



# Current Perspectives on the Application of Bayesian Networks in Different Domains

Galina M. Novikova<sup>(✉)</sup>  and Esteban J. Azofeifa 

RUDN University, Moscow, Russia  
novikova\_gm@rudn.university, esteban.azofeifa@gmail.com

**Abstract.** Bayesian networks are powerful tools for representing relations of dependence among variables of a domain under uncertainty. Over the last decades, applications of Bayesian networks have been developed for a wide variety of subject areas, in tasks such as learning, modeling, forecasting and decision-making. Out of hundreds of related papers found, we picked a sample of 150 to study the trends of such applications over a 16-year interval. We classified the publications according to their corresponding domain of application, and then analyzed the tendency to develop Bayesian networks in determined areas of research. We found a set of indicators that help better explain these tendencies: the levels of formalization, data accuracy and data accessibility of a domain, and the level of human intervention in the primary data. The results and methodology of the current study provide insight into potential areas of research and application of Bayesian networks.

**Keywords:** Bayesian networks · Uncertainty · Domain Formalization · Human intervention · Data accuracy  
Data accessibility

## 1 Background

Uncertainty is a constant in most aspects of everyday life, making it necessary to develop tools that can aid in minimizing it. Bayesian networks (BNs) [16] are nowadays a popular and important method for reasoning under conditions of uncertainty in artificial intelligence (AI). BNs are popular for modeling a wide variety of domains because they facilitate both the construction of the models and the understanding of the domain.

From the modeling point of view, a combination of empirical data and judgment from experts can be used to build BNs. This flexibility is a very useful feature that has attracted researchers from diverse fields. In addition, BNs can be represented as network graphs to provide a visualization of the components and dependencies of a subject area. From the knowledge point of view, they represent domain variables in a probabilistic way, allowing inference under uncertainty and making it possible to run a model with missing data. Therefore, due

to the characteristics of BNs as tools for modeling uncertainty, their possible areas of application can widely vary.

Surveys regarding the application of BNs in specific subject areas can be found nowadays (e.g. [13, 14]). However, to the best of our knowledge, the development of BNs throughout multiple domains has not been a subject of research yet. This work aims at providing a general picture about the application of BNs along multiple subject areas, underlining the reasons that have made possible the development of these statistic tools in each domain. While being far from comprehensive, this study is expected to point at current trends in the application of BNs, as well as future potential areas for their development.

## 2 Introduction to Bayesian Networks

A Bayesian network (BN) [16] is a directed acyclic graph (DAG) in which random variables correspond to nodes and edges between nodes are conditional dependencies between the variables. These random variables reflect knowledge uncertainty in a subject area. If the variables are continuous, a common approach is to discretize them by dividing them into intervals. Thus, nodes are partitioned into a set of possible states that are able to represent numerical and non-numerical values.

An arc or edge from node  $X_i$  to node  $X_j$  represents, intuitively, that  $X_i$  has a direct effect or impact on  $X_j$  [19].  $X_i$  is defined as a parent of  $X_j$ ; thus, all those nodes that have arcs directed to a specific node are considered its parents. Each node in the network has a conditional probability distribution associated of the form  $P(X_i|parents(X_i))$ . This is represented as a conditional probability table (CPT), which contains all the parent influences that act upon the variable  $X_j$ . A joint probability distribution (JPD), which is the likelihood of each possible event as defined by the CPTs, can be calculated using the chain rule:

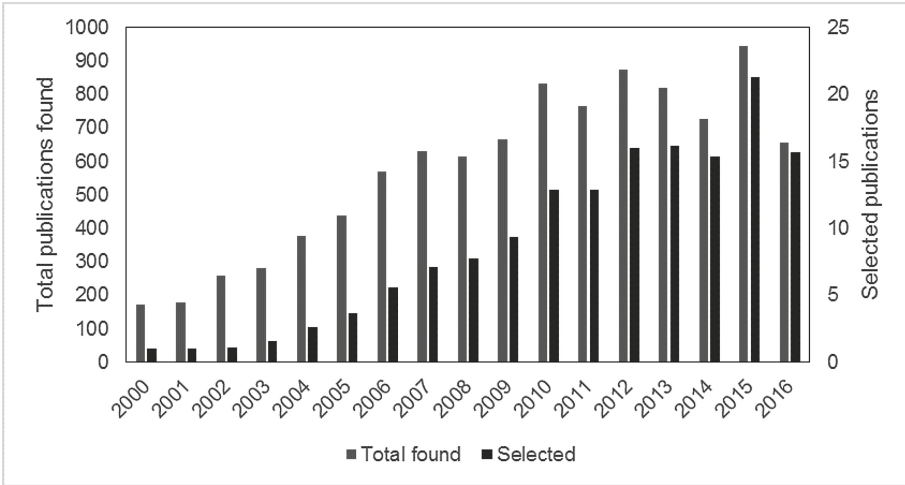
$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i|Parents(X_i))$$

The process of calculation in a BN model is based on the Bayes theorem, where for two uncertain variables  $A$  and  $D$ :

$$P(A|D) = \frac{P(D|A)P(A)}{P(D)}$$

## 3 Subject of Study and Data Structure

A total of 9048 related papers published between the years 2000–2016 were found using the search engine for academic publications Google Scholar [6], filtering results by requiring the appearance of synonyms for the term “Bayesian network” in the title. The distribution of related publications found per year can be visualized in Fig. 1.



**Fig. 1.** Total number of related publications found vs selected number of papers (per year)

Due to limitations regarding time and resources, we randomly selected an initial sample of 422 articles from all the publications found. We excluded merely descriptive papers (without any real-world application of a BN), and thoroughly searched the remaining papers for properly presented BNs, i.e. articles that not only mention the utilization of a BN, but also present a detailed structure of the network (nodes, conditional dependencies, etc.). Considering these restrictions, our sample was reduced to 150 papers concerning real-world applications of BNs to multiple subject areas, distributed according to Fig. 1.

It is important to notice that the list of selected publications resulting of this process cannot be considered to be comprehensive or unbiased. For example, relevant publications may not have been detected using the current search terms, search results are strongly biased toward publications written in English, and books and theses were excluded from this review. Despite these caveats, it is likely that the publications in study provide a representative overview of the different areas of application of BNs.

The main question of the current investigation is, how is a BN related to a domain of research? We tackled this question from two angles. First, after examining each publication in detail, we extracted information concerning the applied BN: (i) the task trying to be solved by the BN, to gather knowledge about the source of uncertainty in the domain; (ii) the kind of variable represented by Bayesian nodes, to determine if the BN implementation requires human interaction; (iii) the stage of development of the BN, to determine if the subject area is relatively new or old; and (iv) the domain of application of the BN.

This last point connects the first approach with the second, which focuses on analyzing the domain of research in search for the answer. For this purpose, we characterized the domains by means of five criteria; namely, the dimension of

the domain, its citation trend, and the levels of formalization, data accessibility and data accuracy. Observed differences in the application of BNs justified the inclusion of an extra characteristic: the level of human intervention. Regarding this analysis, judgment was often required to interpret the information provided in the papers, because many of the methods used were not fully described in the text. Therefore, it should be taken into account that the information presented here is based on an interpretation of what the authors presented.

## 4 Analysis of BN Publication Activity in Different Domains

The publications included in the review were grouped by subject area (domain). This classification was performed intuitively, taking into account factors as keywords in the text, the research institutions involved, and the publication journal. It should be noticed that the domain categories are not rigorously mutually exclusive: for example, artificial intelligence and informatics are considered as separate categories, but they can arguably overlap. Concretely, because of the usefulness of AI in a wide array of subject areas (BNs are considered part of it), we interpret the term AI as encompassing the tools and techniques not covered by other domains. For example, Multi-agent systems, Semantic Web and Conversation analysis, which can be considered as AI sub-domains, are classified separately due to their greater use of BNs in relation to other subject areas.

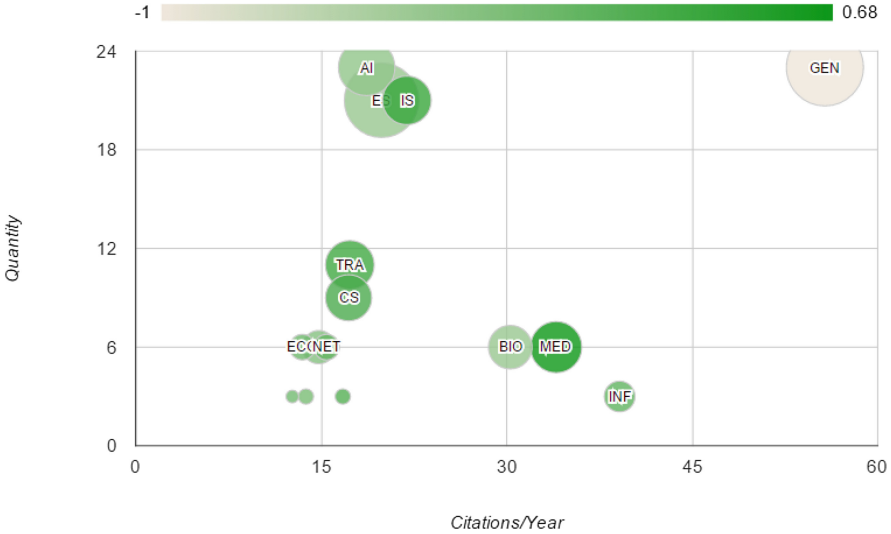
Table 1 shows the quantity of selected papers over 15 domains, along with the dimension of each domain, the averaged citations per year (per publication), and the citation trend as well. Citations per year is an indicator relative to a single paper; it is obtained by dividing the citation number of the paper by how many years passed from publication up until today. The average citations per year, multiplied by the total expected papers for a determined year, results in the total expected citation number in the domain for that period. Forecasts can be made by including the citation trend.

The trend indicator in Table 1 is the slope of the simple regression over the scatter plotting the sum of citations per year from 2000 to 2016. A negative trend indicates citation stagnation, while a positive trend indicates citation growth. The dimension is an indicator of the size of the domain: it is calculated as the area under the simple regression curve where the slope is the citation trend and the axis intersection is the sum of citations, and is afterwards normalized to take values between a minimum of 0 and a maximum of 1. A visual representation of Table 1 is presented in Fig. 2.

We consider that the indicators presented in Table 1 can give us a starting point for the analysis of each subject area in terms of growth, size and influence in the scientific community. However, it should be noted that conclusions based on these indicators must consider an important bias, which we will call compatibility bias. It refers to the presence or absence of likelihood between the characteristics of the subject area and the tasks commonly solved by means of BN. In other words, we cannot jump into conclusions about a domain as a whole, because these

**Table 1.** Publication and citation metrics per domain (QTY=Quantity; DIM=Dimension; CY=Citations/year; CT=Citation trend)

Subject area	SA	QTY	DIM	CY	CT
Artificial Intelligence	AI	23	0.58	18.68	-0.3
Biology	BIO	6	0.34	30.29	-0.35
Computer science	CS	9	0.39	17.22	0.23
Conversation analysis	CA	3	0.08	13.77	-0.11
Economics	ECO	6	0.14	13.48	-0.03
Environmental science	ES	21	0.95	19.88	-0.35
Genetics	GEN	23	1	55.73	-1
Industrial systems	IS	21	0.42	21.93	0.36
Informatics	INF	3	0.2	39.13	0.07
Law	LAW	3	0.08	16.75	0.16
Medicine	MED	6	0.49	34	0.68
Multi-agent systems	MAS	6	0.22	14.81	-0.2
Networks	NET	6	0.15	15.44	0.05
Semantic Web	SW	3	0.07	12.67	-0.07
Transportation	TRA	11	0.42	17.32	0.31



**Fig. 2.** Publication and citation metrics per domain. The color scale corresponds to the citation trend: a light color indicates a negative trend, while a dark color indicates citation growth. The circle diameter corresponds to the domain dimension.

indicators only reveal information about a subject area from the point of view of BN applications. Nevertheless, the scientific status of the current subject areas can be estimated without considering the compatibility bias (for the moment).

Independent from the number of citations, high number of publications may relate to a well-consolidated domain, whereas low number of publications may point to an undeveloped, possibly new scientific domain. High number of citations suggests considerable scientific interest in the field, while low number of citations indicate either lack of scientific interest or the generation of new, previously nonexistent knowledge.

A high dimension and citation trend suggests that the domain is on the rise with constant new research, but at the same time is well-consolidated throughout the last years. This can be seen in the case of artificial intelligence, industrial systems, transport, computer science and medicine. Alternatively, a low dimension and trend, like in biology and economics, points to a stagnation in the domain. A positive citation trend with a low dimension is most probably a new subject area with considerable scientific potential, as in law, informatics and networks.

A different dynamic is noticed in genetics, where the dimension is considerably higher than the citation trend. In this case, two possible explanations can be derived. On the one hand, the domain could already be well-established but research has been halted or closed, possibly by a comprehensive translation of theoretical models to practical appliances. On the other hand, the domain could be recently discovered and the information space is wide but unexplored, giving researchers the opportunity to generate new knowledge from different focal points without the need to rely on citations to previous work.

Besides their inherent probabilistic modeling function, the reviewed BNs in this study serve different tasks or purposes. Figure 3 presents an overview of the domains with the addition of the tasks that are solved by BNs.

It can be noticed in Fig. 3 that the task of learning through BNs is present only in genetics and biology, underlining the knowledge-gathering character of these fields. The tasks of learning, classification and recommendation don't make use of human decision-making, while the rest vary in relation to its utilization.

Different stages of development were observed in the reviewed BNs. Structure learning is present in 15.4% of the reviewed papers, mostly in the field of genetics. This suggests an abundance of primary data in that domain, together with an absence of existing knowledge from where to build the models. Parameter learning is reported in the areas of environmental science, biology, genetics, medicine and artificial intelligence, comprising 13.5% of the papers. Bayesian network models ready up to the inference stage represent the remaining 71.2%, where 44.2% are validated and the remaining 26.9% are not.

## 5 Impact of Domain Properties on the Application of BNs

As we already mentioned, we consider that the indicators already presented do not represent a transparent picture of the "growth" or "progress" of a domain.

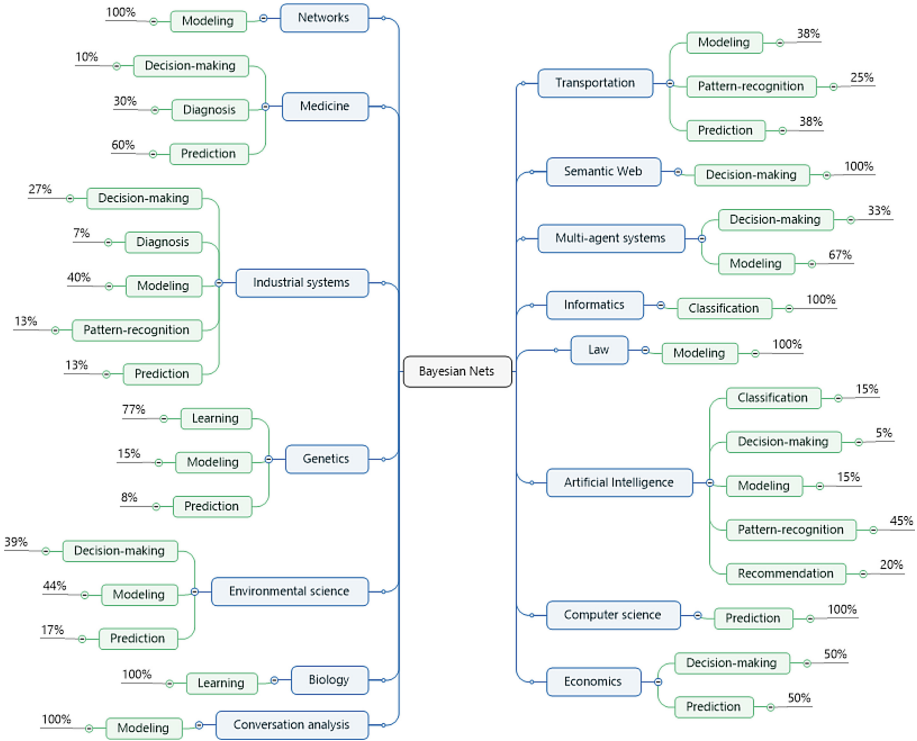


Fig. 3. Tasks addressed by BNs for each subject area

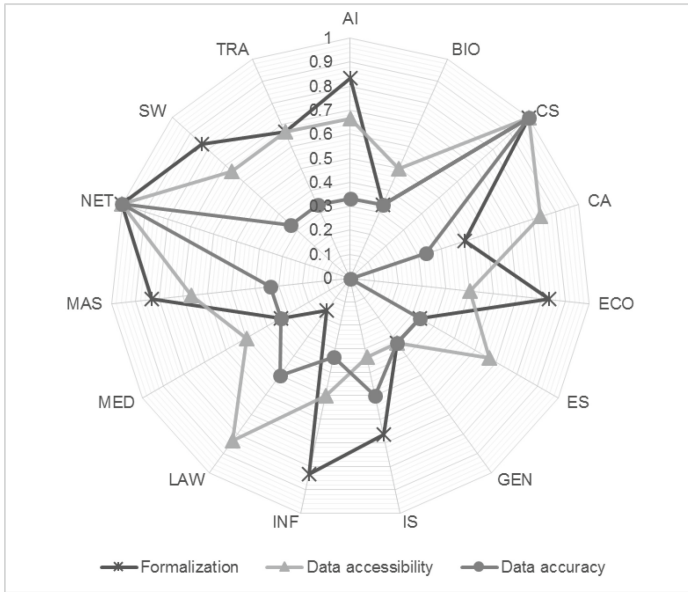
Instead, they portray an inclination to develop BN applications in a determined area of research. This inclination can be described as a compatibility bias, or the likelihood of applying BNs in a domain that facilitates their implementation. In order to analyze this bias, we firstly hypothesized that there are specific attributes of a domain that determine its suitability for BN applications; namely, levels of formalization, data accessibility and data accuracy.

We use the formalization indicator as the level to which sets of symbols, formulas and rules are used to describe objects, events and their interrelationships in the domain. Accordingly, the level of formalization is defined by the presence or absence of standards, mathematical formulas, languages, or any kind of unified coding of the primary data of a domain.

Data accessibility is the level to which it is possible for an independent researcher to obtain data from primary sources in the domain. Monopolization of information in a domain, for example, is considered as a negative factor regarding accessibility to information. Open access to primary data or the need for specialized equipment to obtain it also affect this indicator.

Data accuracy is the unlikelihood of finding contradictions, ambiguity and noise in the information gathered from the domain sources. In a domain with low data accuracy, for example, the researcher knows which information to retrieve

from the primary source but is unable to fully obtain it due to unavoidable error in measurement or interpretation. It represents the gap between the real data and the one gathered by the researcher.



**Fig. 4.** Formalization, data accessibility and data accuracy in different domains

Figure 4 presents an attempt to quantify the mentioned attributes in the domains of the present study. It is necessary to underline that arguments found in the literature were used (when possible) to obtain a qualitative score for these indicators and subjective judgments were needed in the rest of the cases.

### 5.1 Formalization

The areas of computer science and networks are rated with the highest level of formalization because the languages in which they are expressed were created for themselves, not trying to imitate or model behaviors outside their subject area. However, formalization does not mean universality, which means that these fields can be expressed in a variety of languages and symbolizations that can express the same concepts, possibly with mappings between each other.

The next level of formalization comprises fields as artificial intelligence, informatics and industrial systems. Although artificial intelligence shares aspects with computer science, the difficulty in formalizing lies in the attempt to model human intelligence. As an example, [12] raises the issue that no formal theory of common sense can get by without some formalization of context. Informatics is



less formalized than computer science because it involves a broader arrange of subfields, each of which tries to apply established formalizations of computer science to itself. Economics relies heavily in mathematical and statistical formalizations, to the extent that there is influence to diminish this trend [9].

The fields of transportation and industrial systems possess a medium-high level of formalization. Transportation problems in operations research are tightly related to mathematical optimization, and dedicated formalization attempts have been developed [7]. Concerning industrial systems, there is a substantial number of international standards regarding manufacturing and other industrial systems, in which common languages are established among specific disciplines. Formalization efforts of a medium level have been made in the field of conversation analysis [8], however not widespread.

A low-medium level of formalization was assigned to the fields of biology, environmental science, genetics and medicine, as mathematical formalization of the concepts on the processes in living systems represents considerable difficulties [18]. Environmental science is tightly related to public policy, therefore formalization attempts are in early stages [10]. The domain of law is behind the rest of fields with a low formalization level. However, progress has been made in the field of argumentation [3, 17].

## 5.2 Data Accessibility

The areas of computer science and networks are rated once again with the highest level, in this case, of data accessibility. This is because the body of related data is artificial, constantly updated and open for research and application. Both conversation analysis and law share a high level of data availability because audio-visual media, a source of conversational interactions, is widespread and freely available in the Internet, while the body of the law is in the public domain for public access.

Environmental science, artificial intelligence and transportation were rated with a medium-high level of data availability. Environmental data is in the public domain, without commercial restrictions. However, only recently has it become crucial in research with issues like climate change. Artificial intelligence, although is a wide field with considerable quantity of information sources, critical data is kept outside of the public domain, as it is a commercial competitive advantage (e.g. Google, Microsoft). Transportation shares similar characteristics with artificial intelligence, in this respect.

Medicine, biology, economics and informatics share a medium level of data accessibility. In medicine, research results are openly available and constantly scrutinized by governmental authorities. However, primary data is not publicly available in as much as two thirds of the research performed [1]. In biology, independent researchers are still able to gather information possibly without the need for specialized technology. In economics, a considerable percentage of data is publicly available, but it can be argued that its veracity depends on hidden factors (e.g. political). In informatics, knowledge management capabilities are

significantly related with competitive advantage [2], which is a constraint to data openness.

Genetics and industrial systems were rated with a low-medium score on data availability. Information generation in genetics depends on private and governmental funding in specialized laboratories, under specialized research programs. Concerning industrial systems, data is generally not publicly open, especially in industry areas high in competitive advantage.

### 5.3 Data Accuracy

The area of economics possesses a very low level of data accuracy, because there is inherent uncertainty in dealing with out of control, external behavior in a big scale market involving a mass of agents. Industrial systems and law are situated in a medium level of uncertainty. Data accuracy of the law is no less than moderate and it is a much less serious defect in the law than it is often thought to be [11]. In relation to industrial systems, uncertainty is present in the creation, operation and control of industrial processes.

Data accuracy in computer science and networks is virtually the highest, resting on the fact that methods are constantly developed to deal with uncertainty in a fully artificial domain. The rest of subject areas were rated with a low-medium level of data accuracy. In medicine, biology and genetics, there is significant doubt on the prospects for highly deterministic and basically similar mechanisms between individuals [5]. In environmental science, the interconnectedness of complex systems keeps a considerable level of uncertainty. In fact, representing uncertainty in environmental policy is an important subject of study [20].

## 6 Impact of Primary Data Human Intervention on the Application of BNs

During the course of this study, however, we noticed that not only the attributes of a subject area as a whole can determine its suitability for BN applications, but also the characteristics of the primary data source (we consider primary data as the concepts represented by the BN variables and their respective states in each reviewed publication). For this reason we introduce an additional indicator, the human factor, which refers to the level in which human intervention, and conversely, external factors, change or alter the primary data of the domain.

The indicator can take three levels:  $-1$ , meaning that there is no decision-making or human control influencing the domain data;  $0$ , which means that there is a perceptible level of human influence on the data; and  $1$ , when practically all the domain data is product of human actions and interactions. A thorough revision of the Bayesian variables in each selected model was the basis for the human factor scoring (see Table 2). A classification of the subject areas by this factor is presented in Fig. 5.

The areas of artificial intelligence, biology, genetics, informatics and law are devoid of any perceptible human factor in the application of Bayesian networks,

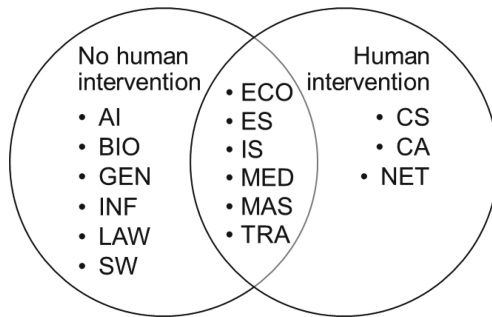
**Table 2.** Distribution of publications (NP) according to Bayesian variables, domain (SA) and human factor (HF)

Bayesian variables	SA	NP	HF
Accident factors	IS	6	0
Accident factors	TRA	3	0
Acoustic-linguistic patterns	AI	2	-1
Audio-visual patterns, previous audio-visual patterns	AI	3	-1
Bankruptcy factors	ECO	3	0
Biological variables	BIO	6	-1
Body part poses, spatial-visual factors	AI	3	-1
Capture factors	ES	2	0
Classes	AI	3	-1
Component wear, cutting parameters	IS	3	1
Concepts	SW	3	-1
Defect proneness, software metrics	CS	3	1
Defect resolution process stages	CS	3	1
Disease progression, treatment	MED	3	0
Economic impact factors	ECO	3	0
Ecosystem health indicators, biochemical properties	ES	2	0
Environmental drivers, ecological responses, decision costs and benefits	ES	3	0
Environmental impact factors	ES	2	0
Environmental viability factors	ES	3	0
Exploits	NET	6	1
Fish population impact factors	ES	3	0
Genes	GEN	20	-1
Hypothesis, Evidence	LAW	3	-1
Machining parameters	IS	3	1
Physical scenario variables	TRA	3	0
Physical symptoms	MED	3	0
Pollution factors	MAS	3	0
Population stressor factors	ES	3	0
Protein features	GEN	3	-1
Risk factor	IS	3	0
Road link vehicle flow	TRA	2	1
Rules, actions	AI	3	-1
Software quality metrics	CS	3	1
Structural, durational, lexical agreement features	CA	3	1

*(continued)*

**Table 2.** (continued)

Bayesian variables	SA	NP	HF
Sustainability factors	ES	3	0
System component status	IS	3	1
System components	IS	3	1
Traffic conditions	TRA	3	1
Trust factors	MAS	3	-1
User/actual experience	INF	3	-1
Users, items, features	AI	3	-1
Vehicle class features	AI	6	-1



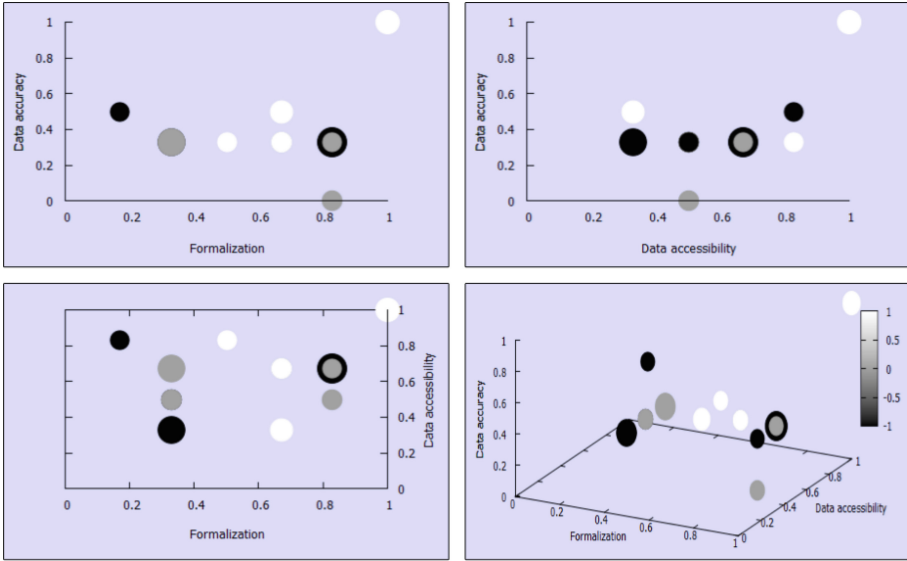
**Fig. 5.** Domain classification by human factor, based on the reviewed publications. The overlap indicates partial human intervention.

with a level of  $-1$ . Arguably, law is a field with an inherent human factor. However, the subfield of argumentation [3] deals with the logic of argumentation instead of the legal confrontations themselves.

Computer science, conversation analysis and networks are fully artificial domains in the samples gathered for the present study, with a human factor level of 1. Software defects [15], network exploits [21] and conversation agreement features [4] depend on the subjective, conscious or unconscious, behavior of humans for the existence of the corresponding sub-domains. The remaining subject areas correspond to a mix of subjective human influence and objective observation of external factors in the domain. For example, in our sample, the highly formalized field of economics presents an undetectable human factor, but a factor of 0 is assigned due to its dependence on human-determined parameters.

## 7 Discussion

Instead of representing different domains, Fig. 6 distributes the totality of publications into the spectrum of data accuracy, accessibility, formalization and



**Fig. 6.** Distribution of the number of publications (size of the spheres) according to formalization, data accuracy, data accessibility and human factor (color scale, dark being negative and white positive).

human factor. It shows that BN applications fully involving human intervention are roughly associated to high levels of formalization, high levels of data accuracy and varied levels of data accessibility. Examples are computer science, conversation analysis and networks. Applications with a negative human factor (not involving human intervention) are related to low levels of data accuracy, varied levels of formalization but at the same time high levels of data accessibility. Examples are artificial intelligence, biology, genetics, informatics, and law. The rest of BN applications have a mixed level of human intervention and are related to low levels of data accuracy, medium levels of data accessibility, and varied levels of formalization.

Domains that present a high citation dimension and a low citation trend, such as environmental science and AI, are considered suitable for BN development and are expected to present a stable utilization of such networks in the near future. On the other hand, it was found that the tasks of learning, classification and recommendation are associated only to a negative human factor. It can be argued that these tasks reflect an early development of the subject area: for example, the domain of genetics is dominated by learning tasks and BN structure learning applications. Thus, a positive citation trend in this and other domains suggests that publications involving human intervention in the primary data can be expected in the near future, along with new decision-making tasks.

## 8 Conclusions

This paper has presented a qualitative and quantitative analysis of the application of BNs in different subject areas. Common indicators such as number of publications reviewed and their citations were used to create more general indicators for each subject area, like the dimension and citation trend of a domain. The purpose of these indicators was not to represent “growth” or “progress” of a domain of research. Instead, they portray a compatibility bias, or an inclination to develop BN applications in a determined area of research, mainly because of the suitability of BNs.

Our strategy to quantify this compatibility bias was to introduce three additional criteria for domain analysis; namely, levels of formalization, data accessibility and data accuracy. The final step was to verify if these three criteria are suitable to explain the trends of BN applications in each field. At this point, we found that also the characteristics of the primary data source must be taken into account. Therefore, an additional factor was introduced into the analysis: the human intervention factor.

The analysis of the four resulting indicators gave us a list of conclusions.

- Full human intervention is associated to high levels of formalization, high levels of data accuracy and varied levels of data accessibility.
- Absence of human intervention is related to low levels of data accuracy, varied levels of formalization but at the same time high levels of data accessibility.
- Mixed level of human intervention are related to low levels of data accuracy, medium levels of data accessibility, and varied levels of formalization.

We expect that domains that comprise the mentioned combinations of formalization, data accessibility, data accuracy and human intervention will be suitable for the development of BN applications in the future. These indicators are meant to facilitate the analysis and development of BN applications. The dimension and citation trends that we presented provide the current trends in developing BNs, but also give possible research opportunities: by applying the present methodology in subject areas not present in this study, new possibilities for BN development can be found.

**Acknowledgment.** The research presented in this paper was supported by the RUDN University Program 5-100.

## References

1. Alsheikh-Ali, A.A., Qureshi, W., Al-Mallah, M.H., Ioannidis, J.P.A.: Public availability of published research data in high-impact journals. *PLoS ONE* **6**(9), e24357 (2011). <https://doi.org/10.1371/journal.pone.0024357>
2. Chuang, S.H.: A resource-based perspective on knowledge management capability and competitive advantage: an empirical investigation. *Expert Syst. Appl.* **27**(3), 459–465 (2004). <https://doi.org/10.1016/j.eswa.2004.05.008>

3. Fenton, N., Neil, M., Lagnado, D.A.: A general structure for legal arguments about evidence using Bayesian networks. *Cogn. Sci.* **37**(1), 61–102 (2012). <https://doi.org/10.1111/cogs.12004>
4. Galley, M., McKeown, K., Hirschberg, J., Shriberg, E.: Identifying agreement and disagreement in conversational speech: Use of Bayesian networks to model pragmatic dependencies, pp. 669–676 (2004)
5. Greenspan, R.J.: Biological indeterminacy. *Sci. Eng. Ethics* **18**(3), 447–452 (2012). <https://doi.org/10.1007/s11948-012-9379-2>
6. Harzing, A., van der Wal, R.: Google scholar as a new source for citation analysis. *Ethics Sci. Environ. Polit.* **8**, 61–73 (2008). <https://doi.org/10.3354/esepe00076>
7. Xu, S.J., Nourinejad, M., Lai, X., Chow, Y. J.: Network learning via multiagent inverse transportation problems. *Transp. Sci.* (2017). <https://doi.org/10.1287/trsc.2017.0805>
8. Kasper, G., Wagner, J.: Conversation analysis in applied linguistics. *Ann. Rev. Appl. Linguist.* **34**, 171–212 (2014). <https://doi.org/10.1017/S0267190514000014>
9. Katzner, D.W.: *Unmeasured Information and the Methodology of Social Scientific Inquiry*, 1st edn. Kluwer Academic Publishers, Boston (2001)
10. van Kerkhoff, L.: Integrated research: concepts of connection in environmental science and policy. *Environ. Sci. Policy* **8**(5), 452–463 (2005). <https://doi.org/10.1016/j.envsci.2005.06.002>
11. Kress, K.: Legal indeterminacy. *Calif. Law Rev.* **77**(2), 283 (1989). <https://doi.org/10.2307/3480606>
12. McCarthy, J.: Generality in artificial intelligence. *Commun. ACM* **30**(12), 1030–1035 (1987). <https://doi.org/10.1145/33447.33448>
13. Mkrtchyan, L., Podofilini, L., Dang, V.: Bayesian belief networks for human reliability analysis: a review of applications and gaps (2015)
14. Newton, A.C.: *Bayesian Belief Networks in Environmental Modelling: A Review of Recent Progress*, 1st edn, pp. 13–50. Nova Science Publishers, New York (2009)
15. Okutan, A., Yildiz, O.T.: Software defect prediction using Bayesian networks. *Empir. Softw. Eng.* **19**(1), 154–181 (2012). <https://doi.org/10.1007/s10664-012-9218-8>
16. Pearl, J.: *Probabilistic Reasoning in Intelligent Systems*, 1st edn. Kaufmann, San Mateo (1988)
17. Prakken, H.: A logical framework for modelling legal argument, pp. 1–9. ACM Press (1993)
18. Rubin, A., Riznichenko, G.Y.: *Mathematical Biophysics*, 1st edn. Springer, Boston (2014). <https://doi.org/10.1007/978-1-4614-8702-9>
19. Russell, S.J., Norvig, P., Davis, E.: *Artificial Intelligence*, 1st edn. Prentice Hall, Upper Saddle River (2010)
20. Shackley, S., Wynne, B.: Representing uncertainty in global climate change science and policy: boundary-ordering devices and authority. *Sci. Technol. Hum. Values* **21**(3), 275–302 (1996). <https://doi.org/10.1177/016224399602100302>
21. Xie, P., Li, J.H., Ou, X., Liu, P., Levy, R.: Using Bayesian networks for cyber security analysis, pp. 211–220. *IEEE/IFIP* (2010)