ITAIS 2021 Proceedings

Annual conference of the Italian Chapter of AIS (ITAIS)

2021

A Taxonomy of Academic Abstract Sentence Classification Modelling

Connor Stead

Stephen Smith

Peter Busch

Savanid Vatanasakdakul

A TAXONOMY OF ACADEMIC ABSTRACT SENTENCE CLASSIFICATION MODELLING

Connor Stead[1], Stephen Smith[2], Peter Busch[3] and Savanid Vatanasakdakul[4]

[1] Macquarie University, NSW, Australia
[2] Carnegie Mellon University, Doha, Qatar
connor.stead@hdr.mq.edu.au stephen.smith@mq.edu.au
peter.busch@mq.edu.au savanid@cmu.edu

Abstract.
Background: Abstract sentence classification modelling has the potential to advance literature discovery capability for the array of academic literature information systems, however, no artefact exists that categorises known models and identifies their key characteristics.
Aims: To systematically categorise known abstract sentence classification models and make this knowledge readily available to future researchers and professionals concerned with abstract sentence classification model development and deployment.
Method: An information systems taxonomy development methodology was adopted after a literature review to categorise 23 abstract sentence classification models identified from the literature. Corresponding dimensions and characteristics were derived from this process with the resulting taxonomy presented.
Results: Abstract sentence classification modelling has evolved significantly with state-of-the-art models now leveraging neural networks to achieve high performance sentence classification. The resulting taxonomy provides a novel means to observe the development of this research field and enables to consider how such models can further improved or deployed in real-world applications.

Keywords: abstract sentence classification modelling, taxonomy, classification, design science

1 Introduction

As the volume of academic literature continues to grow at an unprecedented rate (Rawat and Meena, 2014, Ware and Mabe, 2015, Khabsa and Giles, 2014), researchers are increasingly relying on information systems to provide them with access to relevant literature. The increasing reliance of researchers on these systems warrants the exploration of how machine/deep learning and artificial intelligence can innovate the researcher human computer interaction with such systems, ultimately leading to an enhancement in the efficiency for a researcher to perform the activity of literature discovery. This innovation may materialise through the deployment of machine/deep learning models capable of classifying academic abstract sentences into literature characteristic classes, such as a paper's purpose, methodology, findings and contributions. These models, emanating from the research field of abstract sentence classification modelling (ASCM), offer a novel means to improve the discovery of literature relevant to research. Searching specifically for literature which utilise a certain methodology or have determined similar findings is difficult given the current functionality of academic literature indexes, such as Emerald Insights, ScienceDirect, PubMed, Google Scholar and Microsoft Academic. The adoption of ASCM into the infrastructure of these platforms may enhance their literature discovery capability offering, thereby improving the ability of researchers to identify meaningful and related literature. Only the information systems discipline can offer the knowledge to bring the idea of deploying ASCM into the complex infrastructure of large-scale platforms into reality, however, the ASCM research domain is complex and no research exists offering a comprehensive outline of the key dimensions and characteristics of developed ASCM models.

2 Information Systems Taxonomy Development

Taxonomies are present in almost all facets of human life. From categorization of the animal kingdom to laws governing our conduct (Joudrey and Taylor, 2009), taxonomies assist us in our journey through the universe. Taxonomies also guide us through the information systems body of knowledge. Nickerson et al. (2013) introduced a taxonomy development approach focused on providing optimal benefit to information systems research, deemed necessary given the adoption of taxonomies in a range of other disciplines (Alter, 1977: decision support systems, Eldredge and Cracraft, 1980: biology, Bailey, 1994: social sciences and Kennedy-Eden and Gretzel, 2012: mobile applications in tourism). The taxonomy development process (Nickerson et al., 2013) is grounded in design science, a research paradigm concerned with building and evaluating artefacts (Recker, 2013). Nickerson et al. (2013) derive a specific definition of a taxonomy, being:

> Taxonomy (T) = “a set of n dimensions Di (i=1, … , n) each consisting of ki (ki≥2) mutually exclusive and collectively exhaustive characteristics Cij ¬(j=1, … , ki) such that each object under consideration has one and only one Cij for each Di” (p. 340).

The taxonomy definition can be notated as (Nickerson et al., 2013):

$$T = \{Di, I = 1,\ldots,n|Di = \{Cij,j = 1,\ldots,ki; ki\geq2\}\}$$

Nickerson et al. (2013) do not explicitly define a dimension or characteristic in relation to a taxonomy, however they do refer to similar terms used in cluster analysis literature (Anderberg, 2014). The mutually exclusive component of the definition reflects the need for objects to not have “two different characteristics in a dimension” (Nickerson et al., 2013, p. 341), and the collectively exhaustive criteria requires that each object features a characteristic in each dimension. The purpose of a taxonomy is dependent on the anticipated users of the resulting artefact, a determination which can be made either explicitly or implicitly by the researcher (Nickerson et al., 2013). The usefulness of a taxonomy is dependent on whether it is concise, robust, comprehensive, extendible and explanatory. These factors are known as a taxonomy's key qualitative attributes (Nickerson et al., 2013). Conciseness is achieved by limiting the number of dimensions and the characteristics within each dimension. By ensuring conciseness, the likelihood of the taxonomy to be understood and applied by researchers is greater. (Nickerson et al., 2013). Robustness is achieved by ensuring that a taxonomy contains "enough dimensions and characteristics to clearly differentiate the objects of interest" (p. 341).

Nickerson et al.'s (2013) taxonomy development process is iterative in nature, in that several cycles of dimension and characteristic discovery are expected to occur. Therefore, ending conditions are required, which come in two forms: objective and subjective. The principal objective ending condition is that the resulting taxonomy adheres with the previously stated definition, particularly that it "consists of dimensions each with mutually exclusive and collectively exhaustive characteristics" (Nickerson et al., 2013, p. 343). Additional objective ending conditions include what proportion of the object population has been examined, whether each object is categorized within each characteristic of the taxonomy's dimensions and that no new dimensions or characteristics are added, merged or split in the most recent iteration (Nickerson et al., 2013, p. 344). The minimum subjective ending conditions that should be consulted are reflective of the key qualitative attributes previously discussed, specifically that the taxonomy under development "is concise, robust, comprehensive, extendible, and explanatory" (Nickerson et al., 2013). Nickerson et al. (2013) provide a comprehensive set of objective and subjective ending conditions in their study (p. 343).

3 Abstract Sentence Classification Modelling

ASCM is the development of machine learning models capable of classifying academic abstract sentences into classifications representative of key literature characteristics, such as the explicit section headings observed in structured abstracts (such as 'purpose', 'method', 'findings' etc.). The term 'structured abstracts' is used to describe abstracts featuring explicit headings reflecting key characteristics of a study. The composition of headings utilised is commonly referred to as a format (Nakayama et al., 2005), which can vary significantly, particularly across academic disciplines. An example of a structured abstract format is known as IMRAD, an acronym for 'introduction', 'methods', 'results' and 'discussion'. This format is endorsed by the International Committee of Medical Journal Editors and the Consolidated Standards of Reporting Clinical Trials Group. IMRAD's adoption is evident across the medical academic literature domain (Nakayama et al, 2005). There are a number of other formats used within literature, such as Emerald Group Publishing's headings: 'purpose', 'design/methodology/approach', 'findings' and 'originality/value'. If applicable, the headings 'research limitations', 'practical limitations' and 'social implications' are also included (Emerald Group Publishing, 2019).

There are several examples from the literature of how machine/deep learning and natural language processing methods have been used to classify sentences from academic literature abstracts into structured abstract heading classifications. Gonçalves et al. (2018), for example, developed a "novel deep learning approach based on a convolutional layer and a bidirectional gated recurrent neural network" (p. 479) to classify randomized controlled trial abstract sentences obtained from PubMed. Yamamoto and Yamamoto and Takagi (2005) developed a sentence classification system using support vector models to categorise abstract sentences also obtained from PubMed queries. Chung (2009) used conditional random fields methods to classify randomized controlled trial abstract sentences into four structured abstract components: 'Intervention', 'Participant' and 'Outcome Measures'. Further examples exist in the literature of attempts to develop abstract sentence classification models, however, no artefact exists which can provide a comprehensive gateway to this research domain.

4 Taxonomy Development

This section will detail our efforts to develop a taxonomy of known attempts at ASCM model development. Due to limitations in the length of this paper the taxonomy is available to view at this website:

https://www.abstractsentenceclassification.com/taxonomy.html

Our working definition of ASCM is:

> The leveraging of machine learning, deep learning or other artificial intelligence methods to classify sentences originating from academic abstracts into structured abstract subheadings, such as a study's introduction, method, results or discussion.

There is no artefact available which comprehensively surveys current ASCM models. Therefore, our intended taxonomy may be of use to researchers and professionals seeking to develop novel ASCM models or to deploy these into production platforms. As a result, the intended audience of the taxonomy is researchers and professionals seeking to rapidly understand the current state of ASCM model development. Characteristics captured should reflect the manner in which ASCM models have been developed and how they function to perform classification. Accordingly, models should be differentiated from one another based on how classification is achieved. The taxonomy should also enable researchers to determine the novelty of forthcoming attempts to develop models.

To identify literature that should be included within the scope of our analysis we queried the academic literature indices Google Scholar, Microsoft Academic, Scopus, Emerald Insight, IEEE Xplore and ScienceDirect using search queries such as: "abstract sentence classification modelling", "sequential sentence classification", "abstract sentence machine learning" and "abstract sentence natural language processing". With literature identified we conducted bidirectional citation searching (Hinde and Spackman, 2015) to identify additional literature to retrieve as a result. Some material identified also conducted literature reviews to varying degrees which provided additional support in our literature acquisition. This multi layered approach resulted in the assessment of literature believed to be representative of the research field of ASCM, providing generalisability for our taxonomy development process.

From our literature acquisition efforts, we identified 23 studies concerned with the development and evaluation of ASCM models. These studies are the objects considered in the identification of dimensions and characteristics.

Objects: 23 ASCM models identified from the literature examining abstract sentence classification model development and evaluation

Step 1: Meta-characteristic: The enablement of abstract sentence classification via machine/deep learning or other artificial intelligence means.

The second step in the taxonomy development process comprises the specification of ending conditions. There are two variants of ending conditions: objective and subjective. Nickerson et al. (2013, p. 344) provide examples of objective ending conditions in their study, of which we adopt 5. Nickerson et al. (2013, p. 344) also provide 5 subjective ending conditions, reflecting the key qualitative attributes taxonomies should feature. We implement all these subjective ending conditions in our study.

Step 2: Ending conditions: The taxonomy development process will conclude when both the specified objective and subjective ending conditions have been met:

Objective Ending Conditions
- All objects have been examined
- At least one object is classified under every characteristic of every dimension
- No new dimensions or characteristics were added in the most recent iteration
- No dimensions or characteristics were merged or split in the most recent iteration
- Every dimension is unique

Subjective Ending Conditions
- Concise: Does the number of dimensions allow the taxonomy to be meaningful without being unwieldy or overwhelming?
- Robust: Do the dimensions and characteristics provide for object differentiation?
- Comprehensive: Can all objects or a random sample of objects within the domain of interest be classified? Are all dimensions of the objects of interest identified?
- Extendible: Can a new dimension or characteristic of an existing dimension be easily added?
- Explanatory: What do the dimensions and characteristic explain about an object?

Iteration 1

3: Approach: We begin with the empirical-to-conceptual approach, as we have identified 23 studies from the literature exploring abstract sentence classification modelling.

4e: We randomly select 6 studies to start the first iteration of the taxonomy development process:

Ruch et al. (2007)	Hassanzadeh (2014)	Chung (2009)
Yamamoto and Takagi (2005)	Ito et al. (2004)	Jiang et al. (2019)

5e: The first dimension identified is that of the structured abstract format used as a classification structure. A format is the composition of explicit headings used to structure an abstract (Nakayama et al., 2005). A common example of a structured abstract format is IMRAD, an acronym for 'Introduction', 'Methods', 'Results' and 'Discussion'. For the purpose of developing ASCM, formats serve as the intended classifications. In this iteration we identify 5 formats. We also identify the use of multiple formats, as is the case with Jiang et al. (2019).

6e: These characteristics are grouped informally into the 'Format' dimension. Their count and nature do not warrant subgrouping. Our first taxonomy is structured as per below:

Dimension D_1 = Format

Characteristics

C_{11} = PURPOSE-METHODS-RESULTS-CONCLUSIONS

C_{12} = BACKGROUND-INTERVENTION-OUTCOME-POPULATION-STUDY DESIGN-OTHER

C_{13} = BACKGROUND-PURPOSE-METHOD-RESULT-CONCLUSION

C_{14} = AIM-METHOD-RESULTS-CONCLUSION

C_{15} = BACKGROUND-OBJECTIVE-METHOD-RESULTS-CONCLUSION

C_{16} = Multiple

7e: Ending conditions: One dimension has been added to the taxonomy and additional objects remain to be reviewed. Therefore, the method requires an additional iteration. Subjectively, at this stage the taxonomy is concise, extendible and explanatory, however, it is not robust nor comprehensive, further reasons for a subsequent iteration.

Iteration 2

3: Approach: Due to objects remaining for review, we employ the empirical-to-conceptual approach again in this subsequent iteration.

4e: We randomly select 5 studies not reviewed in the first iteration:

Jin and Szolovits (2018)	Shimbo et al. (2003)	Kim et al. (2011)
Dernoncourt et al. (2016)	Hirohata et al. (2008)	

5e: In this iteration we identify an additional dimension, reflecting the datasets utilised for the training and evaluation phases of the model development process. These datasets are typically associated with academic literature indexes, such as PubMed and Medline. We identified two variants of datasets, those constructed within the scope of the study and those adopted from the literature. We also identified additional characteristics for the ‘Format’ dimension.

6e: The characteristics identified in this iteration are grouped into the ‘Dataset’ dimension. Their nature does not warrant subgrouping, as doing so would result in objects in the taxonomy with no characteristics. The second taxonomy is structured as per below:

Dimension D_2 = Dataset

Characteristics

C_{21} = NICTA-PIBOSO (Kim et al., 2011) and PubMed RCT (Dernoncourt and Lee, 2017)

C_{22} = Computation and Language Archive - Study Developed

C_{23} = MEDLINE - Study Developed

The additional characteristic identified for the ‘Format’ dimension result in an adjustment of the first taxonomy:

Dimension D_1 (Format) Additional Characteristics

C_{17} = BACKGROUND-TOPIC/ABOUTNESS-RELATED WORK-PURPOSE/PROBLEM-SOLUTION/METHOD-RESULT CONCLUSION/CLAIM

e: Ending conditions: Subjectively, the taxonomy remains concise, extendible and explanatory. Its robustness and comprehensiveness are evolving, however, is not yet satisfactory. Furthermore, additional objects require examination. The method requires reiteration.

Iteration 3

3: Approach: As was the case with the last iteration, there are objects remaining for review. Hence, this iteration will take the empirical-to-conceptual approach.

4e: We randomly select 6 studies not reviewed in the previous iterations:

Gonçalves et al. (2018)	Nam et al. (2016)	McKnight and Srinivasan (2003)
Teufel and Moens (1998)	Wu et al. (2006)	Cohan et al. (2019)

5e: In this iteration we identify an additional dimension reflecting the machine learning modelling algorithms used to facilitate sentence classification. Examples of such algorithms are Support Vector Machines (SVM) and Linear Classifier models employed by McKnight and Srinivasan (2003). We also identify 3 additional characteristics for the ‘Format’ dimension, demonstrating at this stage of the study the diverse nature of formats utilised in ASCM. Additional characteristics for the ‘Dataset’ dimension are also identified.

6e: The modelling algorithm related characteristics identified in this iteration are grouped informally into the ‘Modelling Algorithm’ dimension. At this stage, their nature does not necessitate subgrouping, however, this may materialise as the taxonomy process proceeds. The third taxonomy is structured as follows:

Dimension D_3 = Modelling Algorithm
Characteristics
C_{31} = Neural Network
C_{32} = Conditional Random Fields
C_{33} = Support Vector Machines
C_{34} = Support Vector Machines and Linear Classifier
C_{35} = Bespoke - Study Developed
C_{36} = Bidirectional Encoder Representations from Transformers (BERT)
The additional characteristics identified for the 'Format' and 'Training and Evaluation' dimensions result in an adjustment of the taxonomy:
Dimension D_1 (Format) Additional Characteristics
C_{18} = BACKGROUND-POPULATION-INTERVENTION-OUTCOME-STUDY DESIGN-OTHER
C_{19} = INTRODUCTION-METHOD-RESULT-CONCLUSION
C_{110} = INTRODUCTION-METHOD-RESULT-DISCUSSION
Dimension D_2 (Dataset) Additional Characteristics
C_{24} = PubMed – Study Developed
C_{25} = Citeseer – Study Developed
C_{26} = PubMed RCT (Dernoncourt and Lee, 2017)

7e: Ending conditions: At this stage of the taxonomy development process we have reviewed 17 objects (studies), resulting in 6 remaining for review. Therefore, the objective ending condition regarding the examination of all objects have not been met and another iteration is required.

Iteration 4

3: Approach: The objective ending conditions were not met in the previous iteration as 6 objects remain. Consequently, this iteration will adopt an empirical-to-conceptual approach.

4e: We select the 6 studies not reviewed in the previous iterations for examination:

Lin et al. (2006)	Lui (2012)	Verbeke et al. (2012)
Liu et al. (2013)	Gonçalves et al. (2019)	Xu et al. (2006)

5e: In this iteration we determine that studies exploring ASCM may be distinguished by the research domains from which literature selected to comprise the training and evaluation datasets originate. We also identify 2 additional characteristics for the 'Format' dimension, 3 characteristics for the 'Dataset' dimension and 3 characteristics for the 'Modelling Algorithm' dimension.

6e: The research domain related characteristics identified in this iteration are grouped informally into the 'Research Domain' dimension. It is possible that in a future iteration this dimension is reduced to two characteristics: Biomedical and Non-Biomedical, however, at this stage the characteristics will be embedded into the taxonomy as described in the preceding stage.

The fourth taxonomy is structured as follows:

Dimension D_4 = Research Domain
Characteristics
C_{41} = Cross Discipline
C_{42} = Biomedical – Gene Ontology
C_{43} = Biomedical – Randomized Controlled Trials
C_{44} = Biomedical – EBM Generic
The additional characteristics identified for the 'Format', 'Dataset' and 'Modelling Algorithm' dimensions result in an adjustment of the taxonomy:
Dimension D_1 (Format) Additional Characteristics
C_{111} = BACKGROUND-GOAL-METHOD-RESULT
C_{112} = INTRODUCTION-OBJECTIVE-METHOD-RESULT-CONCLUSION
Dimension D_2 (Dataset) Additional Characteristics
C_{27} = ScienceDirect – Study Developed
C_{28} = Origin Unknown – Study Developed
C_{29} = PubMed RCT (Dernoncourt and Lee, 2017) and Study Developed Computer Science
C_{210} = NICTA-PIBOSO (Kim et al., 2011)
Dimension D_3 (Modelling Algorithm) Additional Characteristics
C_{37} = Hidden Markov Model
C_{38} = Logistic Regression
C_{39} = kLog (Frasconi et al., 2014)

7e: Ending conditions: We have reviewed all studies, thereby meeting the first objective ending condition, however, the third objective ending condition is not satisfied as additional characteristics and a dimension were added. Therefore, the process must re-iterate.

Iteration 5

3: Approach: At this stage of the taxonomy development process, we have reviewed all studies, however, they have not all been categorised under every characteristic of every dimension. Therefore, we will adopt the empirical-to-conceptual approach in this iteration.

4e: Due to the enhanced familiarity with the studies under examination, a consequence of the detailed assessment undertaken in the preceding iterations, we select 12 of the 23 objects:

Chung (2009)	Ito et al. (2004)	Hassanzadeh et al. (2014)
Dernoncourt et al. (2016)	Jin and Szolovits (2018)	Hirohata et al. (2008)
Gonçalves et al. (2018)	Jiang et al. (2019)	Kim et al. (2011)
Lin et al. (2006)	Liu et al. (2013)	Cohan et al. (2019)

5e: In this iteration we determine that objects may be distinguished from one another through a consideration of the features observed and engineered from the observations for classification. We decided not to deploy these using Nickerson et al.'s (2013) characteristic and dimension structure, rather opting to enrich the final taxonomy with these observations. This was deemed appropriate due to the likelihood for each study to be unique in this respect. Should these have been adopted as characteristics, the resulting taxonomy may have been unwieldy. We also identify two additional characteristics for the 'Modelling Algorithm' dimension. The first is that of a Transductive Support Vector Machine (TSVM) leveraged by Ito et al. (2004). The second is termed a SR-RCNN approach employed by Jiang et al. (2019), which is combination of both a text convolutional neural network (CNN) and a bidirectional recurrent neural network (bi-RNN). In addition, we discover 3 additional characteristics for the 'Research Domain' dimension. The first of these we refer to as 'Biomedical – EBM Generic and Biomedical – RCT'. This characteristic reflects the use of multiple datasets in the studies conducted by Dernoncourt et al. (2016) and Jin and Szolovits (2018). The second additional characteristic identified is termed 'Biomedical – Subfield Unknown', as we are unable to determine the specific research domains targeted by Hirohata et al. (2008) and Ito et al. (2004). We also identify a third research domain being 'biomedical RCT and computer science', evident in Cohan et al. (2019). This is a result of their use of the NICTA-PIBOSO, PubMed RCT (Dernoncourt et al., 2017) and their own bespoke CSAbstruct dataset, the later of which contains "2,189 manually annotated computer science abstracts with sentences annotated according to their rhetorical roles in the abstract, similar to the PUBMED-RCT categories" (Cohan et al., 2019, p. 3). Cohan et al.'s (2019) dataset combination also leads to the identification of a new dataset characteristic.

6e: The additional characteristics identified for the 'Dataset' dimension result in an adjustment of the second taxonomy:

Dimension D_2 (Dataset) Additional Characteristics

C_{211} = NICTA-PIBOSO (Kim et al., 2011), PubMed RCT (Dernoncourt et al., 2017) and CSAbstruct (Cohan et al., 2019)

The additional characteristic identified for the 'Modelling Algorithm' dimension result in an adjustment of the third taxonomy:

Dimension D_3 (Modelling Algorithm) Additional Characteristics

C_{310} = Transductive Support Vector Machine

C_{311} = SR-RNN (text-CNN + bi-RNN)

The additional characteristics identified for the 'Research Domain' dimension result in an adjustment of the fourth taxonomy:

Dimension D_4 (Research Domain) Additional Characteristics

C_{45} = Biomedical - EBM Generic and Biomedical - RCT

C_{46} = Biomedical - Subfield Unknown

C_{47} = Biomedical RCT and Computer Science

7e: Ending conditions: We fail to meet the third objective ending condition, as additional characteristics were identified. Another iteration is necessary.

Iteration 6

3: Approach: We have 1 objects remaining for secondary review after the formation of the 4 dimensions. Consequently, we adopt the empirical-to-conceptual approach in this iteration.

4e: We select the 11 remaining objects for secondary review in this iteration:

Gonçalves et al. (2019)	Teufel and Moens (1998)	Shimbo et al. (2003)
McKnight and Srinivasan (2003)	Wu et al. (2006)	Ruch et al. (2007)
Nam et al. (2016)	Xu et al. (2006)	Lui (2012)
Verbeke et al. (2012)	Yamamoto and Takagi (2005)	

5e: The performance of the modelling observed in the studies under examination became evident as a key characteristic in our assessment. The corresponding performance metrics, however, were similar in nature to the features adopted in each study as they vary significantly. Although it would be possible to indicate their success using an arbitrary benchmark, doing so may impact the ability of the taxonomy to be meaningful without being unwieldy or overwhelming. Further, forcing performance observations to adhere with the characteristic-dimension structure may require further explanatory remarks, complicating the taxonomy and reducing its utility. We decide to follow the approach adopted for communication of the features leveraged for performance metrics, by including them in the taxonomy as complementary observations. We also identify an additional characteristic for the 'Modelling Algorithm' dimension, and for the 'Research Domain' dimension.

6e: The additional characteristic identified for the 'Modelling Algorithm' dimension result in an adjustment of the third taxonomy:
Dimension D_3 (Modelling Algorithm) Additional Characteristics
C_{312} = Naïve Bayesian classifiers
The additional characteristic identified for the 'Research Domain' dimension result in an adjustment of the fourth taxonomy:
Dimension D_4 (Research Domain) Additional Characteristics
C_{48} = Computational Linguistics and Cognitive Science

7e: Ending conditions: We have examined all 23 studies twice, however, additional characteristics were observed in this iteration. Accordingly, we fail to meet the third objective ending condition and must reiterate the process.

Iteration 7

3: Approach: The last iteration did not meet required ending conditions resulting in the need to reiterate. Accordingly, we adopt the empirical-to-conceptual approach again

4e: Due to our strong ability to navigate the studies under examination at this late stage of the taxonomy development process, we review all 23 in this iteration.

5e: No additional characteristics or dimensions were identified in this iteration. Step 6e will be skipped.

7e: Ending conditions: This iteration marks the third time studies have been examined in this process. Due to the failure to discover additional characteristics or dimensions, we now meet all objective ending conditions. We also believe that all subjective ending conditions have been satisfied, as we have now produced a taxonomy that is concise, robust, comprehensive, extendible and explanatory.

5 Conclusion and Future Research

We present a taxonomy of ASCM models derived using the taxonomy development process championed by Nickerson et al. (2013) and grounded in design science theory (Nickerson et al., 2013, p. 337). The taxonomy is presented at this website:

https://www.abstractsentenceclassification.com/taxonomy.html

Our taxonomy communicates to future researchers and professionals concerned with ASCM and its utilisation the key dimensions and characteristics of known models. Dimensions discovered are the structured abstract format selected for sentence classification, datasets used for model training and evaluation, algorithms adopted to enable classification and the research domain of which the models emanate. By categorising models in this manner, the taxonomy meets its intended purpose. The taxonomy also meets the 5 qualitative attributes taxonomies should feature (Nickerson et al., 2013). Firstly, it is concise and robust, as it features clearly distinguished characteristics and dimensions. It is also comprehensive through its classification of all known ASCM models and is extendible as it can scale to include future models. It was also developed using a logical methodology and understandable language; therefore, it is explanatory.

The absence of an artefact serving as a gateway into the research domain of ASCM is a significant barrier to entry for researchers and professionals. The desire to resolve this problem was the primary motivation for this study, as without it researchers from both information systems and computer science disciplines would find it difficult to become aware of the current state of the research field and are at a disadvantage in seeking to contribute to or implement the research area's capability. We believe we have addressed this research gap through the development of a concise; robust, comprehensive, extendible and explanatory taxonomy - developed using Nickerson et al.'s (2013) information systems taxonomy development methodology.

The purpose of this taxonomy is to benefit future researchers and professionals seeking to contribute to this research domain and explore the deployment of such models into production environments. Accordingly, we recommend the following lines of inquiry for future research: Firstly, a disproportionately large number of modelling efforts emanate from biomedical related research domains. We recommend research to address this imbalance, by exploring the development and utilisation of multidisciplinary abstract sentence datasets. Secondly, we call on research exploring the implementation of ASCM capability into a real-world information system, due to both the mature state of modelling performance and the absence of known adoptions of this capability outside of ongoing efforts in the domain to exceed prior performance benchmarking. Thirdly, we call for research to extend this taxonomy and to conduct more comprehensive comparative analysis of the varying dimension characteristics of each model.

Please also note that the taxonomy will continue to be updated online as new models are identified in the literature.

References

1. Alter, S. 1977. "A Taxonomy of Decision Support Systems," Sloan Management Review (19:1), p. 39.
2. Anderberg, M. R. 2014. Cluster Analysis for Applications: Probability and Mathematical Statistics: A Series of Monographs and Textbooks. Academic press.
3. Bailey, K. D. 1994. Typologies and Taxonomies: An Introduction to Classification Techniques. Sage.
4. Cohan, A., Beltagy, I., King, D., Dalvi, B., & Weld, D. S. (2019). Pretrained language models for sequential sentence classification. arXiv preprint arXiv:1909.04054.
5. Chung, G. Y. 2009. "Sentence Retrieval for Abstracts of Randomized Controlled Trials," BMC Medical Informatics Decision Making (9:1), p. 10.
6. Dernoncourt, F., Lee, J. Y., & Szolovits, P. 2016. Neural networks for joint sentence classification in medical paper abstracts. arXiv preprint arXiv:1612.05251.
7. Dernoncourt, F., & Lee, J. Y. 2017. Pubmed 200k rct: a dataset for sequential sentence classification in medical abstracts. arXiv preprint arXiv:1710.06071.
8. Eldredge, N., and Cracraft, J. 1980. "Phytigenetic Patterns and the Evolutionary Process; Methods and Theory in Comparative Biology."
9. Emerald Group Publishing. 2019. "A 6 Step Guide to Writing an Abstract."
10. Gonçalves, S., Cortez, P., and Moro, S. 2018. "A Deep Learning Approach for Sentence Classification of Scientific Abstracts," International Conference on Artificial Neural Networks: Springer, pp. 479-488.
11. Hassanzadeh, H., Groza, T., and Hunter, J. 2014. "Identifying Scientific Artefacts in Biomedical Literature: The Evidence Based Medicine Use Case," Journal of Biomedical Informatics (49), pp. 159-170.
12. Hinde, S., & Spackman, E. 2015. Bidirectional citation searching to completion: an exploration of literature searching methods. Pharmacoeconomics, 33(1), 5-11.
13. Hirohata, K., Okazaki, N., Ananiadou, S., and Ishizuka, M. 2008. "Identifying Sections in Scientific Abstracts Using Conditional Random Fields," Proceedings of the Third International Joint Conference on Natural Language Processing: Volume-I.
14. Frasconi, P., Costa, F., De Raedt, L., & De Grave, K. 2014. klog: A language for logical and relational learning with kernels. Artificial Intelligence, 217, 117-143.
15. Ito, T., Shimbo, M., Yamasaki, T., and Matsumoto, Y. 2004. "Semi-Supervised Sentence Classification for Medline Documents," Methods (138), pp. 141-146.
16. Jiang, X., Zhang, B., Ye, Y., & Liu, Z. 2019, October. A hierarchical model with recurrent convolutional neural networks for sequential sentence classification. In CCF International Conference on Natural Language Processing and Chinese Computing (pp. 78-89). Springer
17. Jin, D., & Szolovits, P. 2018. Hierarchical neural networks for sequential sentence classification in medical scientific abstracts. arXiv preprint arXiv:1808.06161.
18. Joudrey, D. N., & Taylor, A. G. 2017. The organization of information. ABC-CLIO.
19. Kennedy-Eden, H., and Gretzel, U. 2012. "A Taxonomy of Mobile Applications in Tourism," E-review of Tourism Research, 10 (2), 47-50.
20. Khabsa, M., & Giles, C. L. 2014. The number of scholarly documents on the public web. PloS one, 9(5), e93949.
21. Kim, S. N., Martinez, D., Cavedon, L., & Yencken, L. 2011, December. Automatic classification of sentences to support evidence based medicine. In BMC bioinformatics (Vol. 12, No. 2, pp. 1-10). BioMed Central.
22. Lin, J., Karakos, D., Demner-Fushman, D., and Khudanpur, S. 2006. "Generative Content Models for Structural Analysis of Medical Abstracts," Proceedings of the hlt-naacl bionlp

workshop on linking natural language and biology: Association for Computational Linguistics, pp. 65-72.
23. Liu, Y., Wu, F., Liu, M., and Liu, B. 2013. "Abstract Sentence Classification for Scientific Papers Based on Transductive Svm," Computer Information Science (6:4), p. 125.
24. Lui, M. 2012. "Feature Stacking for Sentence Classification in Evidence-Based Medicine," Proceedings of the Australasian Language Technology Association Workshop 2012, pp. 134-138.
25. McKnight, L., and Srinivasan, P. 2003. "Categorization of Sentence Types in Medical Abstracts," AMIA Annual Symposium Proceedings: American Medical Informatics Association, p. 440.
26. Nakayama, T., Hirai, N., Yamazaki, S., and Naito, M. 2005. "Adoption of Structured Abstracts by General Medical Journals and Format for a Structured Abstract," Journal of the Medical Library Association (93:2), p. 237.
27. Nam, S., Kim, S.-K., Kim, H.-G., Ngo, V., and Zong, N. 2016. "Structuralizing Biomedical Abstracts with Discriminative Linguistic Features," Computers in Biology Medicine (79), pp. 276-285.
28. Nickerson, R. C., Varshney, U., and Muntermann, J. 2013. "A Method for Taxonomy Development and Its Application in Information Systems," European Journal of Information Systems (22:3), pp. 336-359.
29. Rawat, S., & Meena, S. 2014. Publish or perish: Where are we heading? Journal of research in medical sciences: Journal of Isfahan University of Medical Sciences, 19(2), 87.
30. Recker, J., 2013 Scientific Research in Information Systems: A Beginner's Guide Springer Berlin.
31. Ruch, P., Boyer, C., Chichester, C., Tbahriti, I., Geissbühler, A., Fabry, P., Gobeill, J., Pillet, V., Rebholz-Schuhmann, D., and Lovis, C. 2007. "Using Argumentation to Extract Key Sentences from Biomedical Abstracts," International Journal of Medical Informatics (76:2-3), pp. 195-200.
32. Shimbo, M., Yamasaki, T., and Matsumoto, Y. 2003. "Using Sectioning Information for Text Retrieval: A Case Study with the Medline Abstracts," Second International Workshop on Active Mining (AM'03).
33. Teufel, S., and Moens, M. 1998. "Sentence Extraction and Rhetorical Classification for Flexible Abstracts," AAAI Spring Symposium on Intelligent Text summarization, pp. 89-97.
34. Verbeke, M., Van Asch, V., Morante, R., Frasconi, P., Daelemans, W., and De Raedt, L. 2012. "A Statistical Relational Learning Approach to Identifying Evidence Based Medicine Categories," Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning: Association for Computational Linguistics, pp. 579-589.
35. Ware, M. and Mabe, M. 2015, "The STM Report: An overview of scientific and scholarly journal publishing" (2015). http://digitalcommons.unl.edu/scholcom/9
36. Wu, J.-C., Chang, Y.-C., Liou, H.-C., and Chang, J. S. 2006. "Computational Analysis of Move Structures in Academic Abstracts," COLING/ACL: Association for Computational Linguistics, pp. 41-44.
37. Xu, R., Supekar, K., Huang, Y., Das, A., and Garber, A. 2006. "Combining Text Classification and Hidden Markov Modeling Techniques for Structuring Randomized Clinical Trial Abstracts," AMIA Annual Symposium Proceedings: American Medical Informatics Association, p. 824.
38. Yamamoto, Y., and Takagi, T. 2005. "A Sentence Classification System for Multi Biomedical Literature Summarization," 21st International Conference on Data Engineering Workshops (ICDEW'05): IEEE, pp. 1163-116