ARTICLES IN ENGLISH

INFORMATION AND DIGITAL TECHNOLOGIES

COMPUTER SCIENCE, COMPUTING AND CONTROL

Received: 11.10.2026
Accepted: 18.10.2026
Published: 28.10.2026

УДК 004.8+658.3+519.816

DOI: 10.32743/UniTech.2026.151.10.23576

Classification of candidate information uncertainty in AI-based resume screening

Babayev Samir Cahid
PhD student
Azerbaijan State Oil and Industry University
Azerbaijan, Baku
E-mail: samircbabayev@gmail.com

Классификация неопределённости сведений о кандидатах при отборе резюме с использованием искусственного интеллекта

Бабаев Самир Джахид
аспирант
Азербайджанский государственный университет нефти и промышленности
Азербайджанская Республика, г. Баку

Abstract

Large language models and multi-criteria decision methods rank job applicants from resumes, although resume content is written by the applicant and can be verified only in part. The aim of the study is to classify how AI-based resume screening represents the uncertainty of candidate information and to derive research gaps and design requirements from this classification. Five methods closest to pre-interview shortlisting are described and compared on five criteria that follow

Библиографическое описание: Бабаев С.Д. Классификация неопределённости сведений о кандидатах при отборе резюме с использованием искусственного интеллекта // Universum: технические науки : электрон. научн. журн. 2026. 10(151). URL: https://7universum.com/ru/tech/archive/item/23576

For citation: Babayev S.C. Classification of candidate information uncertainty in AI-based resume screening // Universum: Technical Sciences : electronic scientific journal. 2026. No. 10(151). Available at: https://7universum.com/ru/tech/archive/item/23576

the path from a resume claim to a rank: the form of the criterion score, the representation of uncertainty, the treatment of claim reliability, the ranking procedure and the possibility of asking the candidate. The proposed classification crosses two types of uncertainty, the imprecision of a value and the reliability of a claim, with three sources, the candidate, the language model that extracts the data and the evaluator. In the compared methods reliability is either absent or expresses the evaluator's confidence, and language-model pipelines accept resume content as given, so two of the six cells of the classification have no modelling device. Four gaps are stated as research questions, and provisions of the EU AI Act on high-risk recruitment systems are translated into verifiable design requirements. The classification and the requirements can serve as a checklist for selecting and auditing screening tools.

Аннотация

Большие языковые модели и методы многокритериального принятия решений применяются для ранжирования соискателей по резюме, хотя содержание резюме формирует сам соискатель и проверить его удаётся лишь частично. Цель исследования состоит в том, чтобы классифицировать способы представления неопределённости сведений о кандидате в системах отбора резюме на основе искусственного интеллекта и вывести из этой классификации исследовательские пробелы и требования к проектированию. Описаны пять методов, наиболее близких к формированию списка кандидатов на собеседование, и проведено их сопоставление по пяти критериям, отражающим путь от утверждения в резюме до места в рейтинге: форма оценки по критерию, представление неопределённости, учёт достоверности утверждений, процедура ранжирования и возможность запросить сведения у кандидата. Предложена классификация, в которой два типа неопределённости, неточность значения и недостоверность утверждения, сопоставлены с тремя источниками, а именно кандидатом, языковой моделью, извлекающей данные, и экспертом. В рассмотренных методах достоверность либо не учитывается, либо выражает уверенность эксперта, а системы на основе языковых моделей принимают содержание резюме без проверки, поэтому для двух из шести ячеек классификации средства моделирования отсутствуют. Четыре выявленных пробела сформулированы как исследовательские вопросы, а положения Регламента ЕС об искусственном интеллекте о системах найма высокого риска переведены в проверяемые требования к проектированию. Классификация и требования могут служить контрольным перечнем при выборе и аудите систем отбора резюме.

Keywords: resume screening, candidate information reliability, large language models, multi-criteria decision making, fuzzy numbers, Z-numbers, EU AI Act.

Ключевые слова: отбор резюме, достоверность сведений о кандидате, большие языковые модели, многокритериальное принятие решений, нечёткие числа, Z-числа, Регламент ЕС об искусственном интеллекте.

1. Introduction

Resume content is written by the applicant, and respondents admit to some degree of exaggeration (72 %), omission (61 %) and fabrication (31 %) in resumes, and organisations verify only selected items [4]. The same study shows that deception in resumes is not a rare anomaly but a measurable behaviour that differs between applicants [4]. A screening method that treats every statement in a resume as a fact therefore ranks claims rather than verified qualifications.

Multi-criteria decision making has become the standard tool for ranking candidates, and TOPSIS and its variants are used in half of the personnel-selection studies covered by a recent systematic review [8]. In this line of work the uncertainty of a rating is usually modelled with fuzzy numbers. Alguliyev et al. convert linguistic expert ratings into triangular fuzzy numbers and rank candidates with a modified fuzzy VIKOR method [1], and Öztürk and Kaya apply fuzzy VIKOR to personnel selection in an automotive supplier [9]. In both cases the fuzziness describes how precisely the expert can express a judgement, while the information on which the judgement rests is taken as correct. Existing reviews of AI in recruitment address its ethics and its use in employee recruitment [2, 6]. They discuss fairness, transparency and acceptance by applicants, but do not ask how screening methods represent the possibility that a claim in the resume is not true.

Zadeh introduced Z-numbers as a pair of a fuzzy restriction and a measure of its reliability [15], which in principle offers a device for the second kind of uncertainty. Gottwald et al. use Z-numbers in personnel selection, but the reliability component there expresses how confident the evaluator is in the rating [3], not whether the candidate's statement is true. Language models shift the problem to an earlier stage. Hoque et al. let a language model classify profile text into classes that are mapped to fixed fuzzy numbers and ranked by fuzzy TOPSIS [5], so the content of the profile enters the ranking without any check. Stuart et al. take the opposite direction and build an interview system that keeps a belief over rubric levels and asks the candidate for missing information [13], yet this system does not rank candidates. The human reviewer is not a reliable safeguard either, since in an experiment people followed the preferences of a biased model in up to 90 % of decisions [14].

The legal context makes the question practical. Annex III, point 4(a), of the AI Act classifies systems intended to analyse and filter job applications and to evaluate candidates as high-risk, and the Act attaches obligations on data, logging, transparency, human oversight and robustness to such systems [11, 12]. Article 22 of the GDPR and the Azerbaijani law on personal data add protection against decisions based solely on automated processing [7, 10]. Table 1 shows that these obligations can be stated as concrete requirements for a pre-interview screening tool, and several of them concern exactly how the tool handles claims and their evidence.

Table 1. Provisions of the AI Act and verifiable design requirements for a pre-interview screening tool

Provision	Design requirement
Art. 10(2)(f)–(g), bias in data	An annotated test set of known origin is used for name and attribute substitution tests after each model or prompt change
Art. 12(1)–(2), 26(6), logs	Each extracted claim, its evidence, the model version, each rank and each recruiter action are logged and kept for at least six months
Art. 13(1), 13(3)(b), transparency	The interface shows each claim with its evidence and score uncertainty, and the instructions state the measured extraction accuracy

Provision	Design requirement
Art. 14(4)(b), (d), (e), oversight	The recruiter can override any rank with a logged reason, and no candidate is rejected without a human action
Art. 15(1), (5), robustness	Resumes with hidden prompts are tested before deployment, and extracted evidence must appear in the visible text

Taken together, the sources show three separate lines of work, fuzzy multi-criteria ranking, language-model screening and regulation, none of which treats the truth of a resume claim as a quantity to be modelled. The aim of the study is to classify how AI-based resume screening represents the uncertainty of candidate information and to derive research gaps and design requirements from this classification.

2. Materials and Methods

The analysis draws on the studies of personnel selection, multi-criteria ranking and AI-based recruitment listed in the references. No systematic database search was carried out, so the comparison is narrative rather than systematic. From these studies, five methods closest to the formation of an interview shortlist were selected, that is, methods that turn information about each applicant into a score and either rank the applicants or collect information for ranking them. Regulation (EU) 2024/1689 and its amendment by Regulation (EU) 2026/1744 were read in the Official Journal text on EUR-Lex on 10 October 2026 [11, 12].

The modified fuzzy VIKOR method of Alguliyev et al. [1] expresses expert linguistic ratings of candidates as triangular fuzzy numbers and selects the candidate that is closest to the ideal in the sense of a compromise between group utility and the largest individual shortfall. Öztürk and Kaya [9] solve the same selection problem with fuzzy VIKOR on expert ratings of candidates in the automotive supply industry. Gottwald et al. [3] rate candidates with Z-numbers, in which each fuzzy rating is accompanied by a fuzzy reliability given by the evaluator, and rank them with fuzzy AROMAN. Hoque et al. [5] use a language model to read candidate profiles and assign them to classes, map the classes to fixed fuzzy numbers and rank the candidates with fuzzy TOPSIS. Stuart et al. [13] do not rank candidates but conduct an automatic interview that keeps a probability distribution over rubric levels and chooses questions that elicit the missing information.

The purpose of the comparison is narrower than the aim of the study. It is to establish, for each method, where the uncertainty of candidate information enters the selection procedure and whether the truth of a claim can be represented at all. The five criteria in Table 2 follow the path of a claim from the resume to the shortlist. The criterion score shows what the method actually evaluates, an expert rating, a model output or a belief. The uncertainty representation shows which formal device carries imprecision. The reliability of claims is the property under study and shows whether doubt about a statement is represented and whose doubt it is. The ranking procedure shows whether the method produces a shortlist, and the last criterion shows whether the method can reduce uncertainty by asking the candidate. Together the criteria are sufficient to place each method in the classification of Section 3.2.

3. Results and Discussion

3.1. Comparison of the closest methods

Table 2. Comparison of methods closest to pre-interview shortlisting

Work	Criterion score	Uncertainty	Reliability of claims	Ranking	Asks the candidate
Alguliyev [1]	Expert linguistic ratings	Triangular fuzzy numbers	Not considered	Modified fuzzy VIKOR	No
Öztürk [9]	Expert ratings	Fuzzy numbers	Not considered	Fuzzy VIKOR	No
Gottwald [3]	Expert ratings with reliability	Z-numbers	Evaluator's confidence	Fuzzy AROMAN	No
Hoque [5]	Language-model classes	Fixed fuzzy numbers	Content accepted as given	Fuzzy TOPSIS	No
Stuart [13]	Belief over rubric levels	Probabilities over levels	Honesty not modelled	No	Yes

The Z-number method is the only one with an explicit reliability component, but it expresses the hesitation of the person who rates [3]. The language-model hybrid takes the profile text as given [5]. Table 2 thus shows that the five methods differ in the device used for imprecision, but agree in that none of them represents the possibility that the candidate's claim is untrue. The only method that requests information from the candidate does not rank, and the ranking methods do not request information.

3.2. Classification by type and source of uncertainty

Fig. 1 crosses two types of uncertainty, the approximate knowledge of a value (imprecision) and the doubt whether a claim is true (reliability), with three sources.

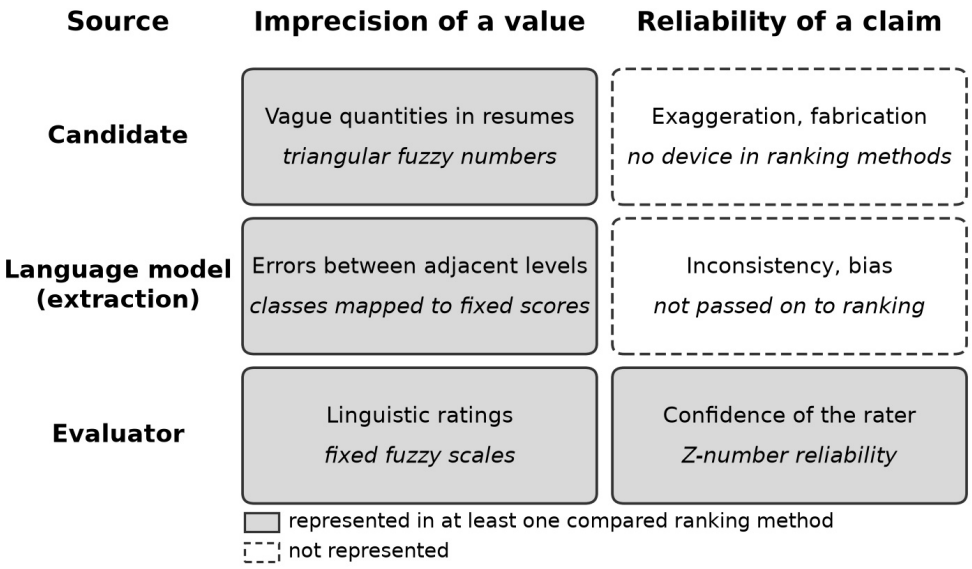


Figure 1. Two types and three sources of uncertainty in candidate information

Most ranking methods place uncertainty in the evaluator row and treat it as imprecision. No compared ranking method represents whether a claim is true, although Z-numbers provide a device for it [15]. The uncertainty of language-model extraction is not passed on to the ranking.

Figure 1 is built from the columns of Table 2. The column of uncertainty representation fills the imprecision cells. Fuzzy VIKOR [1, 9] and the Z-number method [3] place it in the evaluator row through linguistic ratings on fixed fuzzy scales, and the language-model hybrid [5] places it in the extraction row through classes mapped to fixed scores. The candidate cell of the imprecision column can in principle be covered by the same fuzzy numbers, but in the compared methods they express the evaluator's judgement rather than the imprecise value that the candidate states, so this cell also remains open in practice. The column of claim reliability fills the reliability cells. Only the Z-number method has an entry there, and it belongs to the evaluator row, while the entries “not considered”, “content accepted as given” and “honesty not modelled” leave the candidate and extraction cells empty. The main finding of the comparison is therefore that the cell of the candidate's reliability, which corresponds to the exaggeration and fabrication reported in [4], has no device in any compared ranking method.

3.3. Research gaps and questions

The first gap is that no compared method measures how far a claim is supported by evidence in the resume [3, 5, 13]. RQ1 asks how this support can be measured and turned into a fuzzy or probabilistic score. The second gap is that fuzzy parameters come from fixed linguistic tables or expert judgement and do not differ between candidates [3, 5]. RQ2 asks whether candidate-specific parameters derived from claims and evidence rank candidates more accurately.

The third gap is that the ranking methods do not request additional information, while the system that requests it does not rank [1, 3, 5, 9, 13]. RQ3 asks which information should be requested from which candidate so that verification changes the interview shortlist most. The fourth gap is that the methods were tested on 100 profiles, four candidates or simulated candidates [3, 5, 13]. RQ4 asks how the methods perform on real resumes with annotated claims and evidence.

3.4. From legal provisions to design requirements

As noted in Section 1, resume screening systems are high-risk under Annex III, point 4(a), of the AI Act [11]. Regulation (EU) 2026/1744 amends Article 10, leaves Articles 12–15 and 26 unchanged and moves the date of application for such systems to 2 December 2027 [12]. The five requirements in Table 1 (Section 1) follow most directly from the classification, because each of them concerns a claim, its evidence or the uncertainty of its score.

The requirements specify what the recruiter sees and can change, both because Article 22 of the GDPR and Article 7.3 of the Azerbaijani law on personal data protect the data subject against decisions based solely on automated processing [7, 10] and because people followed the preferences of a biased model in up to 90 % of decisions in an experiment [14].

4. Conclusion

The study classified the uncertainty of candidate information in AI-based resume screening by two types, the imprecision of a value and the reliability of a claim, and by three sources, the candidate, the language model and the evaluator. The comparison of five methods closest to

shortlisting showed that the imprecision of values is covered by fuzzy devices, whereas the reliability of what the candidate writes has no modelling device in the compared ranking methods, and the uncertainty of language-model extraction is not passed on to the ranking. Four research questions follow from these gaps, on the support of claims by evidence, candidate-specific fuzzy parameters, the choice of information to verify and testing on real annotated resumes. The AI Act provisions on high-risk recruitment systems were translated into five verifiable requirements that link the classification to the design and audit of screening tools. The comparison is narrative and was coded by one person, so the gaps may not hold for every published method.

References

1. Alguliyev R.M., Aliguliyev R.M., Mahmudova R.S. Multicriteria personnel selection by the modified fuzzy VIKOR method // The Scientific World Journal. – 2015. – Vol. 2015. – Art. 612767. – DOI: 10.1155/2015/612767.
2. Dadaboyev S.M.U., Abdullayeva J., Abbosova N., Suleymenova A., Mamadjanova K. Role of artificial intelligence in employee recruitment: systematic review and future research directions // Discover Global Society. – 2025. – Vol. 3. – DOI: 10.1007/s44282-025-00246-w.
3. Gottwald D. Chocholáč J., Kayacı Çodur M., Čubranić-Dobrodolac M., Yazir K. Z-numbers-based MCDM approach for personnel selection at institutions of higher education for transportation // Mathematics. – 2024. – Vol. 12, No. 4. – Art. 523. – DOI: 10.3390/math12040523.
4. Henle C.A., Dineen B.R., Duffy M.K. Assessing intentional resume deception: Development and nomological network of a resume fraud measure // Journal of Business and Psychology. – 2019. – Vol. 34, No. 1. – P. 87–106. – DOI: 10.1007/s10869-017-9527-4.
5. Hoque S., Karim A.A.J., Alam M.G.R., Gope N. When LLM meets Fuzzy-TOPSIS for personnel selection through automated profile analysis // arXiv preprint arXiv:2601.22433. – 2026. – DOI: 10.48550/arXiv.2601.22433.
6. Hunkenschroer A.L., Luetge C. Ethics of AI-enabled recruiting and selection: A review and research agenda // Journal of Business Ethics. – 2022. – Vol. 178, No. 4. – P. 977–1007. – DOI: 10.1007/s10551-022-05049-6.
7. Law of the Republic of Azerbaijan “On personal data” No. 998-IIIQ of 11 May 2010 (In Azerbaijani).
8. López-González A., Babiloni E., Canós-Darós L. Multi-criteria decision-making methods applied to personnel selection: A systematic literature review // Universidad y Empresa. – 2026. – Vol. 28, No. 50. – P. 1–40. – DOI: 10.12804/revistas.urosario.edu.co/empresa/a.14895.
9. Öztürk F., Kaya G.K. Bulanık VIKOR ile personel seçimi: Otomotiv yan sanayinde uygulama [Personnel selection with fuzzy VIKOR: An application in the automotive supply industry] // Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji. – 2020. – Vol. 8, No. 1. – P. 94–108. – DOI: 10.29109/gujsc.595288.
10. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation) // Official Journal of the European Union. – 2016. – P. 1–88. – URL: http://data.europa.eu/eli/reg/2016/679/oj (accessed: 10.10.2026).

11. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 (Artificial Intelligence Act) // Official Journal of the European Union. – 2024. – No. 12 [Электронный ресурс]. – URL: http://data.europa.eu/eli/reg/2024/1689/oj (accessed: 10.10.2026).
12. Regulation (EU) 2026/1744 of the European Parliament and of the Council of 8 July 2026 amending Regulation (EU) 2024/1689 and other acts (Digital Omnibus on AI) [Электронный ресурс]. // Official Journal of the European Union. – URL: http://data.europa.eu/eli/reg/2026/1744/oj (accessed: 10.10.2026).
13. Stuart H., Kaneko M., Baldwin T. Beyond the Resumé: A rubric-aware automatic interview system for information elicitation // arXiv preprint arXiv:2603.01775. – 2026. – DOI: 10.48550/arXiv.2603.01775.
14. Wilson K., Sim M., Gueorguieva A.-M., Caliskan A. No thoughts just AI: Biased LLM hiring recommendations alter human decision making and limit human autonomy // Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society. – 2025. – Vol. 8, No. 3. – P. 2692–2704. – DOI: 10.1609/aies.v8i3.36749.
15. Zadeh L.A. A note on Z-numbers // Information Sciences. – 2011. – Vol. 181, No. 14. – P. 2923–2932. – DOI: 10.1016/j.ins.2011.02.022.