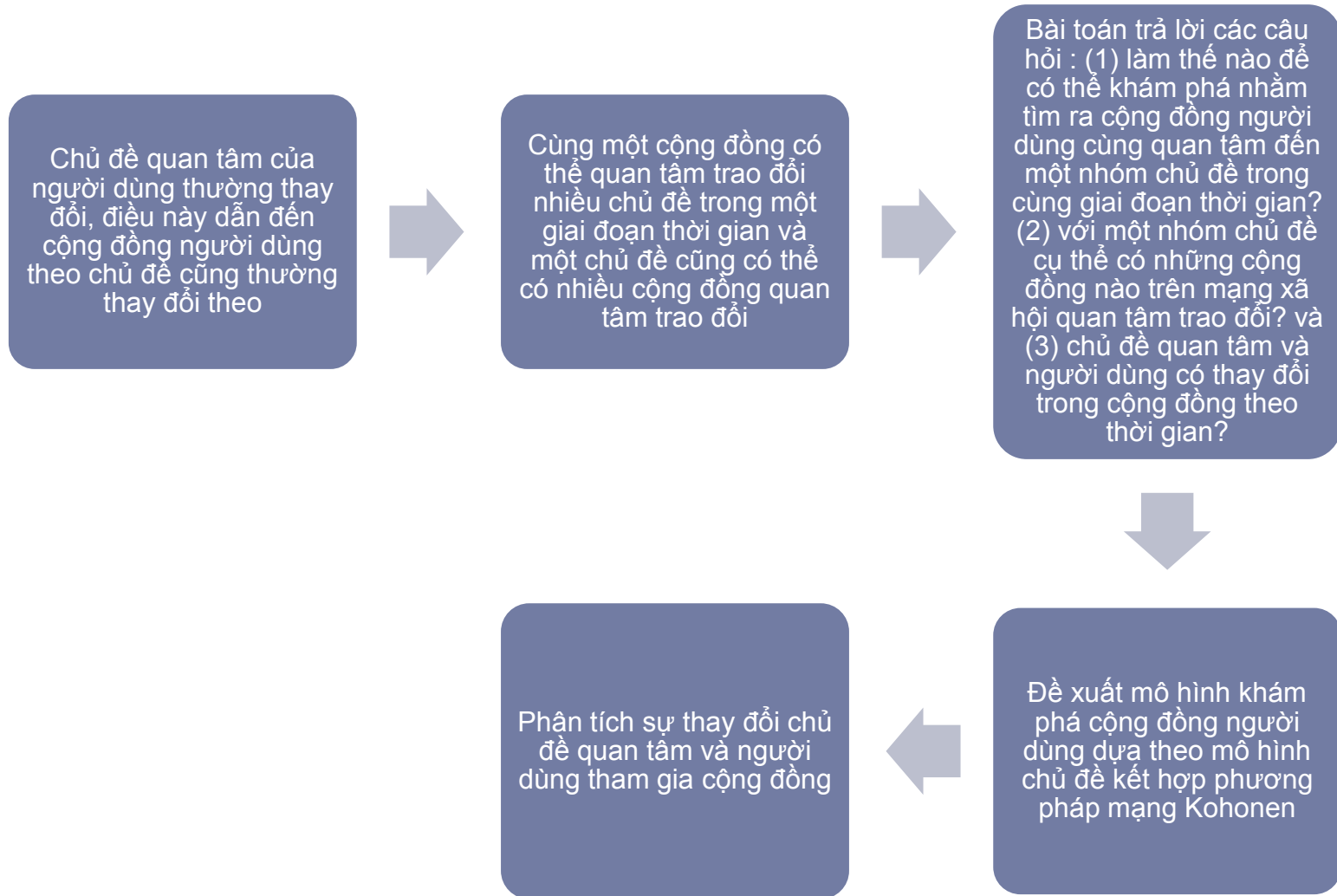
[1] **Thanh Ho**, Phuc Do (2014), *Analyzing Users' Interests with the Temporal Factor Based on Topic Modeling*, ACIIDS 03-2015, Indonesia, Springer, pp. 106-115, ISSN: 0302-9743, ISBN: 978-3-319-15704-7.

[2] **Thanh Ho**, Duy Doan, Phuc Do (2014), *Discovering Hot Topics On Social Network Based On Improving The Aging Theory*, Advances in Computer Science : an International Journal. Volume 3, Issue 3, p. 48-53, ISSN: 2322-5157.

[3] Phan Hồ Viết Trường, **Hồ Trung Thành**, Đỗ Phúc (2013), *Phân tích tầm ảnh hưởng đối tượng theo chủ đề trong mạng xã hội*, Tạp chí Khoa học Công nghệ, Viện Hàn lâm Khoa học Công nghệ Việt Nam, tập 52, số 1B, pp. 101-111, ISSN: 0866-708x.

[4] Nghe Nguyen, **Thanh Ho** and Phuc Do (2015), *Finding the Most Influential User of a Specific Topic on the Social Networks*, Advances in Computer Science : an International Journal, Volume 4, Issue 2, ISSN: 2322-5157.

### 3. Khám phá cộng đồng người dùng theo chủ đề có yếu tố thời gian – Mục tiêu bài toán 3



### 3. Khám phá cộng đồng người dùng theo chủ đề có yếu tố thời gian – Các nghiên cứu liên quan

---

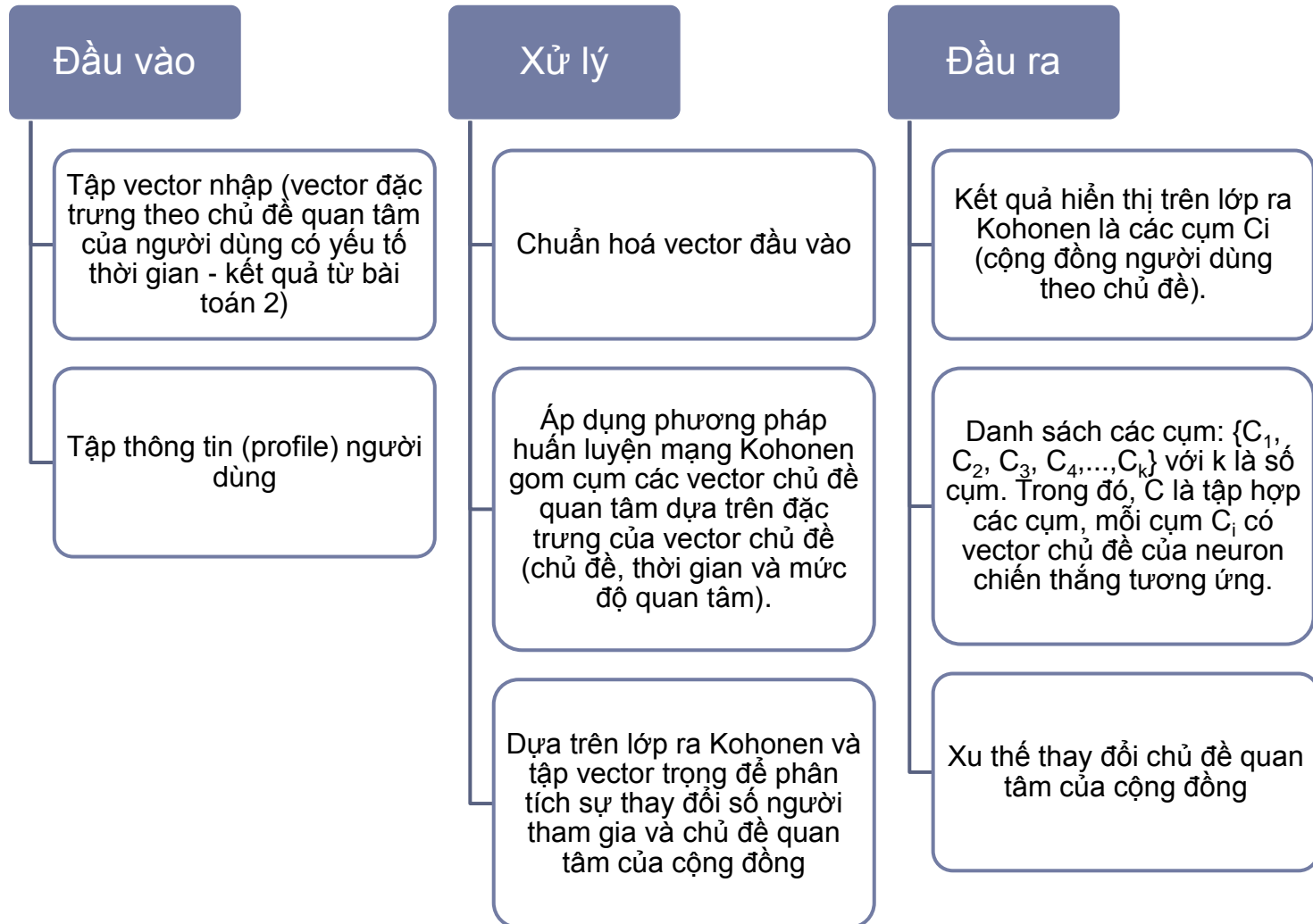
Mô hình Nhóm -  
Chủ đề - GT  
(Andrew McCallum,  
2006, Advances in  
Neural Information  
Processing  
Systems 18)

Mô hình Cộng đồng  
- Người dùng - Chủ  
đề - CUT (D. Zhou,  
et al, WWW)

Mô hình Cộng đồng  
- Người gửi -  
Người nhận - Chủ  
đề -CART (N.  
Pathak, 2008,  
SNA-KDD)

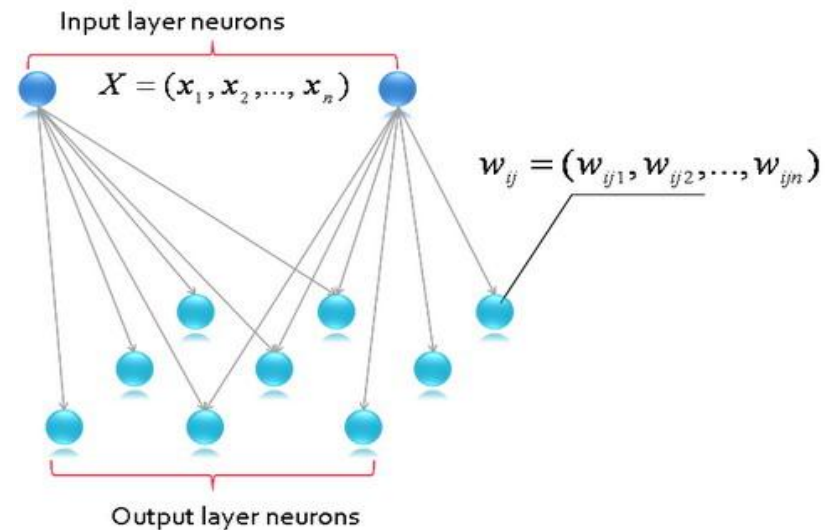
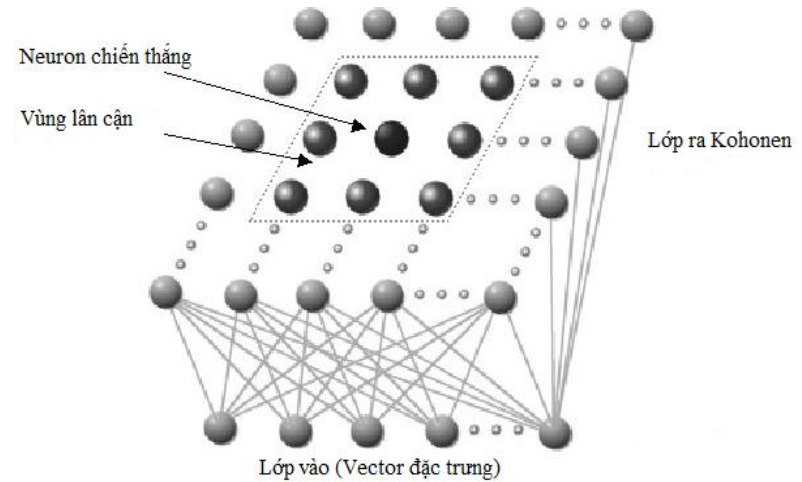
Mô hình Tác giả -  
Chủ đề - Cộng  
đồng - ATC  
(Chunshan Li,  
2014, Springer-  
Verlag London)

### 3. Khám phá cộng đồng người dùng theo chủ đề có yếu tố thời gian – Phát biểu bài toán

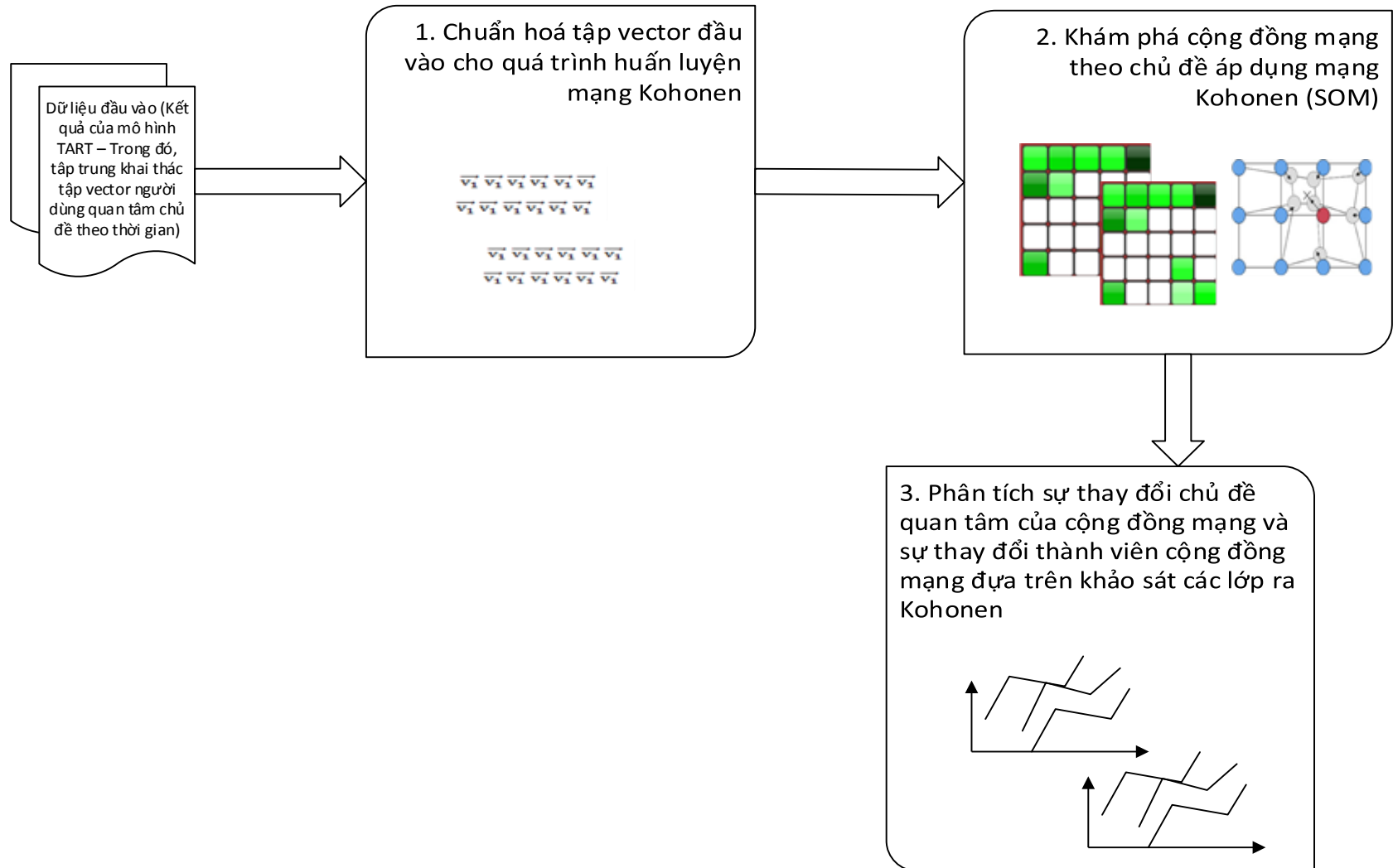


# 3. Phương pháp mạng Kohonen - SOM

- ▶ Có khả năng ánh xạ dữ liệu đầu vào có số chiều lớn thành mạng có số chiều thấp hơn (thường là 2 chiều)
- ▶ Mạng Kohonen có thể gom cụm dữ liệu mà không cần chỉ định trước số cụm
- ▶ Kết quả gom cụm được hiển thị trực quan trên lớp ra của Kohonen



# 3. Khám phá cộng đồng người dùng theo chủ đề có yếu tố thời gian – Mô hình đề xuất



### 3. Khám phá cộng đồng người dùng theo chủ đề có yếu tố thời gian – Công bố

---

[1] **Thanh Ho** and Phuc Do, Analyzing the Changes in Online Community based on Topic Model and Self-Organizing Map, International Journal of Advanced Computer Science and Applications(IJACSA), 6(7), 2015.

[2] **Thanh Ho**, Phuc Do (2015), Discovering Communities of Users on Social Networks Based on the Topic Model Combined with Kohonen Network, KSE, 10/2015, IEEE, Accepted.

[3] Tran Quang Hoa, Vo Ho Tien Hung, Nguyen Le Hoang, **Ho Trung Thanh**, Do Phuc (2014), *Finding the Cluster of Actors in Social Network based on the Topic of Messages*, ACIIDS 04-2014, Thailand, Springer, pp. 183-190, ISBN: 983-3-319-054756-6.

# Tài liệu tham khảo - 1

---

1. Lars Kirchhoff (2010). *Applying Social Network Analysis to Information Retrieval on the World Wide Web: A Case Study of Academic Publication Space*, The University of St. Gallen, Switzerland.
2. Stanley Wasserman, Katherine Faust (1994), *Social Network Analysis: [Methods and Applications](#)*, Cambridge University Press, Nov 25, 1994.
3. Lise Getoor, Christopher P. Diehl. (2005), *Link Mining: A Survey*, SIGKDD Explorations, 7(2), pp. 3-12.
4. Charu C. Aggarwal (2011), *Book: Social Network Data Analytics*, IBM Thomas J. Watson Research Center, Springer.
5. Chong Wang, David Blei and David Heckerman (2009), *Continuous Time Dynamic Topic Models*, Proceedings of ICML. ICML '08.
6. D. Kim, P. Gopalan, D. Blei, and E. Sudderth (2013), *Efficient online inference for Bayesian nonparametric relational models*, Neural Information Processing Systems.
7. Pei Lee, Laks V.S. Lakshmanan, Evangelos Miliotis (2014), *CAST: A Context-Aware Story-Teller for Streaming Social Content*, CIKM'14, November 3–7, 2014, ACM, <http://dx.doi.org/10.1145/2661829.2661859>.
8. Durgesh M. Sharma, Moiz M. Baig (2015), *Sentiment Analysis on Social Networking: A Literature Review*, International Journal on IJRITCC, Volume: 3 Issue: 2, pp. 022-027.
9. David M. Blei, Andrew Y. Ng and Michael I. Jordan (2003), *Latent Dirichlet Allocation*, Journal of Machine Learning Research, pp. 993 – 1022.
10. Tom Griffiths (2004), *Gibbs Sampling in the generative model of Latent Dirichlet Allocation*, Gruffydd@psych.stanford.edu.
11. Hassan Abbas Abdelbary, Abeer Mohamed ElKorany, Reem Bahgat (2014), *Utilizing Deep Learning for Content-based Community Detection*, Science and Information Conference, UK, IEEE, pp. 777-784.
12. Andrew McCallum, Andrés Corrada, Xuerui Wang (2004), *The Author-Recipient-Topic Model for Topic and Role Discovery in Social Networks: Experiments with Enron and Academic Email*, Technical Report UM-CS-2004-096, pp. 44.



# Tài liệu tham khảo - 2

---

13. Paola Velardi, Roberto Navigli, Alessandro Cucchiarelli, Fulvio D'Antonio (2008), *A New Content-Based Model for Social Network Analysis*, IEEE Sixth International Conference on Semantic Computing, pp: 18-25.
14. Angela Bohn, Ingo Feinerer, Kurt Hornik and Patrick Mair (2011), *Content-Based Social Network Analysis of Mailing Lists*, The R Journal. 2011;3(1),pp: 11-18.
15. Xuan-Hieu Phan, Cam-Tu Nguyen, Dieu-Thu Le, Le-Minh Nguyen, Susumu Horiguchi, and Quang-Thuy Ha (2011), *A hidden topic-based framework towards building applications with short Web documents*, IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE), Vol.23, No.7, pp: 961-976.
16. Thorsten Joachims (1999), *Transductive Inference for Text Classification using Support Vector Machines*, International Conference on Machine Learning (ICML), pp. 200-209.
17. Thorsten Joachims (2003), *Transductive learning via spectral graph partitioning*, Proceeding of The Twentieth International Conference on Machine Learning (ICML2003), pp. 290-297.
18. Mccallum, A. (1999), *Multi-label text classification with a mixture model trained by EM*, In AAAI Workshop on Text Learning, pp. 681-687.
19. Nasir Naveed, Sergej Sizov, Steffen Staab (2011), *ATT: Analyzing Temporal Dynamics of Topics and Authors in Social Media*, ACM, WebSci'11, June 14-17, pp. 1-7.
20. D. Zhou, E. Manavoglu, J. Li, C.L. Giles, and H. Zha (2006), *Probabilistic models for discovering e-communities*, In WWW '06: Proceedings of the 15th international conference on World Wide Web, ACM, pp. 173-182.
21. William M. Darling (2011), *A Theoretical and Practical Implementation Tutorial on Topic Modeling and Gibbs Sampling*, In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, pp. 642-647.
22. Jey Han Lau, David Newman, Sarvnaz Karimi (2010), *Best Topic Word Selection for Topic Labelling*, Proceedings of the 23rd International Conference on Computational Linguistics: Posters, ACM, pp. 605-613.
23. Jey Han Lau (2011), *Automatic Labelling of Topic Models*, In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1, Association for Computational Linguistics, 2011, pp. 1536-1545.
24. István Bíró, Jácint Szabó (2008), *Latent Dirichlet Allocation for Automatic Document Categorization*, Research Institute of the Hungarian Academy of Sciences Budapest, pp. 430-441.

# Tài liệu tham khảo - 3

---

25. K. Nowicki and T.A.B. Snijders (2001), *Estimation and prediction for stochastic blockstructures*, Journal of the American Statistical Association, 96(455), pp. 1077–1087.
26. Angela Bohn, Ingo Feinerer, Kurt Hornik and Patrick Mair (2011), *Content-Based Social Network Analysis of Mailing Lists*, The R Journal Vol. 3/1, June 2011, pp. 11-18.
27. Thanh Ho, Duy Doan, Phuc Do (2014), *Discovering Hot Topics On Social Network Based On Improving The Aging Theory*, Advances in Computer Science : an International Journal. Volume 3, Issue 3, p. 48-53.
28. T. L. Griffiths and M. Steyvers (2004), *Finding scientific topics*, PNAS, 1:5228-35.
29. Hồ Trung Thành, Đỗ Phúc (2014), *Mô Hình Tích Hợp Khám Phá Và Gán Nhãn Chủ Đề Tiếp Cận Theo Mô Hình Chủ Đề*, Tạp chí Khoa học Công nghệ ĐHQG-HCM, số K4, tập 17, ISSN: 1859-0128.
30. Muon Nguyen, Thanh Ho, Phuc Do (2013), *Social Networks Analysis Based on Topic Modeling*, The 10th IEEE RIVF International Conference on Computing and Communication Technologies, Hanoi, pp.119-123, ISBN: 978-1-4799-1350-3.
31. Xuerui Wang, Andrew McCallum (2006), *Topics over Time: A Non-Markov Continuous-Time Model of Topical Trends*, ACM SIGKDD-2006, pp. 424–433.
32. David M. Blei, John D. Lafferty (2006), *Dynamic Topic Models*, Appearing in Proceedings of the 23<sup>rd</sup> International Conference on Machine Learning, Pittsburgh, PA, pp. 113-120.
33. Michal Rosen-Zvi, Thomas Griffiths et. al (2004), *Probabilistic AuthorTopic Models for Information Discovery*, 10th ACM SigKDD, Seattle, pp. 306-315.
34. Scott Deerwester, Susan T. Dumais, George W. Furnas, Thomas K. Landauer, Richard Harshman (1990), *Indexing by latent semantic analysis*, Journal American Society for Information Science, 41 (6), pp. 391–407.
35. Hofmann, Thomas (1999), *Probabilistic Latent Semantic Indexing*, Proceedings of the Twenty-Second Annual International SIGIR Conference on Research and Development in Information Retrieval, pp. 50-57.

# Tài liệu tham khảo - 4

---

36. Blei, David M. (2012), *Introduction to Probabilistic Topic Models*, Comm. ACM **55** (4), pp. 77–84.
37. Youngchul Cha, Junghoo Cho (2012), *Social-network analysis using topic models*, The 35th International ACM SIGIR conference, pp. 565-574.
38. Hồ Trung Thành, Đỗ Phúc (2014), *Ontology tiếng Việt trong lĩnh vực giáo dục đại học*, Tạp chí Khoa học Công nghệ - Viện Hàn lâm Khoa học Công nghệ Việt Nam. Tập 52, số 1B, pp. 89-100.
39. X. Wang, N. Mohanty, and A. McCallum (2006), *Group and topic discovery from relations and their attributes*, Advances in Neural Information Processing Systems 18, pp. 1449-1456.
40. N. Pathak, C. DeLong, A. Banerjee, and K. Erickson (2008), *Social topic models for community extraction*, In The 2nd SNA-KDD Workshop, volume 8.
41. Alexandru Berlea<sup>1</sup>, Markus Döhring, Nicolai Reuschling (2009), *Content and communication based sub-community detection using probabilistic topic models*, IADIS International Conference Intelligent Systems and Agents, ISBN: 978-972-8924-87-4 © 2009 IADIS.
42. Wenjun Zhou, Hongxia Jin, Yan Liu (2012), *Community Discovery and Profiling with Social Messages*, KDD'12, August 12–16, 2012, Beijing, China, pp. 388-396.
43. Chunshan Li, William K. Cheung, Yunming Ye, Xiaofeng Zhang, Dianhui Chu, Xin Li (2014), *The Author-Topic-Community model for author interest profiling and community discovery*, Springer-Verlag London 2014, pp. 74-85.
44. The Anh Dang, Emmanuel Viennet (2012), *Community Detection based on Structural and Attribute Similarities*, ICDS 2012 : The Sixth International Conference on Digital Society, pp. 7-14.
45. Yang Zhou, Hong Cheng, Jeffrey Xu Yu (2009), *Graph Clustering Based on Structural/Attribute Similarities*, VLDB '09, August 24-28, 2009, Lyon, France, pp. 718-729.
46. Amer Aljarallah, Kunpeng Zhang (2014), *Using Coauthor Networks to Extract Topics in Information Systems*, Thirty Fifth International Conference on Information Systems, Auckland 2014, pp. 1-9.
47. Do Phuc, Mai Xuan Hung (2008), *Using SOM based Graph Clustering for Extracting Main Ideas from Documents*, RVIF 2008, pp. 209-214.
48. Kohonen T. and Honkela T. (2007), Kohonen network, [http://www.scholarpedia.org/article/Kohonen\\_network](http://www.scholarpedia.org/article/Kohonen_network).

# Tài liệu tham khảo - 5

---

49. Zhijun Yin et. al (2012), *Latent community Topic Analysis: Integration of Community Discovery with Topic Modeling*, ACM Transactions on Intelligent Systems and Technology, pp. 1-21.
50. Kaski, S., Honkela, T., Lagus, K., and Kohonen. T, *WEBSOM--self-organizing maps of document collections*, Neurocomputing, volume 21, (1998), pp. 101-117.
51. Thanh Ho, Phuc Do (2015), *Analyzing Users' Interests with the Temporal Factor Based on Topic Modeling*, 23-25 March 2015, Indonesia, Springer, pp. 106-115.
52. Teuvo Kohonen (1982), *Self-Organized Formation of Topologically Correct Feature Maps*, Biol. Cybern. 43, Springer-Verlag, pp. 59-69.
53. S. Haykin (1999), *Neural Networks. A Comprehensive Foundation*, Second Edition, Prentice-Hall, Inc., New Jersey, 1999, pp.443-465.
54. Kohonen, T. and Somervuo, P. (2002), *How to make large self-organizing maps for nonvectorial data*, Neural Networks 15(8-9), pp. 945-952.
55. Tianbao Yang, Yun Chi, Shenghuo Zhu, Yihong Gong, Rong Jin (2011), *Detecting communities and their evolutions in dynamic social networks—a Bayesian approach*, Mach Learn 82, Springer, pp. 157–189.
56. Ding Zhou, Isaac Councill, Hongyuan Zha, C. Lee Giles (2007), *Discovering Temporal Communities from Social Network Documents*, IEEE ICDM, pp. 745-750.
57. Tran Quang Hoa, Vo Ho Tien Hung, Nguyen Le Hoang, Ho Trung Thanh, Do Phuc (2014), *Finding the Cluster of Actors in Social Network based on the Topic of Messages*, ACIIDS 04/2014, ThaiLan. Springer, pp. 183-190.
58. Thanh Ho and Phuc Do (2015), *Analyzing the Changes in Online Community based on Topic Model and Self-Organizing Map*, International Journal of Advanced Computer Science and Applications (IJACSA), 6(7), pp.
59. Yan Liu, Alexandru N.M et al (2009), *Topic-Link LDA: Joint Models of Topic and Author Community*, Proceedings of the 26 th International Conference on Machine Learning, ACM, pp. 665-672.
60. Mr inmaya Sachan, et al (2012), *Using Content and Interactions for Discovering Communities in Social Networks*, International World Wide Web Conference Com-mittee (IW3C2), Lyon, France, pp. 331-340.

# Tài liệu tham khảo - 6

---

61. Nguyen Le Hoang, Do Phuc, et al (2013), *Predicting Preferred Topics of Authors based on Co-Authorship Network*, The 10th IEEE RIVF International Conference on Computing and Communication Technologies, IEEE, pp. 70-75.
62. Amjad Abu-Jbara, Ahmed Hassan, Dragomir Radev (2012), *AttitudeMiner: Mining Attitude from Online Discussions*, Proceedings of the NAACL-HLT 2012: Demonstration Session, pp. 33–36.
63. [M. Kharratzadeh](#), [B. Renard](#), [M.J. Coates](#) (2015), *Bayesian topic model approaches to online and time-dependent clustering*, Journal: [Digital Signal Processing](#), Elsevier, <http://dx.doi.org/10.1016/j.dsp.2015.03.010>.
64. Nghe Nguyen, Thanh Ho and Phuc Do (2015), *Finding the Most Influential User of a Specific Topic on the Social Networks*, Advances in Computer Science : an International Journal, Volume 4, Issue 2.
65. Ali Daud, Juanzi Li et al (2009), *Exploiting Temporal Authors Interests via Temporal-Author-Topic Modeling*, ADMA 2009, Springer-Verlag Berlin Heidelberg 2009, LNAI 5678, pp. 435–443.
66. Ali Daud (2012), *Using time topic modeling for semantics-based dynamic research interest finding*, Knowledge-Based Systems 26 (2012), Elsevier, pp. 154–163.
67. Tom Fawcett (2005), *Introduction to ROC Analysis*, Elsevier B.V., Available online [www.sciencedirect.com](http://www.sciencedirect.com).
68. H.-L. Yang and J.H. Tang (2003), *Effects of social network on students' performance*, A web-based forum study in Taiwan, Journal of Asynchronous Learning Networks, vol. 7, no. 3, pp. 93-197.
69. A. Calvó-Armengol, E. Patacchini, and Y. Zenou (2009), *Peer Effects and Social Networks in Education*, Review of Economic Studies, vol. 76, no. 4, pp. 1239-2167.
70. A. Mora-Soto, M. Sanchez-Segura, F. Medina-Dominguez, and A. Amescua (2009), *Collaborative Learning Experiences Using Social Networks*, in *International Conference on Education and New Learning Technologies*, EDULEARN09, Barcelona, Spain, pp. 4260-4270.
71. Thanh Ho, Phuc Do (2015), *Discovering Communities of Users on Social Networks Based on the Topic Model Combined with Kohonen Network*, KSE 10/2015, UIT, Vietnam, Accepted 07/2015.

---

**TRÂN TRỌNG CẢM ƠN  
QUÍ THẦY/CÔ**