

## **Co-Concept Analysis in Bioinformatics Field by Yewno Based on Experts' Viewpoints**

### **Masoume Kiani**

Ph.D. candidate in Knowledge & Information Sciences, Department of Knowledge & Information Sciences, Faculty of Education & Psychology, University of Isfahan, Isfahan, Iran, m.kiani@edu.ui.ac.ir

### **Asefeh Asemi**

Associate Prof. Department of Knowledge & Information Sciences, Faculty of Education & Psychology, University of Isfahan, Isfahan, Iran and Doctoral School of Business Informatics, Corvinus University of Budapest, Budapest, Hungary  
Corresponding Author:  
asemi.asefeh@uni-corvinus.hu

### **Mozaffar CheshmehSohrabi**

Associate Prof. Department of Knowledge & Information Sciences, Faculty of Education & Psychology, University of Isfahan, Isfahan, Iran mo.sohrabi@edu.ui.ac.ir

### **Ahmad Shabani**

Professor, Department of Knowledge & Information Sciences, Faculty of Education & Psychology, University of Isfahan, Isfahan, Iran shabania@edu.ui.ac.ir

\* Received: 05 January 2020

Accepted: 27 June 2020

### **Abstract**

In this paper, the opinions of bioinformatics experts were sought to examine the knowledge semantic maps. The research question was whether the mapping of bioinformatics field articles could contribute to the reflection of knowledge map in this field; can knowledge within articles be demonstrated through a knowledge map? The research type was applied and analytical. Data collection was performed in two ways. First, the data were collected from scientific databases by using semantic search software and drawing knowledge maps by the software. Second, Delphi technique was used in two stages in order to obtain the opinion of experts regarding these knowledge semantic maps and co-concepts in bioinformatics field. In this method, first, the co-concepts map of the articles exported by Yewno and the knowledge semantics map were presented to the experts to get their viewpoints. By implementing the Delphi approach, the opinions of experts in examining the application of knowledge semantic maps resulted in the phonological analysis of articles in this field. The analysis results demonstrated that the knowledge map of bioinformatics studies dynamically developed over time. The continuous conduction flow of scientific studies in this field brought continuous changes to the knowledge of this field.

**Keywords:** Knowledge Semantics Map, Bioinformatics, Concepts Relationship, Semantic Relations, Co-concepts, Scientific Studies, Yewno

### **Introduction**

Nowadays, researchers in unique medical fields are showing interest in finding out the concepts' knowledge map governing special areas of research. Most of the previous researches rely on quotation studies and consist of quantitative records to deal with this problem. Now, the

exploration of the knowledge map based totally on records ecology seeks to carry out a deeper analysis of the relation amongst distinct information factors in an informational ecosystem. Within the records ecology attitude, various techniques are applied to explore interrelationships among record elements. This brings forth a pattern that well defines the knowledge map in any field. Records ecology additionally aims at describing and analyzing the statistics structures primarily based on metaphoric uses of such principles in environmental ecology as biotope, species, evolution, ecosystem, and evolution and equilibrium to observe the distribution and abundance of organisms, relationships among them, and how the organisms can have an effect on or be affected by their surroundings (Eryomin, 1998). Global encyclopedia of records and library science defines statistics ecology as: "to have a look at interpersonal communication, sports, technology, and statistics environment" (Feather & Sturges, 2003). As referred to earlier, the records ecology idea has a metaphorical basis, which considers surrounding facts as surroundings.

A knowledge map attempts to investigate the relationships between ecosystems in information environments and the relationships of information elements with such environments by finding their patterns, processes, and relationships. Similar to the ecology of nature and wildlife that investigates processes and relationships governing natural and animal systems, the knowledge map can investigate patterns and relationships governing information environments (Burgin and Zhong, 2018). The knowledge map can provide a new framework for interpreting complicated relationships between organizations, technologies, and information, or between information elements in information environments. In fact, the knowledge map is clearly capable of identifying the knowledge formed in a field and the relationships between knowledge points, indicating the concept features of that field. This includes using methods, models, algorithms, and techniques by which researchers can study the collected knowledge and reveal distributed concept fields, emerging topics, hot topics, and any growing trends in a field. The results of ecological studies conducted in information environments significantly contribute to investigating and understanding the nature and statuses of conducted studies, identifying the research fields of previous studies, modeling study trends, indicating considerable advances, representing the growth and evolution of current studies, and predicting the concept patterns of future studies (Wang, Guo, Yang, Chen & Zhang, 2017).

### **Search tool vs. Yewno**

Yewno provides Augmented Intelligence. Yewno's mission is that of extracting knowledge from an overwhelming quantity of unstructured and structured data. The technology helps to overcome the "Information Overload" problem, to research, and to understand the world in a more natural manner. It is inspired by the way humans process information from multiple sensorial channels, and it leverages state-of-the-art Computational Linguistics, Network Theory, Machine Learning, as well as methods from classical Artificial Intelligence ([www.yewno.com](http://www.yewno.com)). The main difference between keyword search and semantic-based search engines is time consumption by the user. Users spend a long time to sort the retrieved results in keyword engine search. Whereas in case of semantic based search engines, there is no sorting of search results. Keyword-based search engines are unable to give relevant search results as they do not know the exact meaning of the concepts used (Malve & Chawan, 2015). Usually, when we try to search for papers using a standard search engine or search tool, we are not happy with the results we get. There are serious limitations to using keyword searches especially in

some subjects. When we are looking for papers using a generic search engine, we would get articles that mentioned some aspects non-related to our subject. Moreover, articles will be found that are not particularly relevant to our search. However, a semantic search engine like Yewno, powered by scientific vocabularies and a disambiguation system, will just focus on results featuring our search, giving us context specificity. If we need even more accuracy and want to find a specific subject, we would be required to do a search for specific concepts in a search tool. It can be a pretty long list of synonym derivatives. While being a good semantic search system, Yewno does all this for you when it is indexing so that you do not have to worry when searching. Semantic search provides researchers with a variety of features, including Transformative Data Integration, Group Level Searches, and Connections (Semantic Search vs. Keyword Search – Copyright Clearance Center, n.d.):

*Transformative Data Integration:* Having done this, we are then set up for better downstream data analysis because our conversion from unstructured to structured (typed) data is more accurate. We can then connect our enriched, structured data to databases and other systems, giving enhanced data connectivity across the organization and speeding up analysis.

*Group Level Searches:* Great semantic search provides taxonomic relationships between its entities, so higher-order searches are possible. We can find a bunch of articles that would mention things related to our search that was found in an article and where they sit in the taxonomy. What we are looking for is a rich set of taxonomies covering areas. A good semantic search engine will actually embed the concepts within the plain text. The query time is really quick and extremely accurate, all because you do not have to do synonym expansion. In essence, you have far more control over the granularity of your searches than in generic search engines. That is hard to do in a search tool that does not leverage taxonomy data in one step.

*Connections:* Additionally, you could examine the co-occurrence data to get a feel for the landscape. We could also look at co-occurrences on a sentence level. Sentence level co-occurrences are stronger indicators of a real association between entities than document level. Why? Because at a document level, you might find entities in the concepts section that hold spurious and unrelated terms. A comprehensive autocomplete index helps guide your searches. But we are not limited to entities and types that have already been curated. We can build our own vocabulary or use plain text. We can see also an entity type and a plain text query to identify documents that mention either word. The semantically enabled search is as good as the vocabularies it is built on. Against the depth and breadth that semantic search offers, keyword search simply cannot compete in terms of accuracy, full awareness, or efficiency. Semantic search (Yewno) allows us to buy back valuable time that would otherwise be spent sifting through huge amounts of documents, and even convert textual data into something we can integrate across your systems, thanks to entity recognition.

Across various academic fields, bioinformatics has gained attention in different areas such as information technology and biological science in recent years. This is because of high biological data availability and the need for converting these data into usable information and knowledge (Chen, 2005). Bioinformatics has also gained much attention from researchers in the last few years as an interdisciplinary science that, along with biology, can be applied to other fields of studies (Jiang, Zhang, & Zhang, 2013). A definition depicts bioinformatics as “the collection, classification, storage, and analysis of biochemical and biological information using computers especially as applied to molecular genetics and genomics” (Merriam-Webster collegiate Dictionary, 1993). Bioinformatics has been considered as a part of biology, according

to The National Center for Biotechnology Information (NCBI), and usable in developing biological data collection tools to employ the gathered data for biological discovery and prediction. This subject of study provides all of the computational methods and theories for solving a biological problem, such as the problem with using a model and dataset-related problems with the use of a model (Bioinformatics: Computational Biology, 2020).

The main emphasis in the current study was on this fact that, despite the importance of the knowledge semantic maps in bioinformatics, little attention has been paid to this inquiry with respect to previous studies. To conduct a large-scale and effective research in this area, it is necessary to draw a topic map. In doing so, analyses of scientific information are found as a promising method. The analyses are performed in different sizes of information and are applied to determine the relationships among any type of concepts in different types of knowledge.

The review of the literature showed that there was no determined main (cores) and minor (close to the core) concepts of bioinformatics published papers by Yewno (<https://www.yewno.com/about>). Thus, we used Yewno to identify the core and close to core themes in bioinformatics. The knowledge map was confined to the discovered concepts of published papers in the field of bioinformatics. The bioinformatics' knowledge map shows the conceptual relationship of sources (concepts or information elements) in the information environment. We investigated the concepts of knowledge maps and how their relationships influence each other based on in-depth expert viewpoints. These aims were addressed by the following questions:

1. What are the core concepts in the knowledge semantic map of the bioinformatics published papers in the scientific databases?
2. How is the relation or connection between the core concepts in the field of bioinformatics in the specified knowledge maps?
3. What are the experts' viewpoints about the concepts and relationships between them in the knowledge maps in the field of bioinformatics?

In brief, this study helps to support the important concepts of bioinformatics knowledge and the relations among them in the form of knowledge semantic maps.

### Methodology

Carrying out a knowledge map investigation can result in practical benefits. The present research uses an informatics approach to analyze the knowledge map of concept elements of an information environment. The research type was applied and analytical. Data collection was performed in two ways. First, the data were collected from scientific databases by using semantic search software and drawing knowledge maps by the software. Second, Delphi technique was used in two stages in order to obtain the opinion of experts regarding these knowledge semantic maps and co-concepts in bioinformatics field. The research sample included all papers that were published in bioinformatics for the years 1980 to 2018 indexed in Clarivate database (WoS). To exclude the data, a concept search strategy was used. First, the study interviewed bioinformatics experts with the generated concepts. Next, a conceptuallygrouped key terms were created based on the experts' opinions and Medical Subject Headings (MeSH)<sup>1</sup> concepts. After selecting the concepts, the research method was

---

<sup>1</sup> Medical Subject Headings (MeSH) are controlled vocabularies that are arranged in a hierarchy and created by the United States National Library of Medicine (NLM). The purpose of MeSh is for indexing, cataloging, and information searching of biomedical and health-related information. Also it includes the subject headings from

worked out through the following steps:

1. Bioinformatics concept was searched in a specific topic area by Yewno. The following eight concepts excluded two group areas: that is, Life Sciences and Computer Sciences:

Bioinformatics, BioRuby, Multiple sequence alignment, Gene, Chemical biology, Genomics, RNA-Seq, DNA sequencing

2. Yewno can only expand the eight top concepts and presents the semantic knowledge map. Therefore, all eight concepts were inspected and extended by Yewno to find the relation between the basic concepts in bioinformatics.

3. Implementing Delphi technique in two steps

In this study, the researcher invited a number of bioinformatics experts to collaborate, and eventually, five of them collaborated. Therefore, a panel consisting of five bioinformatics experts who volunteered to participate in this research was formed to implement the Delphi method. It should be noted that the method of snowball sampling was used to identify and select the statistical sample of the present study.

Table 1

*Demographics of People participating in Delphi technique*

No	Gender	Age	Degree	work experience	Affiliation
1	Male	56	Ph. D	28	National Institute of Genetic Engineering and Biotechnology
2	Male	40	Ph. D	11	Institute for Studies in Theoretical Physics and Mathematics (IPM)
3	Male	52	Ph. D	23	Shahid Beheshti University
4	Male	55	Ph. D	26	University of Tehran
5	Female	39	Ph. D	8	Institute for Studies in Theoretical Physics and Mathematics(IPM)

The validity and reliability of the Delphi technique depends on the validity and reliability of the questions that are prepared. Unfortunately, it is not easy to control Delphi's validity and reliability (Chas & Bryan, 1998). In research where the members of the study represent the group or field of knowledge in question, the validity of the content is guaranteed (Turoff & Linstone, 2008). In addition, because researchers have not influenced the process of conducting research, the tool of formal narrative is also encountered (Landeta, 2006). Therefore, data collection tools have validity and reliability.

## Results

Because it is not possible to explain all the concepts pertaining to the study findings in this article, ten core concepts that had the highest number of records are presented according to Table 2.

---

databases including MEDLINE/PubMed, the NLM Catalog and other NLM databases (Medical Subject Headings 2020).

Table 2

*Core Concepts*

No	Concepts	Definition	Source
1	Gene Expression	Gene expression is the process by which information from a gene is used in the synthesis of a functional gene product. These products are often proteins, but in non-protein-coding genes such as transfer RNA (tRNA) or small nuclear RNA (snRNA) genes, the product is a functional RNA.	Crick (1958). "On protein synthesis". Symposia of the Society for Experimental Biology. 12: 138–63.
2	DNA sequencing	DNA sequencing is the process of determining the nucleic acid sequence - the order of nucleotides in DNA. It includes any method or technology that is used to determine the order of the four bases: adenine, guanine, cytosine, and thymine.	Behjati & Tarpey (2013). "What is next generation sequencing?". Archives of Disease in Childhood. Education and Practice Edition. 98 (6): 236–8.
3	GeneMark	GeneMark is a generic name for a family of ab initio gene prediction programs developed at the Georgia Institute of Technology in Atlanta.	Orodovsky & McIninch, (1993). "GeneMark: parallel gene recognition for both DNA strands." Computers & Chemistry. 17 (2): 123–133.
4	Sequence assembly	In bioinformatics, sequence assembly refers to aligning and merging fragments from a longer DNA sequence in order to reconstruct the original sequence.	Wolf (2019). "De novo genome assembly versus mapping to a reference genome". University of Applied Sciences Western Switzerland.
5	Genome	In the fields of molecular biology and genetics, a genome is the genetic material of an organism.	Ridley (2006). Genome: the autobiography of a species in 23 chapters. New York: Harper Perennial.
6	Gene	In biology, a gene is a sequence of nucleotides in DNA or RNA that encodes the synthesis of a gene product, either RNA or protein.	Watson, Baker, Bell, Gann, Levine & Losick (2013). Molecular Biology of the Gene (7th ed.). Benjamin Cummings.
7	Metagenomics	Metagenomics is the study of genetic material recovered directly from environmental samples.	Poinar et al. (2006). "Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA". Science. 311 (5759): 392–4.
8	Protein structure prediction	Protein structure prediction is the inference of the three-dimensional structure of a protein from its amino acid sequence—that is, the prediction of its folding and its secondary and tertiary structure from its primary structure.	Zhang & Skolnick (2005). "The protein structure prediction problem could be solved using the current PDB library". Proceedings of the National Academy of Sciences of the United States of America. 102(4):1029–34.

No	Concepts	Definition	Source
9	RNA- Seq	RNA-Seq (named as an abbreviation of "RNA sequencing") is a particular technology-based sequencing technique, which uses next-generation sequencing (NGS) to reveal the presence and quantity of RNA in a biological sample at a given moment, analyzing the continuously changing cellular transcriptome.	Chu & Corey (2012). "RNA sequencing: platform selection, experimental design, and data interpretation". <i>Nucleic Acid Therapeutics</i> . 22 (4): 271–4.
10	Chemical Biology	Chemical biology is a scientific discipline spanning the fields of chemistry and biology.	Terai & Nagano (2008). "Fluorescent probes for bioimaging applications". <i>Current Opinion in Chemical Biology</i> . 12 (5): 515–21.

Figure 1 shows the relationship between bioinformatics concept and other core concepts (10 top) that were revealed by Yewno. These concepts include "Gene Expression", "DNA sequencing", "GeneMark", "Sequence assembly", "Genome", "Gene", "Metagenomics", "Protein structure prediction", "RNA- Seq", and "Chemical Biology". It is notable that the length of the edges does not have a specific meaning. The size of the node of the central concept indicates the centrality of that concept. It shows the connection of the main concept with other nodes (concepts).

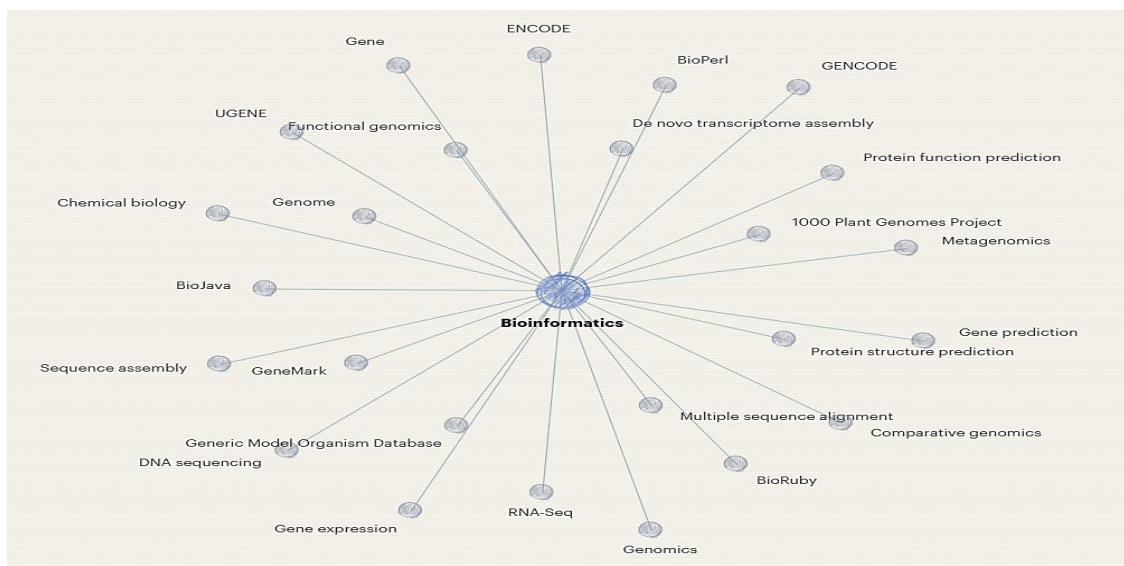


Figure 1. Relationship between bioinformatics concept and other core concepts

Figure 2 shows the relationship between bioinformatics concept and other core concepts in life sciences that were revealed by Yewno. These concepts include "BioRuby", "Multiple sequence alignment", "BioJava", "Generic Model Organism Database", "UGENE", "GENCODE" and "Bioperi".

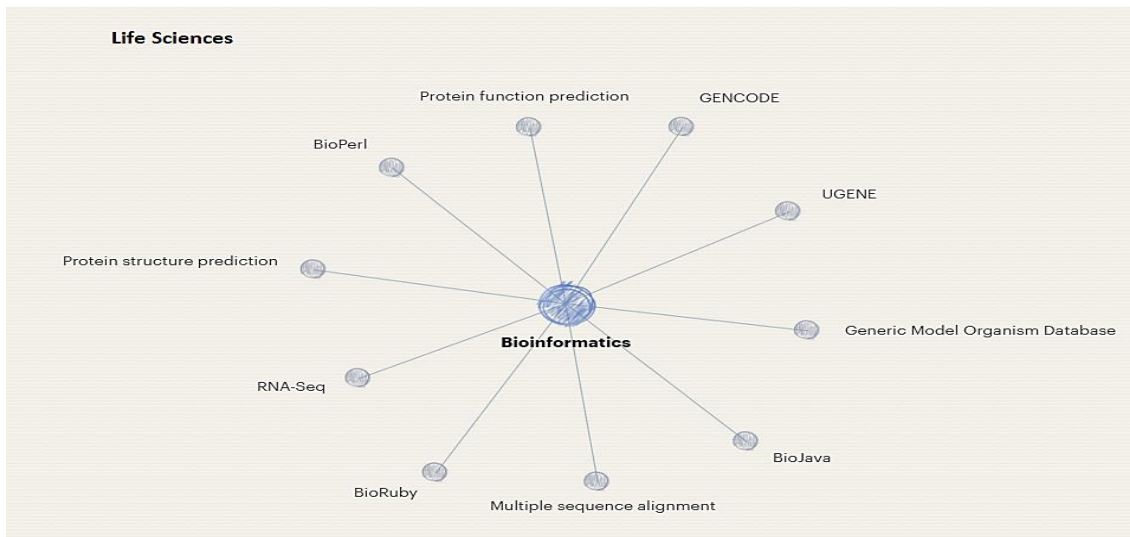


Figure 2. Relationship between bioinformatics concept and core concepts in life sciences

Figure 3 shows the relationship between bioinformatics concept and other core concepts in computers and computer sciences that were revealed by Yewno. These concepts include "RNA-Seq", "Protein function prediction", "Bioperi", "BioJava", and "GENCODE".

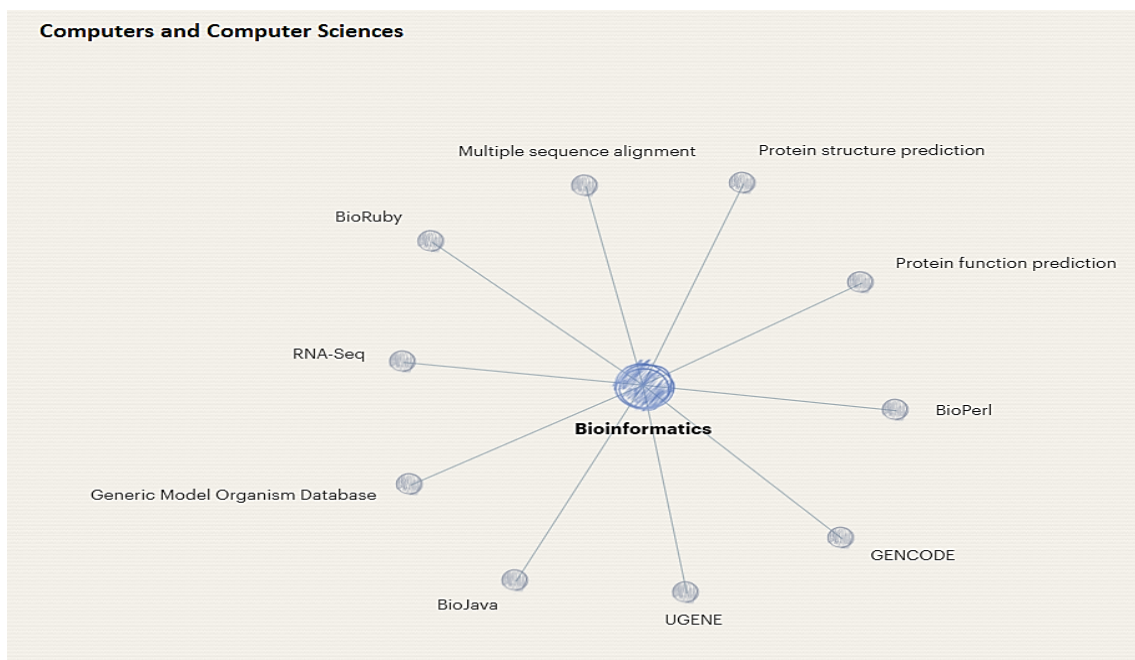


Figure 3. Relationship between bioinformatics concept and core concepts in computers and computer sciences

#### *Ward's Hierarchical Clustering of Concepts of Bioinformatics Papers*

The hierarchical clustering algorithm by Ward was used to determine the clusters that made up the bioinformatics domain. The purpose of using the hierarchical clustering method in the present study was to determine the relationships between the core concepts of the bioinformatics field. Figure 4 shows the relationship between bioinformatics concept and seven inspected and expanded concepts by Yewno. These concepts include "BioRuby", "DNA sequencing", "Multiple sequence alignment", "RNA sequence", "Genomics", "Gene", and



of thematic relationships in bioinformatics fields? Can knowledge within articles be demonstrated through a knowledge map? In order to obtain the opinion of experts regarding these knowledge semantic maps, Delphi technique was used in two stages. First, the seven questions (table 3) were presented to them. In this stage, the researcher reached saturation stage by 4 questions. In order to obtain more information and reach the saturation stage in 3 other questions, we performed the second stage of Delphi among the specialists in bioinformatics. During these steps, the concepts and terminology of the topics obtained from the articles along with the general map of the topics and the network of the most important topics in the field of bioinformatics were presented to the experts. Then, by implementing the Delphi approach, the opinions of experts were sought to examine the application of knowledge semantic maps.

Table 3

*Viewpoints of bioinformatics experts in the first stage of Delphi technique*

No.	Questions	Experts Answer
1	Are words and relations between both the introduction and understanding of the field of bioinformatics effective for the audience area?	<ol style="list-style-type: none"> <li>1. Very much</li> <li>2. Low</li> <li>3. It is enough to introduce the field of bioinformatics.</li> <li>4. These words can be used to help identify initial areas of bioinformatics.</li> <li>5. The words can be so high that people in the area other than the bioinformatics activities can become familiar with the concepts and terminology of the field.</li> </ol>
2	To what extent are the relations of concepts shown in the field of bioinformatics applicable in real world?	<ol style="list-style-type: none"> <li>1. Relations concepts can be used in practice and laboratory.</li> <li>2. A lot</li> <li>3. The relationships between the concepts of the articles are largely applicable.</li> <li>4. Relations concepts are shown to be the other's views.</li> <li>5. It may include some topics but not all of them.</li> </ol>
3	How can you identify the emerging research topics in the field of bioinformatics through the relation concepts shown ?	<ol style="list-style-type: none"> <li>1. It may not.</li> <li>2. Very little</li> <li>3. Not able to have an impact.</li> <li>4. Most of the emerging research topics in the field of bioinformatics can be seen in the relationship between concepts.</li> <li>5. To some extent</li> </ol>
4	How effective are the relationships of the concepts shown in drawing the content of the bioinformatics domain?	<ol style="list-style-type: none"> <li>1. It has been able to fully illustrate the content of the field.</li> <li>2. It could show 50 to 60 percent.</li> <li>3. Draws the content of the domain to an acceptable level.</li> <li>4. Has been very successful in providing image content in this area.</li> <li>5. The general picture shows.</li> </ol>
5	Are you sure words can be considered as specialized vocabulary field of bioinformatics?	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. Yes</li> <li>3. Very much</li> <li>4. To some extent</li> <li>5. Yes</li> </ol>
6	Will the relations between concepts shown be a roadmap for policy of the Bioinformatics?	<ol style="list-style-type: none"> <li>1. Yes, you can</li> <li>2. No</li> <li>3. There is a need for policymakers in this area.</li> <li>4. No, it is not effective</li> <li>5. To some extent</li> </ol>

No.	Questions	Experts Answer
7	Have the relationships of the concepts shown been able to identify the technologies that affect the field of bioinformatics that you were not aware of and add to your knowledge?	1. IT field I know. 2. No. 3. Relations concepts show that technologies affecting the field of bioinformatics should be identified. 4. No. 5. No.

Expert opinions in the field of bioinformatics are shown in the first stage of Delphi technique. The findings show that the opinions of experts in the first stage were convergent in four questions 1, 2, 4 and 5. From the answers of the experts to the first question of the questionnaire, it was concluded that the concepts of the field of bioinformatics and the relations between them could be effective in recognizing and introducing this scientific field. In addition, the results obtained from the second question of the questionnaire show that the thematic relations drawn in the field of bioinformatics are practically applicable. The results obtained from the fourth question of the questionnaire show that the thematic relations shown in the field of bioinformatics could be successful in illustrating the content of this scientific field. Moreover, the results obtained from the fifth question of the questionnaire show that the words used in the thematic maps of the bioinformatics field could be used as specialized words or concepts.

In the next step, in order to reach a consensus on the other three questions, the questions were returned to the experts along with the results of the first stage comments. At this point, there is a consensus among experts. The following is a summary of the second stage of the Delphi technique in Table 4.

Table 4

*Viewpoints of bioinformatics' experts in the second stage of Delphi technique*

No.	Questions	Experts' Answer	Concluding by the Research Team
1	Map of the knowledge presented: how can you identify the emerging research topics to be effective in the field of bioinformatics?	Knowledge semantic maps can be effective for professionals to identify new research in this area.	Knowledge semantic maps can be effective in the identification of emerging issues.
2	Will the relations between concepts shown be a roadmap for policy of the Bioinformatics?	If the deficiencies are resolved, it can be the roadmap.	It can be a roadmap for policy makers and managers.
3	Are relations between the concepts shown in the identification of technologies affecting the field of bioinformatics contributing to your knowledge?	It needs to review. After that, we will be able to identify new technologies affecting this field.	The scientific maps are not able to help in identification of the new technologies affecting the field.

Table 4 shows three questions the experts disagreed with in the first step. The table also shows the answers of the experts and their assumptions of the research team before sending the questions to the experts. These assumptions are based on a literature study on the application of scientific and thematic maps. As can be seen, in some cases, the results are consistent with predictions, and in others, they are inconsistent. The co-word analysis and co-occurrence analysis allow the structure of internal and external relationships of thematic elements to be displayed objectively. This can help to understand the relationship between concepts and concepts in a field. Knowledge semantic maps can be effective in understanding the relationship structure of concepts in scientific fields. The results of the present study in the second phase of Delphi technique showed that the maps obtained from the co-word analysis of articles in the field of bioinformatics could be useful in identifying core concepts for experts. If the knowledge semantic maps of the bioinformatics field become a roadmap, they can have a positive impact on the research policies and programs of this field. Knowledge semantic maps can also help identify non-specialist or beginner-affected technologies in the field, but they cannot give experts in the field of cognitive research more knowledge and add to their scientific knowledge.

Figure 5 shows the research pattern. This pattern has been compiled based on the opinions of bioinformatics experts and with the help of Delphi technique.

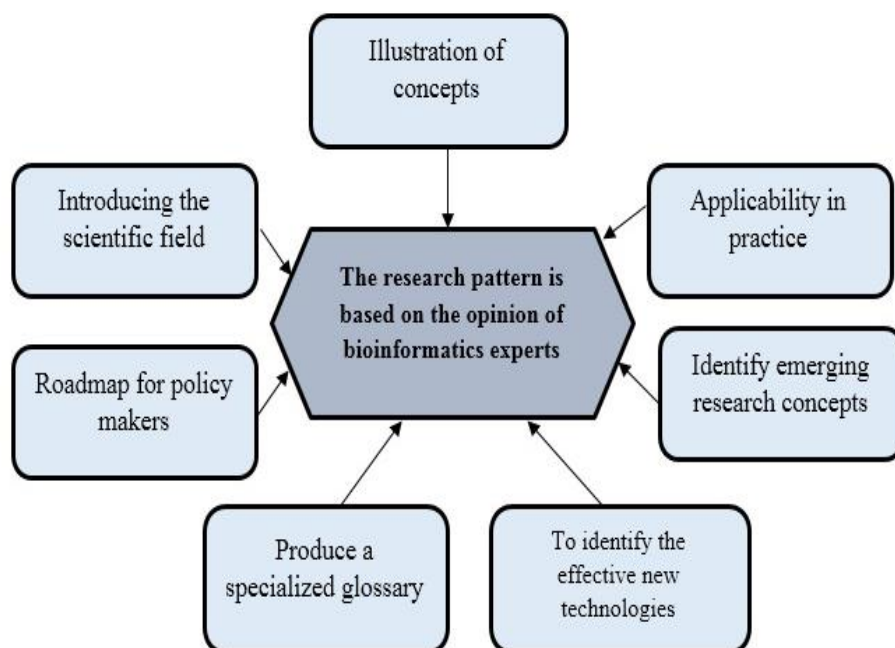


Figure 5. Research pattern based on the viewpoints of the bioinformatics experts

As Figure 5 shows, co-word analysis is a type of content analysis. This analysis focuses on outstanding concepts in scientific maps. In order to identify the basic concepts, it is necessary to have a complete set of thematic words in order to clarify the importance of the basic concepts. One of the functions of drawing scientific maps is to extract terms from a field of knowledge. The pattern presented in this study helps to understand the importance of concepts related to the field of bioinformatics. The pattern also demonstrates the applicability of relationships between concepts, illustration of field content, usability as specialized vocabulary, development of roadmaps for policy makers and planners, and identification of effective technologies in this field.

### Discussion and conclusion

The data analysis led to the formation of seven concept clusters in this study. It was also revealed that concepts such as proteomics, microarray, genomics, gene expression, computational biology, database, and data mining were the main concepts of this field. Using knowledge maps in bioinformatics, the concepts knowledge map between the studies conducted in this field can be analyzed.

The results also demonstrated that concepts such as alternative splicing, *Arabidopsis thaliana*, "Arabidopsis", "bioinformatics analysis", "cloning", "comparative genomics", and "computational biology" had been the most frequently used concepts in bioinformatics studies. Thus, they were identified as the most frequently used concepts. The results of the drawing of co-word maps in different time periods demonstrated changes and stabilities in informatics-related concepts and words. Some words, such as bibliometrics analysis, have been present in the entire studied years, while some others disappeared over time.

The present study attempted to identify the concepts of bioinformatics by drawing a knowledge map. The analysis results demonstrated that the knowledge map of bioinformatics studies dynamically developed over time. The continuous production of scientific studies in this field, the same as other scientific fields, creates continuous changes in the knowledge map of this field. Considering the extensiveness and interdisciplinarity of this field, it continues to interact with other sciences. Some bioinformatics fields are addressed more considering the requirements of scientific communities and countries. The present results also demonstrated that concept maps could contribute to understanding the existing knowledge status and guiding scientific policies in this field. The drawn maps provide a clear representation of research topics in informatics and the relationships between different topics. These maps represent changes and stabilities in bioinformatics-related topics and words over time. Moreover, new concepts have appeared as recombination of the existing words in interaction with new evolutions and technologies. Overall, the results of the present study and of similar studies can respond to these questions: What topics are considered by the scientific community? What are the different topics and sub-topics of scientific fields and how have they evolved?

The results were found to be useful to the planning of education and research and to the development of a framework for collaborative work of experts in researches in order to extend the knowledge of the bioinformatics area of study. The findings also involve providing a good information source for decision- and policy-makers in performing scientific activities in bioinformatics at universities. In addition, it helps researchers to determine policies for clinical research by providing proper guidelines. Considering that this study used scientific databases to obtain resources, it is recommended to conduct a study in this field by other accurate analysis methods on international databases such as Scopus and MEDLINE. The comparison of such a study to those of the present study would yield valuable results.

### Acknowledgments

This study is the result of a doctoral thesis in the field of Knowledge & Information Sciences in the University of Isfahan. The authors of this research from the University of Isfahan are grateful for financial and spiritual support for this thesis.

### References

Behjati, S. & Tarpey, P.S. (2013). What is next generation sequencing?. Archives of Disease in

- Childhood. *Education and Practice Edition*. 98 (6), 236–238.
- Bioinformatics: Computational Biology. (2020). Retrieved from <http://www.ncbi.nlm.nih.gov/mesh/?term=bioinformatics>
- Burgin, M., & Zhong, Y. (2018). Information Ecology in the Context of General Ecology. *Information*, 9(3), 57. <https://doi.org/10.3390/info9030057>
- Chas, C. & Bryan, G. (1998). Utilizing the Delphi technique in policy discussion: a case study of a privatized utility in Britain. *Public Administration*. 76(3), 431–49.
- Chen, Y. P. P. (2005). *Bioinformatics Technologies*. Berlin Heidelberg: Springer-Verlag.
- Chu, Y. & Corey, D.R. (2012). RNA sequencing: platform selection, experimental design, and data interpretation. *Nucleic Acid Therapeutics*. 22 (4), 271–4.
- Crick, F. H. (1958, January). On protein synthesis. In *Symposia of the Society for Experimental Biology* (Vol. 12, No. 138-63, p. 8).
- Eryomin, A. L. (1998). Knowledge map: A viewpoint. *International Journal of Environmental Studies*, 54(3-4), 241-253. doi: 10.1080/00207239808711157
- Feather, J. & Sturges, P. (2003). *International encyclopedia of information and library science*. London: Routledge. <https://doi.org/10.4324/9780203403303>
- Jiang, R., Zhang, X. & Zhang, M. Q. (2013). *Basics of bioinformatics. Lecture Notes of the Graduate Summer School on Bioinformatics of China*. Berlin: Springer-Verlag. <https://doi.org/10.1007/978-3-642-38951-1>
- Landeta, J. (2006). Current validity of the Delphi method in social sciences. *Technological Forecasting and Social Change*, 73(5), 467-82.
- Malve, A., & Chawan, P. P. M. (2015). A comparative study of keyword and semantic based search engine. *International Journal of Innovative Research in Science, Engineering and Technology*, 4(11), 11156-11161.
- Medical Subject Headings. Retrieved from <https://www.nlm.nih.gov/mesh/meshhome.html>
- Merriam-Webster's collegiate dictionary (1993). Springfield, MA: Merriam-Webster.
- Orodovsky, M. & McIninch, J. (1993). GeneMark: parallel gene recognition for both DNA strands. *Computers & Chemistry*. 17 (2), 123–133.
- Poinar, H. N., Schwarz, C., Qi, J., Shapiro, B., MacPhee, R. D., Buigues, B., ... & Rapp, M. (2006). Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *science*, 311(5759), 392-394.
- Ridley, M. (2006). *Genome: the autobiography of a species in 23 chapters*. New York: Harper Perennial.
- Semantic Search vs. Keyword Search – Copyright Clearance Center. (n.d.). Retrieved from <https://www.copyright.com/blog/semantic-search-vs-keyword-search/>
- Terai, T. & Nagano, T. (2008). Fluorescent probes for bioimaging applications. *Current Opinion in Chemical Biology*. 12 (5), 515–21.
- Turoff, M. & Linstone, H.A. (2008). *The Delphi method: techniques and applications*. Retrieved from: <http://is.njit.edu/pubs/delphibook/>
- Wang, X., Guo, Y., Yang, M., Chen, Y. & Zhang, W. (2017). Knowledge map research: past, present, and future. *Information Technology and Management*, 18(1), 27-39. doi: 10.1007/s10799-015-0219-3
- Watson, J.D., Baker, T.A., Bell, S.P., Gann, A., Levine, M. & Losick, R. (2013). *Molecular Biology of the Gene*. (7th ed.). Benjamin Cummings.
- Wolf, B. (2019). *De novo genome assembly versus mapping to a reference genome*. University

of Applied Sciences Western Switzerland.

Yewno. Retrieved from <https://www.yewno.com/about>

Zhang, Y., & Skolnick, J. (2005). The protein structure prediction problem could be solved using the current PDB library. *Proceedings of the National Academy of Sciences*, 102(4), 1029-1034.