Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

LLM Adoption in Business Analysis: A Survey-Based Study of Ukrainian and Polish IT Practitioners

Denys Gobov

d.gobov@kpi.ua

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”

Aleksander Jarzębowicz

Gdańsk University of Technology

Łukasz Radliński

West Pomeranian University of Technology in Szczecin

Sylwia Kopczyńska

Poznań University of Technology

Research Article

Keywords: Large Language Models, Business Analysis, Requirements Engineering, AI Adoption, Survey Study, Generative AI, Software Projects

Posted Date: November 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11708263/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

LLM Adoption in Business Analysis: A Survey-Based Study of Ukrainian and Polish IT Practitioners

Author Names[1], Author Names[2], Author Names[3], Author Names[4]

[1*]Organization, Organization Address.
[2*]Organization, Organization Address.
[3*]Organization, Organization Address.
[4*]Organization, Organization Address.

Contributing authors: author@email; author@email; author@email; author@email;

Abstract

Background: Large Language Models (LLMs) are being rapidly adopted for software engineering tasks. However, little empirical data exists on their adoption patterns in business analysis (BA) practice.

Objective: This study quantitatively assesses the adoption of LLMs and their application to specific BA tasks. It also explores practitioners' concerns and attitudes toward artificial intelligence, as well as the level of organizational readiness.

Method: We conducted a cross-sectional survey of 469 BA professionals in Ukraine ($n = 375$) and Poland ($n = 94$). We analyzed the data using descriptive statistics and associative analyses. We also performed multivariable logistic and Poisson regression analyses.

Results: Adoption of LLMs is high among the surveyed, community-engaged practitioners: 80.8% of the 469 respondents actively use LLMs (UA 82.4%, PL 74.5%). Among the 421 respondents with LLM experience, however, only 54.9% report strong management support and 21.1% have access to company-deployed AI solutions tailored for BA. ChatGPT dominates among respondents with LLM experience. The tasks most often selected for LLM use are drafting requirements and searching for knowledge; the primary concern is accuracy and hallucinations. Screening 450 pairs of contextual factors and adoption outcomes per country with false-discovery-rate control identified 25 significant associations in Ukraine (17 of them also after correction across all 450 pairs) and one in Poland. In both countries, broader BA involvement is associated with more LLM-supported tasks.

Conclusions: LLM adoption in BA is widespread, yet it is distributed unevenly across organizational contexts. Measured contextual factors explain little of the variation in adoption; the few associations found are small and should be treated as hypotheses for confirmatory research. A gap between BA practitioners' individual readiness and organizational support is visible.

Keywords: Large Language Models, Business Analysis, Requirements Engineering, AI Adoption, Survey Study, Generative AI, Software Projects

1 Introduction

1.1 Motivation and Relevance

Since ChatGPT's public release in late 2022, Generative AI solutions, especially Large Language Models (LLMs), have rapidly spread among various business sectors and professions [1, 2]. The ability to generate human-like content, the universal rather than task-specific capabilities, and the ease of use through assistants and chat-like communication sparked much wider adoption than the earlier waves of solutions produced by AI research initiatives [3, 4]. Consequently, LLMs are being integrated into workflows in fields such as software development, legal analysis, medical decision support, documentation processing, financial advising, and customer relationship management, among others [1, 2, 5, 6].

The IT industry is no exception. In recent years, reported applications of LLMs in software engineering have included generating source code, suggesting design and architectural patterns, generating and optimizing test cases and scenarios, supporting code reviews and refactoring, and facilitating project management activities, among others [7–11]. As LLMs reshape industrial practices, scholarly work on LLM applications is growing rapidly. Such work investigates the benefits, challenges, and gaps that future research should address [12, 13]. This article focuses on the upstream activities of software projects, including understanding business and stakeholder needs; eliciting, processing, and specifying requirements; and managing changes to these requirements throughout the project lifecycle. These activities are subjects of requirements engineering (RE), which focuses on the requirements of the intended solution, and the broader discipline of business analysis (BA), which begins by identifying business problems or opportunities and facilitates the selection of the right solution that delivers value to stakeholders.

In this article, we investigate BA activities and the opportunities for using LLM support to complete them. BA practitioners engage in needs analysis, requirements elicitation, analysis, specification and validation, stakeholder communication, and solution evaluation – activities that primarily involve natural language artifacts, such as software requirements specifications, user stories, use cases, meeting summaries, and regulatory texts [14]. These activities are a natural fit for LLM-based assistance because LLMs' core competency is generating, transforming, summarizing, and reasoning over natural language. Therefore, such assistance directly aligns with the everyday tasks of business analysts. It is not surprising that scientific works have

already investigated the use of LLMs in requirements elicitation (e.g., generating stakeholder interview scripts [15], retrieval of information providing input for requirements [16]), analysis (e.g., requirements classification [17], prioritization [18]), specification (e.g., transforming natural text into structured requirements [19], automated UML diagrams creation [20]), validation [21], and management [22].

However, the existing scientific literature on LLM adoption in RE and BA still exhibits a gap regarding the state of practice in LLM usage. Most of the available research studies focus on technical evaluations, such as assessing how well LLMs classify requirements, detect ambiguities, or generate user stories [23, 24]. These studies often involve industry practitioners performing LLM-assisted tasks or evaluating LLM outputs (e.g., [25]). However, a practitioner's real-life perspective remains noticeably absent: how BA professionals use LLM tools in their daily practice, the tasks to which they apply them, their concerns, and how organizational contexts impact adoption patterns. Many surveys dedicated to RE and BA practices have been published, but they predate the widespread adoption of LLMs and do not address this issue. More recent, potentially related studies have limited scope, focusing on specific activities (e.g., requirements prioritization [26]) or drawing conclusions from a small number of respondents (e.g., [27]). The only similar survey studies to ours that cover a wide range of RE/BA tasks and substantial numbers of responses were published by Rani et al. [28] and Hess et al. [29]. These studies both involved different populations from ours.

We focus on Ukraine and Poland because they offer particularly informative contexts for studying LLM adoption in BA. Both countries are major destinations for IT outsourcing and nearshoring, with large BA communities that work on projects for clients across Europe and the United States. Following the onset of the pandemic, both countries experienced rapid digital acceleration, and Russia's full-scale invasion of Ukraine in 2022 further accelerated distributed work practices and the adoption of digital tools. However, the two countries occupy different positions along the EU integration spectrum. Poland is a full EU member, subject to the GDPR [30] and the AI Act [31], now being phased in, while Ukraine is an EU candidate country. This creates a setting in which differences in regulatory context could be expected. We report whether the two samples differ without testing regulation as a cause.

1.2 Research Questions

To address the identified gap, we formulated four research questions that progressively shift from describing the current state of adoption to identifying its contextual determinants:

- **RQ1:** What is the current level and pattern of LLM adoption among BA practitioners in Ukraine and Poland?
- **RQ2:** Which BA tasks are LLMs used for most and least, and what concerns about their use do practitioners report?
- **RQ3:** What attitudes do BA practitioners hold toward AI's role in their profession, and how ready are their organizations for AI adoption?
- **RQ4:** Which project and organizational context factors are associated with variations in LLM adoption and usage patterns?

RQ1 and RQ2 describe the extent to which LLMs are used in BA, the tasks for which they are used, and the concerns associated with their use. RQ3 explores practitioners' attitudes toward AI and the readiness of their organizations to adopt it. RQ4 examines which project and organizational factors are associated with variations in the selection and use of these tools.

1.3 Contributions

This study offers the following contributions:

1. **Empirical baseline of LLM use in BA practice (RQ1–RQ2).** Based on 469 practitioners from Ukraine and Poland, the largest sample among the peer-reviewed RE/BA practitioner surveys on LLM use that we identified (Table 1), we show that LLM use is widespread among community-engaged BA practitioners (80.8% active users) and dominated by general-purpose chat assistants. Use concentrates on text-centric tasks such as requirements drafting and knowledge search. This profile partly mirrors the activities practitioners are most involved in, not only LLM capabilities.
2. **Concerns coexist with use (RQ2).** Accuracy and hallucinations are the dominant concerns (82.7%); we found no evidence that they differ across levels of use. Practitioners thus continue to use LLMs while remaining skeptical of their output, which points to verification as a key practice to support.
3. **Individual vs. organizational readiness (RQ3).** Willingness to learn AI tools is near-universal but often conditional on usefulness. BA-specific organizational AI infrastructure is rare (21.1%), and most organizations provide generic tools at best.
4. **Exploratory analysis of context (RQ4).** Screening 450 context–outcome pairs per country with false-discovery-rate control shows that measured context explains little of the variation in adoption. Associations are few and small, and only one, BA activity breadth with the number of LLM-supported tasks, appears in both countries. We report these associations as hypotheses for confirmatory research.
5. **Open research artifacts.** We provide a trilingual (EN/UA/PL) instrument aligned with BABOK, the anonymized dataset, and results for all tested pairs, enabling replication in other countries and longitudinal follow-up.

1.4 Structure

The remainder of this article is organized as follows. Section 2 reviews related work and identifies research gaps. Section 3 outlines the survey design and analytical methods. Section 4 presents the findings across the research questions, while Section 5 interprets them in light of prior research. Finally, Section 6 summarizes the findings, discusses practical implications, and outlines directions for future research.

2 Related Work

This section reviews relevant research directions and situates our research within this context.

2.1 LLMs in Requirements Engineering and Business Analysis

Using the established RE domain decomposition [32, 33] into requirements elicitation, analysis, specification, validation, and management, it is apparent that for each of these areas, attempts to use LLMs as support have been reported.

Requirements elicitation: LLMs can be used to extract information from contracts and other documents that provides input necessary to perform elicitation tasks, e.g., [16, 34]. As Crowd-based Requirements Engineering (CBRE) becomes more widespread, LLMs have become instrumental in processing user reviews, social media posts, online discussions, and similar sources to identify requirements [35]. Another possible use case is to treat LLMs as additional stakeholders and elicit candidate requirements from them that might otherwise be missed [36, 37]. LLMs can also help prepare interview-based elicitation sessions [15] or be prompted to conduct interviews as autonomous agents [38].

Requirements analysis: As LLMs are particularly suited to processing natural language artifacts, this is probably the most explored RE area in the context of LLM applications. Unstructured results of elicitation sessions can be transformed into User Stories [39] or structured, template-based requirements [19]. Also, the elicited requirements can be classified in an automated way into, e.g., functional and non-functional ones [17, 40] or assigned to several quality characteristics [40, 41]. Requirements verification with respect to various quality criteria such as completeness [42], consistency [23], unambiguity [43], or multiple criteria [41, 44] can be conducted with LLMs assistance.

Requirements specification: The abilities of LLMs to transform input artifacts into different forms can be utilized in requirements specification, which aims to express requirements using a selected notation, technique, or structure. Examples of LLMs applications include generating Use Case descriptions [45] and high-quality user stories that conform to the INVEST criteria [46]. As requirements are often expressed using models, other published research studies deal with sequence diagrams [47], multiple UML diagrams [20], goal-oriented models [48], and even specifications that use formal methods (temporal logic) [49].

Requirements validation: It seems that validation (understood as confirming that specified requirements adequately express stakeholders' needs) is a less-explored application area for LLMs, as such tasks, by definition, must involve stakeholders. Some sources (e.g., [50]) classify the studies dedicated to requirements verification and quality assessment as requirements validation, but in our review of related work, we have already discussed them as part of requirements analysis. Still, some attempts were made. For example, LLMs can generate questions that facilitate better communication between stakeholders and development teams [51]. Another example is a study where original requirements issued by human stakeholders were later revised by an LLM, and the revised versions were positively evaluated by stakeholders with respect to (among others) the alignment with their intent [52].

Requirements management: This area is related to maintaining the set of all requirements and their traceability, as well as handling changes to requirements, including change impact assessment and propagating the necessary modifications to

all affected interlinked requirements. Up to now, such tasks were covered by the functionality of requirements management tools such as DOORS or Polarion, but currently new ideas of applying LLMs have emerged, for example, to establish traceability between requirements on various abstraction levels [53, 54] and to assist in processing changes in regulatory acts affecting compliance requirements [22].

Various LLM applications to RE tasks are summarized in the secondary studies, e.g., [50, 55, 56]. What is interesting is that we found no secondary studies covering BA tasks not included in the RE scope, such as strategy analysis, needs assessment, and solution evaluation (see [14]). As for primary studies, they are rather scarce, but some examples include LLMs applications to business process identification and re-engineering [57] or to strategy management [58].

Despite this growing body of research on LLMs applications to RE and BA, a critical perspective remains largely absent: the practitioner's viewpoint. Existing studies predominantly evaluate what LLMs can do with requirements artifacts under controlled conditions, rather than investigating how practitioners actually use LLMs in their daily work. Some studies involve practitioners who perform tasks with the assistance of LLMs or evaluate LLM output, but these studies remain embedded in experimental settings. The distinction is significant – such studies say little about practitioners' use of LLMs in real-world projects and even less about adoption patterns, which are shaped by practitioners' attitudes as well as by organizational context and professional norms.

2.2 AI Adoption Studies in Adjacent Disciplines

The adoption of new technological solutions is a widely explored topic, especially in the Information Systems domain, where frameworks such as the Technology Acceptance Model (TAM) [59] and its extensions (e.g., UTAUT [60]) provide theoretical lenses for understanding the factors that condition such adoption. These frameworks have since been applied to Generative AI, an emerging technology used in diverse socio-technical contexts. Also, frameworks dedicated to AI adoption, such as the Human-AI Collaboration and Adaptation Framework (HACAF) [3], have recently started to appear.

Such frameworks have been used in the context of Generative AI adoption in IT industry software development activities. Simaremare and Edison [61] used HACAF to investigate adoption factors from the perspective of a general software practitioner. Ge and Wu [62] interviewed 50 professional developers and stakeholders to explore their attitudes and perceptions towards adopting ChatGPT for bug fixing. They applied the UTAUT theory to identify the key factors influencing the system's adoption. Xia and Chen [63] focused on the perspective of software development teams specializing in new product development (NPD) and identified the factors driving the use of and attitudes toward GenAI in NPD tasks, based on the UTAUT. Lambiase et al. [64] used the UTAUT2 framework to explore factors influencing the adoption of LLMs in software development, with a focus on the role of professionals' cultural values. Neumann et al. [65] applied the Technology-Organization-Environment (TOE) framework to learn how agile practitioners adopt GenAI tools in real-world organizational contexts, focusing on regulatory conditions, role-specific use cases, benefits, and barriers.

Stray et al. [66] used HACAF to examine how compatibility with existing workflows shapes the adoption of GitHub Copilot and ChatGPT among software developers in a large-scale public-sector organization, drawing on 13 interviews and a survey of 114 respondents. Shao and Ishengoma [67] applied UTAUT to compare determinants of GenAI tool adoption between software engineering students and professional developers, finding social influence to be the strongest predictor of adoption intention.

All such studies, however, focus on software development or its selected aspects; to the best of our knowledge, no equivalent study of the adoption of Generative AI/LLMs for RE and BA tasks among BA practitioners is available.

2.3 Survey Studies in Requirements Engineering and Business Analysis

Surveys involving practitioners and investigating the practical aspects of RE and BA in industrial projects are an integral part of the research landscape in these domains. The most widespread method is a questionnaire-based survey, but other methods, such as interviews or focus groups, are also used. One of the main objectives of many surveys is to examine the contemporary RE problems and challenges encountered by practitioners. Hall et al. [68] surveyed the employees of 12 (UK and multinational) IT companies to identify and classify requirement problems reported by them. Solemon et al. [69] investigated the RE problems in the Malaysian IT industry. Liu et al. [70] reported the problems contributing to project failures in China. Jarzębowicz and Ślesiński [71] surveyed analysts from Poland to learn about the RE problems they cope with in their work. Gobov and Titlova [72] analyzed the challenges in BA reported by Ukrainian practitioners.

Another focus of industrial surveys is the state of practice in RE and BA, in which respondents are asked about the practices, techniques, and tools they use in their work. Kassab and Laplante [73] investigated a broader spectrum of RE practices and compared the outcomes of three survey editions conducted over a 15-year period. Similarly, Ozkaya et al. [74] surveyed practitioners about the practices and techniques they use in several RE areas. As agile methods become prevalent in the IT industry, Ochodek and Kopczyńska [75] focused on agile RE practices and investigated how practitioners perceive the relative importance of specific practices. Vogelsang et al. [76] explored potential differences in RE perception between scholars and practitioners by asking practitioners to rate selected research statements extracted from the scientific RE literature. Several surveys had a more restricted scope and focused on specific sub-areas or aspects of RE, such as requirements elicitation techniques [77], requirements documentation techniques [78], and strategies for documenting non-functional requirements in agile software projects [79].

Probably the most widespread survey studies in RE, comprising both problems/challenges and practices, are conducted under the auspices of the NaPiRE (Naming the Pain in Requirements Engineering) initiative [80, 81]. There have been four completed runs of the NaPiRE survey: the first in 2012/2013, in only two countries; and the last in 2024, with a worldwide scope. The usage of LLMs (or AI in general) in RE is, however, not covered by any of them [82].

None of the survey studies dedicated to RE problems or practices listed in this subsection include any aspects of LLMs adoption (or AI adoption in general). However, we identified a small number of more recent studies addressing this topic following the surge in GenAI adoption.

Hess et al. [29] designed a questionnaire that inquired about GenAI usage across the main areas of RE as defined by the IREB (International Requirements Engineering Board). Also, they asked about the perceived usefulness/harmfulness of using GenAI in particular tasks and techniques included in those areas. In addition, they asked about practitioners' positive and negative experiences applying GenAI in RE (to identify barriers and opportunities) and their needs for additional training to effectively adopt AI in their work. The survey received 57 complete responses, mostly from European countries and likely from people associated with the IREB community, as this organization's channels were the primary way of distributing invitations. The results show an uneven ratio of GenAI applications in particular areas, as Elicitation and Specification are the ones where GenAI is most commonly used, while Analysis and Negotiation, as well as Validation, are not as strongly supported, and Management is the area where practitioners do not apply GenAI (yet), as only 14% of them confirmed such usage.

Rani et al. [28] questioned 55 software practitioners if they apply AI solutions to particular RE areas (Elicitation, Analysis, Specification, and Validation) and to what extent, with respect to the degree to which decision-making is impacted by AI (human-only decisions, AI validation, human–AI collaboration, full AI automation). In addition, they asked about participants' perceptions of applying AI to RE tasks, as well as the associated challenges and opportunities. They found that AI is used in all investigated areas, mostly as part of a human-AI collaboration model. Also, they report highly positive attitudes of the respondents towards AI. Several challenges and opportunities identified by respondents are listed in this paper, but none are voiced by a significant share of the survey population.

Santos et al. [27] designed a survey on LLMs use in RE. It specifically asked about the LLMs' application in generating User Stories. Apart from that, the questions on possible usage patterns were rather generic ("For which activities do you usually use LLMs?"). An open question about positive and negative impacts of LLMs on practitioners' work was also included. The survey yielded 21 complete responses. Among participants who use LLMs in their professional activities, 75% reported using them to generate User Stories. The answers regarding other LLM-supported activities were very varied, and no trends were found in this area. The study also reports that automated generation of User Stories with LLMs is perceived as promising, especially when LLMs are considered a supporting tool rather than a replacement of a human decision-maker.

Pir Muhammad et al. [26] investigated how LLM recommendations influence human decision-making related to requirements prioritization. The participants in the study first assigned priorities themselves and were afterward provided with the LLM's priority proposal, along with additional justifications. Participants could change their initial choices and were asked to assess the usefulness of LLM support. The results showed that participants rarely changed their decisions (treating LLM as an aid rather

than an authority), but their confidence in choices made increased thanks to the additional input provided by an LLM.

2.4 Research Gap and Positioning

The substantial existing literature on applications of LLMs to RE and BA tasks usually omits practitioners' perspectives on their actual use of LLMs in real-world work and focuses on evaluating LLMs and/or associated tooling, such as prompts, under controlled conditions. The studies on LLM adoption that utilized recognized technology adoption frameworks investigated several aspects of LLM-supported software development and identified relevant adoption factors, but no study applied any of these frameworks to the RE/BA field. Among the surveys investigating the RE/BA state of practice and exploring various processes, practices, and techniques, only a handful of the most recent studies capture practitioners' LLM usage, which stands in striking contrast to the widespread use of LLMs for requirements-related tasks. Thus, our article is one of the very few and helps address the existing gap. As our study targets a different population, includes aspects not covered by the existing LLM-focused surveys (especially LLM adoption in BA practitioners' organizations), and draws its conclusions from a much higher number of responses, we argue that it provides unique contributions with respect to the most closely related work [26–29].

In addition to scope, number of responses, and methodology, our study differs from prior studies along three dimensions. First, we adopt a business analysis perspective. While the closest studies survey requirements engineers or general software practitioners using RE frameworks, we target BA practitioners and align our task taxonomy with BABOK [14], capturing activities that extend beyond the classical RE scope covered by [27–29]. Second, among the LLM-adoption surveys focused specifically on RE/BA practice, we integrate adoption levels, task-level usage patterns, tool choice, concerns, attitudes, and organizational readiness within a single survey. Third, we apply an analytical approach: unlike the descriptive designs of [27–29] and the individual-decision experiment of [26], our study tests associations between contextual variables and adoption-related outcomes using a stratified, scale-appropriate analytical design. Our cross-national design further enables comparison of LLM adoption across two distinct national and institutional contexts, Ukraine and Poland. The comparison to related work is presented in Table 1.

3 Methodology

This study employs a survey-based empirical research design drawing on survey practices used in the NaPiRE initiative [80, 81] and draws on prior BA practice surveys conducted in Ukraine and Poland [78, 83, 84]. The study adopts a cross-sectional design with a single wave of data collection: responses were collected between 15 January and 20 February 2026. The analytical approach combines descriptive statistics to characterize adoption patterns (RQ1–RQ3) with scale-appropriate bivariate tests and multivariable regression to identify context-dependent adoption patterns (RQ4). Figure 1 gives an overview of the study design.

Table 1: Characteristics of the included studies

Study	Population	Respondents	Geography	Scope of LLM use	Statistical tests	Adoption perspective
Hess et al. [29]	RE practitioners	57 (+ 45 incomplete)	Europe (with few exceptions)	RE areas and tasks	No	Individual and organizational
Rani et al. [28]	Software practitioners	55	N/A	RE areas	No	Individual
Santos et al. [27]	Software practitioners involved in RE	21	Brazil	Selected RE practices, mostly User Story generation	No	Individual and team
Pir Muhammad et al. [26]	Software practitioners	33	N/A	Requirements prioritization	No	None
This study	BA practitioners	469	Ukraine and Poland	BA tasks	Yes	Individual and organizational

3.1 Survey Instrument

The survey instrument was prepared using Google Forms services. It comprised seven thematic blocks covering BA practice and the use of LLMs: (B1) respondent and project profile, (B2) BA activities conducted, (B3) LLM usage level and tools, (B4) LLM task usage, (B5) concerns and limitations related to LLM usage, (B6) attitudes toward AI in BA, and (B7) organizational AI readiness. Blocks B1–B2 are based on an instrument used in a prior survey of BA practice [84], while blocks B3–B7 were designed specifically for this study to capture the LLM adoption phenomenon. The LLM-related blocks were part of a broader survey on BA practice; the remaining blocks (e.g., on requirements elicitation and documentation practices) address different research questions and will be investigated separately.

The instrument contains 9 questions (blocks B3–B7) concerning LLMs in BA: five categorical (LLM usage level, AI replacement belief, willingness to learn, management support, company AI solutions), three multiple-choice with binary indicators (AI tools used, BA tasks supported, concerns), one five-point rating scale (LLM effectiveness rating). The BA task categories were derived from the BABOK Guide [14]. Block B2 covers all six BABOK knowledge areas, whereas the list of LLM-supported tasks (B4, Q37) comprises eight tasks that relate mainly to Elicitation and Collaboration and to Requirements Analysis and Design Definition, together with general support activities such as knowledge search and stakeholder communication; it does not include

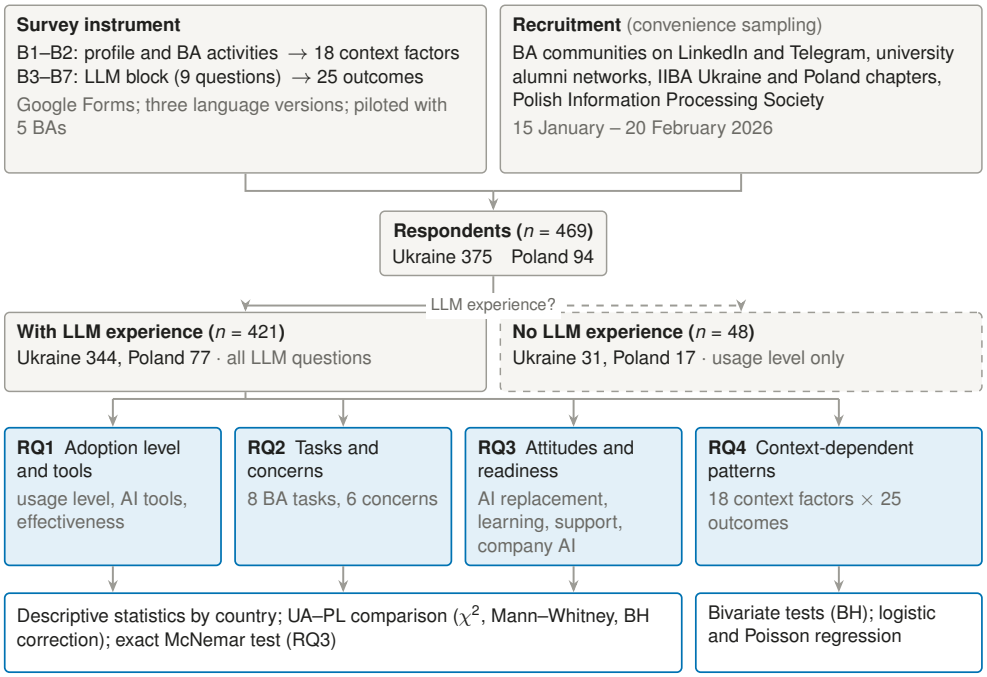

Fig. 1: Study design: survey instrument, recruitment, routing of respondents by LLM experience, and the analyses used for each research question

tasks specific to Strategy Analysis, Solution Evaluation, or Business Analysis Planning and Monitoring (see Section 5.2). We chose BABOK as the reference framework because the survey was distributed mainly through BA communities, including the IIBA chapters that publish it.

The instrument was developed in English and translated into Ukrainian and Polish by native-speaking BA practitioners to ensure terminological accuracy. Before deployment, the instrument underwent pilot testing with five senior BA practitioners (three Ukrainian and two Polish, each with at least 7 years of experience), resulting in refinements to the wording of questions and the clarity of response options.

Participation in the survey was voluntary and anonymous. The first page of the questionnaire informed respondents about the study's purpose, the intended use of the results in aggregated form for research publications, the approximate completion time (20–30 minutes), and their right to stop at any time without giving a reason. Respondents gave their informed consent by confirming that they had read this information and agreed to participate before proceeding to the questions. No email addresses, IP addresses, or other information about respondents that could reveal their identities were collected by default; however, we provided an option to leave an email address for those who wished to receive the report summarizing the survey results. These addresses were stored separately from the survey responses in accordance with the GDPR [30] and are not part of the dataset. The Ukrainian regulations on scientific

research require no detailed ethical review for an anonymous survey of adult professionals with no medical intervention and no collection of sensitive personal data. All three language versions were distributed in parallel.

3.2 Sample and Data Collection

The survey was distributed through professional BA communities on LinkedIn and Telegram, in partnership with university alumni networks in Ukraine and Poland. We also cooperated with the IIBA Ukraine and IIBA Poland chapters, as well as the Polish Information Processing Society, which facilitated the sending of invitations through their internal channels. The inclusion criteria required respondents to perform BA or RE activities in a software project, regardless of job title, and to reside in Ukraine or Poland at the time of survey completion.

Survey responses from all three language versions were merged into a single dataset with standardized English-language labels. The survey instrument (all three language versions, limited to the questions analyzed in this article), the anonymized coded dataset used in the analysis, a description of the data transformations, the Python scripts that reproduce all reported results, and the results for all tested context–outcome pairs (Table S1) are available as supplementary material. If the article is accepted, the supplementary materials will be published in a permanent location (e.g., Zenodo, Figshare, or Mendeley Data).

We received 485 responses; 15 respondents residing outside Ukraine and Poland were excluded, and one respondent who selected no BABOK area (Q13) and no analyst role (Q8) was excluded, so the final sample comprised n = 469 respondents: 375 from Ukraine (80.0%) and 94 from Poland (20.0%). We checked the remaining responses for data quality. Our checks did not reveal suspected duplicate submissions: each response differed from the most similar other response in at least 57 of the 241 questionnaire items. Because the survey was anonymous, repeated participation by the same person cannot be ruled out entirely. No respondent gave identical answers to all items in more than two of the five multi-item grids of the questionnaire, and no respondent with LLM experience selected all tasks and all concerns. 30 respondents (6.4%), mostly product owners or project managers, held no analyst role but were retained because they performed BA activities. No further responses were excluded.

3.3 Data Transformations

The dataset ($n = 469$) was split by country into two independent sub-samples: Ukraine ($n = 375$) and Poland ($n = 94$). All subsequent steps are applied separately to each sub-sample.

Several categorical variables were recoded before testing to meet the minimum expected count assumption (the tests are listed in Table 2); the same coding was applied to both sub-samples. The “Up to 1 year” and “1–3 years” experience categories were merged into “Up to 3 years” due to low cell counts in each category. Company size categories with fewer than 50 respondents in both countries together were consolidated into “Up to 249 employees”. “Agile/Lean” and “Rather Agile” were merged into a single level of the ordinal scale (Agile

→ Hybrid → Plan-driven), as both represent the same end of the methodology spectrum. Outsource and Outstaff company types were merged into the "Service" category. Industry (sector or application domain) was recoded into four categories grouping the following business sectors: Finance (Finance/Banking; a regulated sector kept separate because it has the highest number of responses), Commercial (e-Commerce/Retail/FMCG, Media/Advertising/Entertainment, Travel & Tourism, ERP), Other regulated (Healthcare/Pharmaceuticals, e-Government, Telecommunication), plus an "Other" category encompassing all other answers including the free-text ones. Role selections (multiple choice) were classified into three groups: analyst (a single Business Analyst role, or Business Analyst combined with Requirements Engineer), analyst-deepening (other combinations within the analyst family: business, system, and data analyst), and cross-functional (any non-analyst role present, e.g., Product Owner, Project/Delivery Manager, Scrum Master, Architect, QA). Project execution country and stakeholder location were mapped to four categories relative to the respondent's home country: Domestic, EU+GB, USA, and Other foreign.

The more detailed explanation of all data transformations applied is included in the supplementary materials.

3.4 Statistical Methods

Descriptive statistics.

Frequency distributions and percentages were computed for all LLM-related variables; proportions are reported with Wilson 95% confidence intervals [85]. Cross-tabulations comparing the Ukrainian and Polish sub-samples were generated for key adoption indicators (RQ1–RQ3), and the two sub-samples were compared using the χ^2 test for nominal variables and the Mann–Whitney test with Cliff's δ for ordinal variables; the Benjamini–Hochberg correction [86] was applied across all 25 comparisons between the two countries (23 adoption indicators plus the distribution of usage levels and the share of active users). This set of comparisons is corrected separately from the associative analysis and the regression models described below, which were run within each country. The composition of the two sub-samples was compared in the same way across the 18 context factors (Section 4.1), with a separate correction for these 18 comparisons. The exact McNemar test [87] was applied to explore whether respondents' willingness to learn AI tools was as frequent as strong management support.

Associative analysis.

To identify context-dependent patterns (RQ4), we tested associations between 18 context factors and 25 LLM outcome variables, grouped into six outcome families: AI tools (3), BA tasks (8), concerns (6), attitudes and usage (5: usage level, effectiveness rating, AI replacement belief, willingness to learn, management support), company AI solutions (1), and composite counts (2), i.e., the numbers of tasks and concerns selected by a respondent. The pairs are defined by RQ4: all context factors from blocks B1–B2 were crossed with all LLM outcomes from blocks B3–B7, whereas associations among factors or among outcomes were not tested. Because RQ4 is exploratory, all 450 pairs per country (54, 144, 108, 90, 18 and 36 in the six families) were tested, and the false discovery rate was controlled using the Benjamini–Hochberg procedure within

each outcome family. Respondents without LLM experience ($n = 48$) answered only the usage-level question, so all other outcomes were analyzed for respondents with LLM experience; usage level was analyzed for the full sample, since its lowest levels are among non-users. Tests were selected based on the measurement levels of the two variables in each pair [88]. Because most variables are nominal or short ordinal scales with many ties (three to five response categories for most ordinal questions), we used categorical and rank-based tests rather than methods that assume interval data or normality [88]; the χ^2 test with Cramér's V also keeps the analysis comparable with prior surveys of BA practice [72, 83]. Table 2 lists the test and effect size used for each combination of scales.

Table 2: Statistical methods used (bc - bias corrected)

Outcome → Predictor ↓	**Binary outcome**	**Nominal outcome**	**Ordinal outcome**
Nominal predictor	χ^2 test of independence; Cramér's V (bc)	χ^2 test of independence; Cramér's V (bc)	Kruskal–Wallis test; η^2_H
Ordinal predictor	Cochran–Armitage trend test; Kendall's τ_b	χ^2 test of independence; Cramér's V (bc)	Kendall's τ_b correlation; τ_b (Spearman's ρ for reference)

Additional explanations regarding the methods used:

- Sparse tables ($> 20\%$ of cells with expected count < 5, or any cell with expected count < 1): Fisher's exact test for 2×2 tables and a Monte Carlo χ^2 test (3,000 resamples) for larger tables [87]; only the Monte Carlo branch occurred, as every predictor has at least three categories.
- The Kruskal–Wallis test uses a permutation p-value (3,000 resamples) when the smallest group has $n < 7$; this applied to five Ukrainian pairs with the factor "Project country", none of which was significant after correction. Monte Carlo and permutation p-values are computed as $(b + 1)/(B + 1)$, where B is the number of resamples and b is the number of resampled statistics at least as extreme as the observed one.
- $\eta^2_H = (H - k + 1)/(n - k)$ [89], where H is the Kruskal–Wallis statistic, k the number of groups, and n the number of observations; it is truncated at zero, so a reported value of 0 is a lower bound.
- The 95% confidence intervals for the effect size were obtained using the percentile bootstrap intervals (2,000 resamples). Since V and η^2_H cannot be negative, their intervals can start at zero even for significant associations.
- Multiplicity: Benjamini–Hochberg correction [86] within each of the six outcome families; as a sensitivity check, Table S1 (supplementary materials) also reports the correction across all 450 pairs per country. Because Benjamini–Hochberg thresholds

depend on the number and distribution of p-values in each set, the correction across all pairs is not uniformly stricter than the correction within families.

We interpret Cramér's V (bias-corrected [90]) using Cohen's thresholds for w (0.10, 0.30 and 0.50 for small, medium and large effects) divided by $(\min(r, c) - 1)^{1/2}$ [91]: for a binary outcome the thresholds remain 0.10, 0.30 and 0.50, and for a table with $\min(r, c) = 3$ they are 0.07, 0.21 and 0.35. η^2_H is interpreted using 0.01, 0.06 and 0.14 [91], and Cliff's δ (Table 5) using 0.147, 0.33 and 0.474 [92]. We used these thresholds as descriptive benchmarks, with no implication that they indicate practical importance in other contexts. Kendall's τ_b is a rank correlation adjusted for ties (from -1 to 1), and its sign indicates the direction of the trend [87].

Multivariable analysis.

Binary logistic regression models were estimated separately for each country. In Ukraine, four outcomes were modeled – regular LLM usage (all respondents, $n = 375$) and use of ChatGPT, Gemini and Copilot (respondents with LLM experience, $n = 344$) – with six predictors (nine terms, as company type and industry were entered as two and three dummy variables): company size, company type, experience, development approach, industry and BA activity breadth, with ordinal predictors entered as linear scores, so each coefficient gives the change in odds per category step. This assumes equal effects between adjacent categories, a simplification given that the categories differ in width (e.g., 250–800 vs. 801–1500 employees); likelihood-ratio tests rejected it only for company size in the Gemini model ($p_{\text{adj(BH)}} = 0.008$). All VIFs were below 1.8. Events per variable (the smaller outcome class divided by the nine model terms) ranged from 8.4 to 18.0; only the ChatGPT model fell below the recommended 10 [93], so it was also re-estimated with Firth's penalized likelihood [94], which gave the same conclusions. Benjamini–Hochberg correction was applied to the nine coefficients of each model (primary) and to all 36 coefficients (sensitivity). The Polish sub-sample was too small for these models, so we fitted Firth-penalized logistic regressions [94] containing only the predictors that were nominally significant ($p < 0.05$) for the same outcome in Ukraine, with confidence intervals and p-values from the profile penalized likelihood. Because this selection relies on uncorrected p-values and the Polish models have few events (29–46), they only check whether the direction of the Ukrainian effects is preserved. ChatGPT was not modeled, as only 12 Polish respondents with LLM experience did not use it, and Gemini was not modeled because the Ukrainian Gemini model was not significant overall.

In each country, the numbers of tasks and concerns selected by respondents with LLM experience were modeled using Poisson regression on BA activity breadth, company size, experience, and development approach, with quasi-binomial models (for 8 tasks and 6 concerns) as a robustness check.

The following analyses were not planned and are not covered by the multiplicity correction: the comparisons in Sections 5.2 and 5.3, the robustness check in Section 4.5.2, and, in Section 4.5.4, the Gemini model with categorical company size and the Copilot model extended with company-provided generic AI tools.

The analysis scripts (Python) were developed with the assistance of a large language model (Claude Opus 5.5, Anthropic) and reviewed by the authors. All analytical decisions were made by the authors.

3.5 Threats to Validity

Our analysis of threats to validity is based on the guidelines by Wohlin et al. [88].

Construct validity.

The main threat to construct validity for a questionnaire-based survey is the risk that some questions or concepts may have been misinterpreted. Construct validity was supported by using established constructs from prior BA practice surveys for the profile blocks (B1–B2) and by deriving LLM-specific constructs (B3–B7) from recurring themes in the AI adoption literature (usage level, task application, concerns, attitudes, organizational readiness). Also, the survey instrument was pilot-tested with five BA practitioners, who provided a few minor suggestions regarding some terms and phrases used (addressed in the final version of the questionnaire) and confirmed that, in general, the questions were unambiguous and covered the relevant aspects of LLM use in BA practice. However, residual misinterpretation cannot be ruled out. In particular, LLM usage is measured through self-reported behavior rather than verified tool logs or observational data, which introduces the possibility of social desirability bias: practitioners may overstate their LLM usage to appear more up to date with the contemporary technology. The anonymous survey design mitigates this risk but cannot eliminate it entirely. Because all variables were self-reported by the same respondents in a single questionnaire, common method variance may have inflated some associations [95]; this applies in particular to associations between counts of selected items, such as BA activity breadth and the number of LLM-supported tasks (Section 4.5.2). Also, respondents' interpretation of "regularly" versus "occasionally" is not precisely defined and may vary. Similarly, the AI tool items refer to product names whose underlying models and features changed over time and differed across subscriptions, so a tool item captures the use of a named product at the time of data collection rather than the use of a specific model version or feature (Section 5.7). In particular, the option "Microsoft Copilot" did not distinguish Microsoft 365 Copilot, Microsoft 365 Copilot Chat, GitHub Copilot and the consumer Copilot app, so the Copilot associations cannot be attributed specifically to enterprise licensing (Section 5.5). Moreover, Claude was mentioned in the usage question but was not offered as a tool option and could only be named in free text (Section 4.2.2). Also, task categories are distinguished to gather information about the scope of LLM support and are defined in accordance with BABOK knowledge areas, which usually simplify the complexity of real BA workflows and require respondents to map their responsibilities to the presented list of items. Moreover, the task question (Q37) asked for which tasks respondents use or consider using LLMs, so the task shares may include intended as well as actual use; among regular users the task ranking was the same (Spearman's $\rho = 1.00$ in Ukraine, 0.90 in Poland) and the association between BA activity breadth and the number of tasks was slightly stronger ($\tau_b = 0.225$ and 0.430). We used terminology based on

BABOK, though some respondents may rely on other industry standards with different vocabularies. We assess such a threat as low because the BA terminology is quite common, but it is not possible to dismiss it entirely. In multilingual survey studies, translation errors can pose threats. However, in this case, questionnaire translations into Ukrainian and Polish were produced by native-speaking BA practitioners and reviewed by a second domain expert in each language. No formal back-translation was performed during instrument development; this is a limitation, particularly for the attitudinal items in blocks B6–B7.

Internal validity.

Because the design is cross-sectional and observational, the associations reported for RQ4 cannot be interpreted causally, and the descriptive comparisons between the countries (RQ1–RQ3) are not adjusted for the differences in sample composition reported in Section 4.1. Associations with stakeholder location and the location of key decisions were not adjusted for company type, with which both are strongly related, so they may partly reflect differences between company types. Respondents were recruited through professional communities, so several of them may work for the same employer. Such clustering violates the independence assumption of the tests and may produce associations that reflect a single organization's policy, for example, which AI tools are licensed, rather than an industry-wide pattern. The pattern in the Finance sector (the lowest ChatGPT and the highest Gemini use) is consistent with this explanation. We could not check this because employer names were not collected. The usage question served as a filter: respondents without LLM experience were not shown questions about AI tools, tasks, concerns, attitudes, or company AI solutions, thereby preventing logically inconsistent answers in the LLM block. As a consequence, concerns, attitudes and organizational readiness were measured only among respondents with LLM experience (Section 5.7).

External validity.

As we conducted our research as a sample study aiming to make inferences about a target population from a sample of volunteering respondents, the external validity may be threatened by several factors. The survey intentionally targeted BA practitioners from two specific countries; thus, its results cannot be generalized to the global BA population without replication in other geographic and professional contexts. In particular, both Ukraine and Poland have strong IT outsourcing sectors with internationally oriented projects, which may impose adoption patterns (e.g., tool diversity, confidentiality concerns) that differ from those in domestic-focused markets. Moreover, the sampling strategy relied on convenience sampling via professional networks, introducing potential self-selection bias. Practitioners who are more active in professional communities (IIBA) and BA networks on social media (LinkedIn, Telegram) were more likely to receive survey invitations. Also, the survey topic itself may have attracted respondents with a greater interest in AI tools, thus involving those more technologically engaged rather than those who are not. All adoption rates reported in this study – including the 80.8% active usage rate – should therefore be interpreted as characteristic of professionally active, community-engaged BA practitioners

in Ukraine and Poland, rather than as representative estimates of the broader BA population. Another threat concerns the composition of the sample. The intended respondent profile was clearly indicated in survey invitations, and we assumed that members of the target population (BA professionals) would not have difficulty understanding the questions. However, we cannot rule out that people without the necessary expertise decided to participate, and this cannot be verified given the anonymity of respondents. Response rate is a useful metric for assessing the quantitative aspects of the sampling process; however, computing a precise response rate is not feasible in our case because dissemination included open distribution channels.

Conclusion validity.

All associative tests were run within each country, so statistical power depends on the sub-samples rather than on the total sample: for a 2×2 association among respondents with LLM experience, the minimum effect detectable with 80% power at $\alpha = 0.05$ is $w \approx 0.15$ in Ukraine ($n = 344$) but $w \approx 0.32$ in Poland ($n = 77$). The lack of a significant result in Poland, therefore, does not indicate the absence of an association. A significant association in one country and a non-significant one in the other does not by itself show that the two countries differ [96]; we therefore do not interpret such contrasts as evidence of country differences. If none of the 450 associations tested per country existed, about 22.5 of them would be expected to reach $p < 0.05$ by chance alone, so all conclusions rest on Benjamini–Hochberg-adjusted p-values.

4 Results

This section presents findings organized by research question. Section 4.1 describes the respondent profile. Section 4.2 reports LLM adoption levels and tool preferences (RQ1). Section 4.3 covers task-level usage and concerns (RQ2). Section 4.4 examines practitioner attitudes and organizational readiness (RQ3). Section 4.5 identifies context-dependent adoption patterns (RQ4).

As our survey yielded an imbalance in the number of Ukrainian and Polish respondents (375 and 94, respectively), we present the results separately for each country in this section.

4.1 Respondent Profile

Demographic details of survey respondents are provided in Table 3, while Table 4 presents respondents' BA responsibilities at a high level of abstraction (by main BA areas defined according to BABOK).

The sample ($n = 469$) comprises respondents from Ukraine ($n = 375$, 80.0%) and Poland ($n = 94$, 20.0%). Most respondents have more than five years of experience in BA or similar roles, and about half work for large organizations with over 1,500 employees. Respondents with more than 10 years of experience (34.0% vs. 22.9%) and those working for organizations with over 1,500 employees (60.6% vs. 48.3%) are more frequent in Poland, but neither experience nor company size differs between the countries after correction (Table 5). About half of the projects follow an agile approach

Table 3: Respondents' demographics (all respondents; $n_{\mathrm{UA}} = 375$, $n_{\mathrm{PL}} = 94$)

Characteristic	Category	Ukraine		Poland	
		n	%	*n*	%
Experience	Up to 3 years	65	17.3%	12	12.8%
	3–5 years	91	24.3%	18	19.1%
	5–10 years	133	35.5%	32	34.0%
	Over 10 years	86	22.9%	32	34.0%
Company type	IT Outsource	117	31.2%	39	41.5%
	IT Outstaff	59	15.7%	24	25.5%
	IT Product	107	28.5%	18	19.1%
	Non-IT In-house	92	24.5%	13	13.8%
Company size	Up to 249 employees	93	24.8%	14	14.9%
	250–800 employees	56	14.9%	14	14.9%
	801–1500 employees	45	12.0%	9	9.6%
	Over 1500 employees	181	48.3%	57	60.6%
Development approach	Agile/Lean	69	18.4%	17	18.1%
	Rather Agile	118	31.5%	30	31.9%
	Hybrid	130	34.7%	41	43.6%
	Plan-driven or rather plan-driven	58	15.5%	6	6.4%
Team organization	Distributed	352	93.9%	89	94.7%
	Co-located	22	5.9%	4	4.3%
	I don't know	1	0.3%	1	1.1%

Note. For the statistical analyses, IT Outsource and IT Outstaff were merged into Service (Ukraine 46.9%, Poland 67.0%), and Agile/Lean and Rather Agile into Agile (Ukraine 49.9%, Poland 50.0%).

Table 4: Respondents' involvement in BA activities (all respondents; $n_{\mathrm{UA}} = 375$, $n_{\mathrm{PL}} = 94$)

Area	Ukraine		Poland	
	n	%	*n*	%
Business analysis planning & monitoring	209	55.7%	54	57.4%
Strategy analysis	219	58.4%	50	53.2%
Elicitation & collaboration	353	94.1%	83	88.3%
Requirements analysis and design definition	329	87.7%	89	94.7%
Solution evaluation	179	47.7%	41	43.6%
Requirements life cycle management	293	78.1%	76	80.9%

in both countries; hybrid approaches are reported more often in Poland and plan-driven ones in Ukraine, although the position of projects on the agile–plan-driven scale does not differ significantly between the countries. Almost all respondents work in distributed teams (about 94%). Service companies (IT Outsource and IT Outstaff) account for 67.0% of Polish and 46.9% of Ukrainian respondents. Finance/Banking

is the most represented industry (27.2% Ukraine, 24.5% Poland), followed by e-Commerce, ERP, Media & Entertainment, and Healthcare/Pharmaceuticals (5–14% each); 21.9% and 24.5% selected "Other".

Key stakeholders are most often located in the home country in Ukraine (48.4%) and in other EU countries or the UK in Poland (47.3%); the USA accounts for about a quarter of answers in both countries (24.2% and 29.0%). Key BA decisions are made mainly in the home country in Ukraine (56.7%), whereas in Poland they are most often distributed across locations (37.2%) or made mainly in another country (26.6%).

The most common roles declared are Business Analyst (46.9% Ukraine, 41.5% Poland), Business Analyst combined with System Analyst (18.7%, 11.7%), and Business Analyst combined with Product Owner/Product Manager (11.5%, 12.8%).

Respondents are mainly involved in Elicitation & collaboration, Requirements analysis and design definition, and Requirements life cycle management (78–95%), and less often in Business analysis planning and monitoring, Strategy analysis, and Solution evaluation (44–58%). Involvement in Solution evaluation increases with experience in both countries, and involvement in Strategy analysis and Business analysis planning and monitoring only in Ukraine (Kendall's $\tau_b = 0.10$–0.26). Overall, the number of areas in which respondents are involved (BA activity breadth, 0–6) also increases with experience ($\tau_b = 0.17$ in Ukraine and 0.23 in Poland); this variable is used as a context factor in Section 4.5.

BA practices are most often adapted at the project level (43.8% Ukraine, 50.0% Poland) rather than following organization-wide standards strictly (35.2%, 13.8%) or partially (21.0%, 36.2%). We also asked to what extent BA practices are defined by five sources. In both countries, practices are most often defined to a large or very large extent by the project team's own practices (66.1% in Ukraine, 63.8% in Poland) and by the individual experience of team members (62.9%, 56.4%), followed by client-defined standards (46.0%, 40.4%) and global organizational standards (34.4%, 34.0%). Local organizational standards play this role more often in Ukraine than in Poland (37.6% vs. 18.1%). These five sources are used as context factors in Section 4.5.

Because the two sub-samples differ in size, we compared their composition for every context factor used in the analysis (Table 5). Most differences are small ($|\delta| < 0.20$, $V \leq 0.15$). Moderate differences appear only for the three geographic factors – project execution country ($V = 0.36$), location of key decisions ($V = 0.30$) and stakeholder location ($V = 0.29$) – which reflect the more export-oriented project portfolio of the Polish respondents. Seven of the 18 factors differ between the countries after Benjamini–Hochberg correction; company size and BA experience differ only at the nominal level. These differences motivate the country-stratified analysis in Section 4.5; the descriptive comparisons between the countries (RQ1–RQ3) are not adjusted for them.

4.2 LLM Usage Patterns (RQ1)

4.2.1 Overall Adoption Rate

Table 6 presents the distribution of LLM usage levels among respondents.

Table 5: Country × characteristic: comparability of the two samples

Factor	Type	Test / meas.	$p_{\text{adj(BH)}}$	Effect
Company type	N	χ^2 / V	0.009*	0.148
Industry (4 categories)	N	χ^2 / V	0.932	0.000
Project category	N	χ^2 / V	0.645	0.000
Project country (4 categories)	N	χ^2 / V	< 0.001*	0.362
Company size	O	MW / δ	0.050	0.141
BA experience	O	MW / δ	0.054	0.142
Project size	O	MW / δ	0.012*	0.185
Development approach	O	MW / δ	0.645	−0.047
Stakeholder location (4 categories)	N	χ^2 / V	< 0.001*	0.289
Key-decisions location	N	χ^2 / V	< 0.001*	0.301
Practice autonomy	O	MW / δ	0.025*	0.160
Source: global standards	O	MW / δ	0.645	−0.043
Source: local standards	O	MW / δ	0.009*	−0.196
Source: client standards	O	MW / δ	0.863	−0.015
Source: team practices	O	MW / δ	0.427	−0.075
Source: individual experience	O	MW / δ	0.645	−0.049
BA activity breadth (0–6)[a]	O	MW / δ	0.762	−0.027
Role group[a]	N	χ^2 / V	0.721	0.000

Note. Type: N, nominal; O, ordinal. Test / meas.: χ^2, chi-square test with bias-corrected Cramér's V (truncated at zero); MW, Mann–Whitney test with Cliff's δ ($\delta > 0$: higher in Poland). $p_{\text{adj(BH)}}$: Benjamini–Hochberg correction across the 18 factors; highlighted rows and *: $p_{\text{adj(BH)}} < 0.05$; company size and BA experience differ only at the nominal level ($p_{\text{raw}} = 0.022$ and 0.027; $p_{\text{adj(BH)}} = 0.0502$ and 0.054). $n = 375$ (UA) and 94 (PL), except 372 (UA) for the location and BA practice factors and 92–93 (PL) for project country and stakeholder location, owing to missing or uncodable answers. [a] Derived variables (Sections 3.3 and 4.1).

Table 6: LLM usage level (all respondents; $n_{\text{UA}} = 375$, $n_{\text{PL}} = 94$)

Usage level	Ukraine		Poland	
	n	%	*n*	%
Yes, regularly	162	43.2%	46	48.9%
Yes, occasionally	147	39.2%	24	25.5%
Tried, but do not use regularly	35	9.3%	7	7.4%
No, but I plan to try	19	5.1%	11	11.7%
No, and I do not plan to use	12	3.2%	6	6.4%

Active LLM users (regular and occasional) constitute 80.8% of the sample (n = 379; 95% CI [77.0%, 84.1%]). Further, 9.0% have tried LLMs but do not use them regularly, while 6.4% plan to try them. Only 3.8% of respondents entirely reject the use of LLMs in their BA work.

Active usage is higher in Ukraine than in Poland (82.4% vs. 74.5%), but this difference is not significant ($\chi^2(1) = 3.05$, $p = 0.08$). The distribution across the five usage levels differs between the countries only at the nominal level ($\chi^2(4) = 11.9$, $p = 0.018$; $p_{\text{adj(BH)}} = 0.30$ across the 25 country comparisons described in Section 4.2.2), and its central tendency does not differ (Mann–Whitney $p = 0.90$):

Polish respondents are somewhat more often both regular users and non-users, whereas Ukrainian respondents more often use LLMs occasionally (Table 6). We therefore treat this difference as tentative.

For all subsequent analyses of tool preferences (Section 4.2.2), task-level usage (Section 4.3.1), and concerns (Section 4.3.2), the analytical base is restricted to $n = 421$ respondents who have direct experience with LLMs (i.e., those who reported using LLMs regularly, occasionally, or having tried them). Respondents who indicated no experience with LLMs ($n = 48$: 30 planning to try, 18 with no plans) were not presented with the corresponding questions, as it would be pointless given their lack of experience. Tables 7–10 therefore report percentages of respondents with LLM experience in each country (Ukraine $n = 344$, Poland $n = 77$; $n = 421$ in total).

4.2.2 AI Tools Used

Table 7 reports adoption rates for individual AI tools based on multi-select responses ($n = 421$; respondents with LLM experience).

Table 7: AI tools used (respondents with LLM experience; $n_{\mathrm{UA}} = 344$, $n_{\mathrm{PL}} = 77$; multiple responses permitted)

AI tool	**Ukraine**		**Poland**	
	n	**%**	***n***	**%**
ChatGPT / GPT-4 / GPT-5	268	77.9%	65	84.4%
Google Gemini	183	53.2%	30	39.0%
Microsoft Copilot	116	33.7%	29	37.7%
Jira / Confluence AI features	73	21.2%	14	18.2%
Miro AI	31	9.0%	5	6.5%
Notion AI	17	4.9%	7	9.1%

Cross-national tool preferences show ChatGPT dominance in both countries (Ukraine: 77.9%; Poland: 84.4% of respondents with LLM experience). Microsoft Copilot is used by 33.7% of respondents with LLM experience in Ukraine and 37.7% in Poland. Gemini is used less often in Poland (39.0% vs. 53.2% of respondents with LLM experience; χ^2 $p = 0.024$, uncorrected). We compared the two countries on 25 indicators: 23 LLM-related variables (with usage level compared by its central tendency), plus the distribution of usage levels and the share of active users (Section 4.2.1). Only the Gemini difference and the distribution of usage levels are nominally significant, and neither survives multiple-comparison correction ($p_{\mathrm{adj(BH)}} = 0.30$ for both), so we treat them as tentative.

The tool question also offered a free-text "Other" field. Claude, which was mentioned in the usage question but was not listed among the tool options, was named there by 22 respondents with LLM experience (5.2%), and Perplexity by 9 (2.1%). Because unlisted tools had to be typed in, these shares are lower bounds and are not comparable with those of the listed tools.

4.2.3 LLM Effectiveness Rating

Respondents with LLM experience ($n = 421$) rated the effectiveness of LLMs for their BA tasks on a five-point scale (Table 8). In both countries, most rated LLMs as rather effective or neutral, about one in seven as very effective, and a small minority (6.4% in Ukraine, 10.4% in Poland) as rather not or not effective. Polish respondents chose "rather effective" somewhat less often, but the rating distributions do not differ between the countries (Mann–Whitney $p = 0.42$).

Table 8: LLM effectiveness rating (respondents with LLM experience; $n_{\mathrm{UA}} = 344$, $n_{\mathrm{PL}} = 77$)

Rating	Ukraine		Poland	
	n	**%**	***n***	**%**
Very effective	48	14.0%	12	15.6%
Rather effective	152	44.2%	28	36.4%
Neutral	122	35.5%	29	37.7%
Rather not effective	16	4.7%	8	10.4%
Not effective	6	1.7%	0	0.0%

4.3 Task-Level LLM Usage and Concerns (RQ2)

4.3.1 BA Tasks Supported by LLMs

Table 9 presents the share of respondents who selected each task in response to "For which BA tasks do you use or consider using LLMs?" among respondents with LLM experience ($n = 421$).

Table 9: BA tasks for which respondents use or consider using LLMs (respondents with LLM experience; $n_{\mathrm{UA}} = 344$, $n_{\mathrm{PL}} = 77$; multiple responses permitted)

Task	Ukraine		Poland	
	n	**%**	***n***	**%**
Drafting requirements (user stories, use cases, specs)	249	72.4%	55	71.4%
Knowledge search / competitor research	239	69.5%	61	79.2%
Stakeholder communication (emails, notes, summaries)	210	61.0%	54	70.1%
Document analysis (contracts, RFPs, regulations)	195	56.7%	42	54.5%
Brainstorming elicitation questions or techniques	147	42.7%	38	49.4%
Visual modeling support (UML/BPMN from text)	115	33.4%	24	31.2%
Jira ticket / Confluence page generation	94	27.3%	20	26.0%
Wireframe / mockup creation	85	24.7%	19	24.7%

The most common BA tasks with LLM support are requirements drafting (72.4% in Ukraine, 71.4% in Poland) and knowledge/competitor search (69.5% in Ukraine,

79.2% in Poland), though the ranking order of these two tasks differs between countries. The tasks of stakeholder communication (61.0% Ukraine, 70.1% Poland) and document analysis (56.7% Ukraine, 54.5% Poland) follow, with over half of respondents reporting LLM support when performing them. Brainstorming elicitation questions or techniques is reported less often (42.7% in Ukraine, 49.4% in Poland). The remaining tasks, visual modeling support (33.4%, 31.2%), Jira/Confluence generation (27.3%, 26.0%) and wireframe/mockup creation (24.7%, 24.7%), are reported by about a quarter to a third of respondents with LLM experience.

4.3.2 Concerns and Limitations

Table 10 presents the concerns reported by practitioners regarding LLM use in BA tasks.

Table 10: Concerns about LLM use (respondents with LLM experience; $n_{\mathrm{UA}} = 344$, $n_{\mathrm{PL}} = 77$; multiple responses permitted)

Concern	**Ukraine**		**Poland**	
	n	**%**	***n***	**%**
Accuracy / hallucinations	285	82.8%	63	81.8%
Confidentiality / security of data	228	66.3%	52	67.5%
Lack of domain knowledge in LLM	144	41.9%	32	41.6%
Limited integration with existing BA tools	105	30.5%	26	33.8%
Ethical concerns (role replacement, bias)	49	14.2%	9	11.7%
Cost of usage	32	9.3%	10	13.0%

Accuracy and hallucinations represent the dominant concern, cited by 82.8% of respondents with LLM experience from Ukraine and 81.8% from Poland. They are followed by the challenge related to confidentiality and data security (66.3% Ukraine, 67.5% Poland). The second tier of concerns – lack of domain knowledge in LLMs (41.9% Ukraine, 41.6% Poland) and limited integration with BA tools (30.5% Ukraine, 33.8% Poland) – relates to practical limitations of current technologies. Ethical concerns (14.2% Ukraine, 11.7% Poland) and cost of LLM use (9.3% Ukraine, 13.0% Poland) rank lowest among the measured concern categories. This low share should not be interpreted as low importance of ethics to practitioners: the category combines heterogeneous issues (role replacement and bias) in a single answer option, and ethically relevant issues such as data confidentiality were offered as a separate category. The shares of the concerns do not differ significantly between the countries, and their ranking is the same, except that in Poland, cost and ethical considerations swap places in the last two positions.

4.4 Attitudes and Organizational Readiness (RQ3)

4.4.1 AI Replacement Belief

When asked whether AI could partially or fully replace a business analyst, respondents who declared experience with LLMs ($n_{\text{UA}} = 344$, $n_{\text{PL}} = 77$) express predominantly moderate views: most of them (60.8% Ukraine, 57.1% Poland) believe AI will partially replace BA tasks. The second group (33.4% Ukraine, 32.5% Poland) predicts that AI will remain only a support tool. Only 4.1% of respondents from Ukraine and 9.1% from Poland believe in full replacement. The remaining respondents with LLM experience (1.7% Ukraine, 1.3% Poland) were uncertain and selected the "not sure" answer. The dominant view is thus one of partial automation rather than wholesale displacement; our data do not show which parts of the BA role respondents expect to persist.

4.4.2 Willingness to Learn AI Tools

Practitioner openness to AI tools is overwhelmingly positive. There are two main groups: the first one (44.8% Ukraine, 48.1% Poland) comprises practitioners who are very willing to learn and already dedicating time to this task, the second one (48.8% Ukraine, 44.2% Poland) represents people somewhat willing if tools prove useful. Together, 93.3% of respondents with LLM experience (95% Wilson CI [90.6%, 95.4%]) express positive willingness. Only a small fraction (5.8% Ukraine, 7.8% Poland) declares a neutral stance and almost none (0.6% Ukraine, 0.0% Poland) express unwillingness.

4.4.3 Management Support

Organizational endorsement of AI adoption shows a more differentiated picture. Most respondents (54.7% Ukraine, 55.8% Poland) report strongly supportive management that encourages AI adoption and provides necessary resources. The others (27.9% Ukraine, 20.8% Poland) describe moderate support, where LLM usage is allowed but not prioritized. The less frequently chosen answers include: neutral management with no clear position (13.7% Ukraine, 13.0% Poland), uncertainty about management's stance (2.6% Ukraine, 6.5% Poland), and unsupportive management that discourages AI use (1.2% Ukraine, 3.9% Poland).

4.4.4 Company AI Solutions

The availability of organizational AI infrastructure further accentuates the readiness gap. Among respondents with LLM experience, only a fraction of them (21.2% Ukraine, 20.8% Poland) report that their company has custom AI solutions dedicated to BA (e.g., internal GPT instances, chatbots, automated assistants). The largest group of respondents (30.5% Ukraine, 36.4% Poland) are provided by their organizations with AI tools, but generic ones, not specifically tailored for BA tasks (e.g., enterprise Copilot licenses). The other answers include: organization's plans to implement AI solutions in future (16.0% Ukraine, 6.5% Poland); no such plans available in respondent's organization yet (22.1% Ukraine, 26.0% Poland); and uncertainty (10.2% Ukraine, 10.4% Poland).

Taken together, the three indicators show a descending pattern: individual readiness is high (93.3% willing to learn; Ukraine 93.6%, Poland 92.2%), organizational endorsement is partial (54.9% strongly supportive; Ukraine 54.7%, Poland 55.8%), and infrastructure deployment – particularly BA-specific tooling (21.1%; Ukraine 21.2%, Poland 20.8%) – is the weakest link. Figure 2 contrasts these indicators in both countries and shows the paired answers on willingness to learn and strong management support.

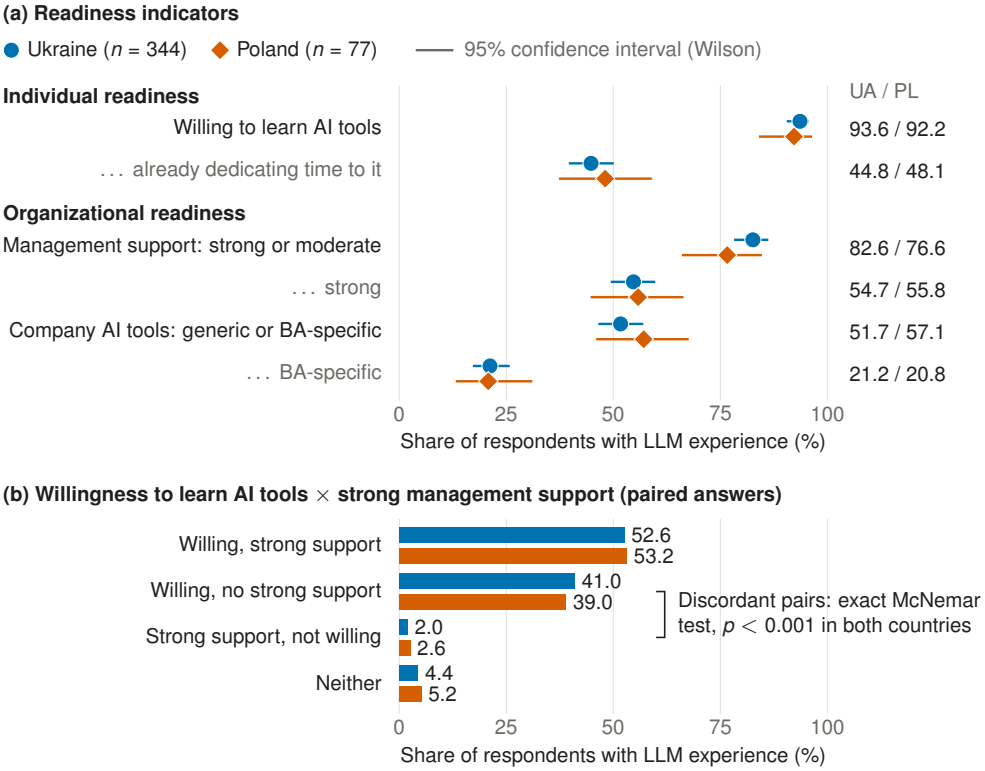

Fig. 2: Individual versus organizational readiness for AI adoption among respondents with LLM experience (Ukraine $n = 344$, Poland $n = 77$). (a) Share of respondents for each indicator with 95% Wilson confidence intervals; rows starting with "..." are subsets of the row above. (b) Paired answers on willingness to learn AI tools and strong management support; the two discordant groups were compared with the two-sided exact McNemar test

4.5 Context-Dependent Adoption Patterns (RQ4)

We tested the 450 context–outcome pairs described in Section 3.4 for each country separately. In Ukraine, 73 pairs were nominally significant ($p < 0.05$) and 25

remained significant after the Benjamini–Hochberg correction within outcome families; in Poland, 28 pairs were nominally significant and 1 remained significant. Table 11 lists these 26 associations; 17 of the 25 Ukrainian, but not the Polish one, also remain significant after correction across all 450 pairs of the country (Table S1 in the supplementary materials). Conversely, two Ukrainian pairs were significant only under the correction across all pairs: BA experience with usage level and global organizational standards with the effectiveness rating (Table S1).

Most context factors are described in Section 4.1. In Table 11, "BA activity breadth" is the number of BABOK knowledge areas in which a respondent is involved (Table 4); "Practice autonomy" reflects whether BA practices follow organization-wide standards or are adapted at the project level; and "Source: ..." denotes the extent to which BA practices are defined by the given source. "No. of tasks" and "No. of concerns" are the numbers of tasks (Table 9) and concerns (Table 10) selected by a respondent. Below, the associations are reported by outcome family, followed by the multivariable models.

4.5.1 Tool Choice Associations

In Ukraine, seven associations concerned the choice of specific AI tools; in Poland, none did. The strongest was between industry and ChatGPT use ($V = 0.300$), at the boundary of a medium effect. This association mainly reflects the Finance sector, where 57.6% of Ukrainian respondents with LLM experience report ChatGPT use, compared with 78.7% in Other, 85.8% in Commercial, and 93.0% in Regulated other industries (adjusted standardized residual for Finance: -5.5). ChatGPT use also decreases with company size ($\tau_b = -0.196$), from 88.5% in companies with up to 249 employees to 68.7% in companies with over 1,500 employees.

Another pattern has been identified for the use of Copilot. Copilot use increases from Agile through Hybrid to Plan-driven projects (25.9%, 39.0% and 46.4%; $\tau_b = 0.164$) and with company size (23.0% and 24.0% in companies with up to 800 employees vs. 46.3% and 39.2% in larger ones; $\tau_b = 0.137$) and with BA activity breadth (from 23.3% among respondents involved in up to two BA areas to 41.5% among those involved in five or six; $\tau_b = 0.154$).

Gemini use is associated with the location of key decisions ($V = 0.186$) and with industry ($V = 0.172$). It is highest where key decisions are made domestically (61.9%) and lowest where decision-making is distributed across locations (37.7%); by industry, it is most common in Finance (65.2%) and least common in Regulated other (40.4%).

Three of these tool associations, Copilot use with BA activity breadth and with company size and Gemini use with industry, do not remain significant after correction across all 450 pairs (Table 11); the two Copilot associations are, however, supported by the multivariable model (Section 4.5.4).

Thus, in Ukraine, ChatGPT use is lower in Finance and in larger companies, whereas Copilot use is higher in larger companies and in more plan-driven projects. These associations are examined further in the multivariable models (Section 4.5.4).

Table 11: Bivariate associations between context factors and LLM adoption outcomes that remain significant after the Benjamini–Hochberg correction within outcome families

Factor	Outcome	Test / meas.	$p_{\text{adj(BH)}}$	Effect [95% CI]
Ukraine (respondents with LLM experience, $n = 344$)				
Industry	Tool: ChatGPT	χ^2 / V	$< 0.001^{**\dagger}$	0.300 [0.195, 0.415]
BA activity breadth	No. of tasks	K / τ_b	$< 0.001^{**\dagger}$	0.201 [0.117, 0.283]
Key-decisions location	No. of tasks	KW / η^2_H	$0.002^{**\dagger}$	0.050 [0.018, 0.112]
BA activity breadth	No. of concerns	K / τ_b	$0.002^{**\dagger}$	0.163 [0.079, 0.244]
Stakeholder location	No. of tasks	KW / η^2_H	$0.002^{**\dagger}$	0.048 [0.014, 0.113]
Company size	Tool: ChatGPT	CA / τ_b	$0.005^{**\dagger}$	−0.196 [−0.281, −0.102]
BA activity breadth	Task: Jira/Confluence generation	CA / τ_b	$0.005^{**\dagger}$	0.192 [0.102, 0.280]
Stakeholder location	Task: brainstorming	χ^2 / V	$0.005^{**\dagger}$	0.235 [0.133, 0.349]
Company size	Company AI solutions	χ^2 / V	$0.009^{**\dagger}$	0.149 [0.111, 0.243]
Company type	Task: brainstorming	χ^2 / V	$0.010^{*\dagger}$	0.209 [0.110, 0.313]
Stakeholder location	Task: stakeholder communication	χ^2 / V	$0.012^{*\dagger}$	0.214 [0.126, 0.313]
Company type	Task: stakeholder communication	χ^2 / V	$0.014^{*\dagger}$	0.197 [0.087, 0.304]
BA activity breadth	Task: brainstorming	CA / τ_b	$0.022^{*\dagger}$	0.153 [0.056, 0.241]
Practice autonomy	Task: knowledge search	CA / τ_b	$0.024^{*\dagger}$	0.166 [0.068, 0.262]
Key-decisions location	Task: wireframe creation	χ^2 / V	$0.027^{*\dagger}$	0.191 [0.081, 0.320]
Development approach	Tool: Copilot	CA / τ_b	$0.028^{*\dagger}$	0.164 [0.062, 0.264]
Key-decisions location	Tool: Gemini	χ^2 / V	$0.028^{*\dagger}$	0.186 [0.078, 0.312]
BA activity breadth	Tool: Copilot	CA / τ_b	0.030^{*}	0.154 [0.060, 0.249]
Company size	Tool: Copilot	CA / τ_b	0.031^{*}	0.137 [0.043, 0.229]
Industry	Tool: Gemini	χ^2 / V	0.034^{*}	0.172 [0.061, 0.297]
BA activity breadth	Task: knowledge search	CA / τ_b	0.043^{*}	0.138 [0.043, 0.233]
Role group	No. of concerns	KW / η^2_H	0.044^{*}	0.024 [0.001, 0.073]
Source: individual experience	No. of tasks	K / τ_b	0.044^{*}	0.117 [0.030, 0.206]
Source: team practices	No. of tasks	K / τ_b	0.047^{*}	0.113 [0.027, 0.200]
Company type	No. of tasks	KW / η^2_H	0.047^{*}	0.021 [0.000, 0.070]
Poland (respondents with LLM experience, $n = 77$)				
BA activity breadth	No. of tasks	K / τ_b	0.008^{**}	0.334 [0.167, 0.484]

Note. Test / meas.: χ^2, chi-square test with bias-corrected Cramér's V; CA, Cochran–Armitage trend test, and K, Kendall's correlation, both with Kendall's τ_b ($\tau_b > 0$: the outcome increases with the factor); KW, Kruskal–Wallis test with η^2_H. Effect sizes are not comparable across measures. 95% CI: percentile bootstrap (2,000 resamples). $p_{\text{adj(BH)}}$: Benjamini–Hochberg correction within the outcome family; * $p_{\text{adj(BH)}} < 0.05$, ** $p_{\text{adj(BH)}} < 0.01$; † also significant after the correction across all 450 pairs of the country. Two further Ukrainian pairs (project country and role group with the number of tasks; $p_{\text{adj(BH)}} = 0.054$) lie just above the threshold; the classification of these pairs depends on the Monte Carlo error of a permutation p-value. Raw p-values and the correction across all pairs are given in Table S1.

4.5.2 Task and Concern Associations

The most consistent pattern across both national sub-samples concerns the breadth of respondents' BA involvement. For Ukraine, BA activity breadth was positively associated with the number of LLM-supported BA tasks ($\tau_b = 0.201$). It was also positively associated with the number of reported concerns about LLM use ($\tau_b = 0.163$), with the use of LLMs for Jira/Confluence generation ($\tau_b = 0.192$) and for brainstorming elicitation questions or techniques ($\tau_b = 0.153$), and for knowledge search ($\tau_b = 0.138$; not significant after correction across all pairs).

The same broad pattern was reproduced for Poland. BA activity breadth was positively associated with the number of LLM-supported tasks ($\tau_b = 0.334$). Thus, across both countries, respondents involved in a broader range of BA activity areas also reported a broader range of LLM-supported BA tasks. This association was moderate in Poland and smaller in Ukraine. The two confidence intervals overlap ($[0.167, 0.484]$ and $[0.117, 0.283]$), and we did not formally test the difference. Because both variables are counts of selected items, the association could partly reflect a general tendency to select many options. However, it remains after adjusting for the number of reported concerns and for experience (linear model, Ukraine $b = 0.29$, Poland $b = 0.67$, both $p < 0.001$), and few respondents selected all eight tasks (2.3% in Ukraine, 6.5% in Poland). Breadth also increases with experience (Section 4.1), so it partly captures seniority. In Ukraine, however, the number of tasks was most strongly associated with involvement in Business Analysis Planning and Monitoring, for which no task was listed (IRR = 1.20, 95% CI $[1.06, 1.36]$), so breadth may reflect a broader role rather than area-specific LLM use.

In Ukraine, organizational and project-location variables were also associated with task-level outcomes. Stakeholder location was associated with the number of LLM-supported tasks ($\eta^2_H = 0.048$) and with the use of LLMs for brainstorming ($V = 0.235$) and stakeholder communication ($V = 0.214$): respondents whose key stakeholders are in their home country report fewer LLM-supported tasks (mean 3.53 vs. 4.35 with US stakeholders) and use LLMs less often for brainstorming (31.5% vs. 61.9%) and stakeholder communication (51.5% vs. 75.0%). Similarly, key-decisions location was associated with the number of LLM-supported tasks ($\eta^2_H = 0.050$), which is lowest where key decisions are made domestically (mean 3.61 vs. 4.52 where decision-making is distributed), and with LLM use for wireframe/mockup creation ($V = 0.191$), which is most common where there is no clear central decision-making location (43.6%) and least common where decisions are made abroad (12.8%). A plausible mechanism for the stakeholder-location associations is language: respondents working with foreign stakeholders typically communicate in English, a second language for most of them, and may use LLMs to draft, polish and translate messages and elicitation materials. Our task list did not include translation, so this explanation cannot be tested with the present data.

In Ukraine, company type was associated with LLM use for brainstorming ($V = 0.209$), which is least common in Non-IT companies (23.8%, vs. 45.4% in Product and 50.9% in Service companies). A similar association was observed for stakeholder communication ($V = 0.197$), which is most common in Service companies (71.8%, vs. 49.5% in Product and 53.6% in Non-IT). LLM use for knowledge

search increased as BA practices moved from organization-wide standards towards project-level adaptation ($\tau_b = 0.166$).

Four further Ukrainian associations do not remain significant after correction across all pairs: the extent to which BA practices were shaped by individual experience ($\tau_b = 0.117$) or by team practices ($\tau_b = 0.113$), as well as company type ($\eta^2_H = 0.021$), was associated with the number of LLM-supported tasks, and role group was associated with the number of reported concerns ($\eta^2_H = 0.024$), which is lowest in the analyst-deepening group (mean 2.15 vs. 2.52 for pure analysts and 2.56 for cross-functional respondents). These associations should therefore be interpreted with caution.

No individual concern was associated with any context factor after correction in either country. Most of the associations reported above were observed only in the Ukrainian sub-sample. This may reflect substantive contextual differences, but it may also reflect the substantially smaller Polish sub-sample and the resulting lower statistical power; hence, absence of a significant result in one country should not be interpreted as evidence that the association is absent in that country.

4.5.3 Organizational Readiness Associations

One association directly concerns organizational AI readiness. In Ukraine, company size was associated with the reported availability of company AI solutions ($V = 0.149$). The cross-tabulation shows a clear gradient at the ends of the size scale: in companies with up to 249 employees, 39.1% of respondents with LLM experience report no AI solutions and no plans and 10.3% report BA-specific solutions, whereas in companies with over 1,500 employees these shares are 12.7% and 28.3%, respectively. Management support was not associated with any context factor after correction (smallest $p_{\text{adj(BH)}} = 0.067$, for company size). Accordingly, the present results support a specific association between company size and the organizational availability or planning of AI solutions, rather than a broader claim about management support.

4.5.4 Multivariable Analysis

Table 12 reports the country-specific logistic regressions (Section 3.4). In Ukraine, nine of the 36 coefficients are significant at $p < 0.05$; after Benjamini–Hochberg correction within each model, only the three Copilot coefficients remain, and none survives correction across all 36 coefficients (smallest $p_{\text{adj(BH)}} = 0.060$). Copilot use is more likely in larger companies (OR = 1.33 per size category, 95% CI $[1.08, 1.63]$), in more plan-driven projects (OR = 1.67, $[1.20, 2.32]$) and among respondents involved in more BA areas (OR = 1.31, $[1.09, 1.57]$). Larger companies also show lower odds of regular LLM use (OR = 0.83, $[0.69, 0.99]$) and of ChatGPT use (OR = 0.72, $[0.56, 0.93]$); these and the remaining nominally significant coefficients do not survive correction and are treated as exploratory. The models explain a modest share of the variation (McFadden pseudo-R^2 0.035–0.123; AUC 0.63–0.73), as is common for observational survey data. The Gemini model as a whole is not significant (likelihood-ratio $p = 0.056$), but it is significant when company size is coded categorically (a post hoc respecification) ($p = 0.001$), owing to higher Gemini use in companies with 250–800 employees. Adding to the Copilot model an indicator of whether the company provides generic AI tools

(a post hoc extension) hardly changes the company-size coefficient (OR $1.33 \rightarrow 1.29$), although the indicator itself is associated with Copilot use (OR $= 1.85$, $p = 0.021$).

In Poland, the Firth models point in the same direction as the Ukrainian models for company size: larger companies have lower odds of regular LLM use (OR $= 0.65$, $[0.45, 0.93]$, $p = 0.019$) and higher odds of Copilot use (OR $= 1.69$, $[1.09, 2.81]$, $p = 0.018$). The coefficient for BA activity breadth also points in the same direction (OR $= 1.46$, $[1.01, 2.19]$, $p = 0.047$), and that for development approach (OR $= 1.44$) points in the same direction but is not significant, and the experience effect on regular usage is close to null (OR $= 0.98$).

BA activity breadth is associated with more LLM-supported tasks in both countries (Poisson IRR: Ukraine 1.09, $[1.04, 1.13]$; Poland 1.18, $[1.07, 1.29]$; both $p < 0.001$) and with more reported concerns in Ukraine only (IRR $= 1.07$, $[1.02, 1.12]$, $p = 0.009$; Poland IRR $= 0.98$, $p = 0.70$). The counts are underdispersed relative to the Poisson distribution, so the Poisson standard errors are conservative, and the quasi-binomial models lead to the same conclusions for BA activity breadth.

5 Discussion

This section interprets the key findings in relation to the existing literature and theory, identifies emergent patterns, and discusses the study's limitations.

5.1 The LLM Adoption Tension in BA

In our sample, 80.8% of respondents use LLMs in their BA work regularly or occasionally, and 89.8% have at least tried them. The closest related studies report lower shares of respondents using LLMs in their professional work: 66.7% in Santos et al. [27] and Pir Muhammad et al. [26], 58.2% in Rani et al. [28], and 66% for requirements specification, the most LLM-intensive RE area, in Hess et al. [29]. These figures are not directly comparable: the questions differ (any professional use vs. use for BA tasks), the samples are small and were recruited differently, and the surveys were conducted at different times in a period of rapid LLM uptake. Our convenience sampling through professional communities likely overrepresents early adopters, as we acknowledge in Section 3.5. What our data do indicate is that LLM usage is widespread among professionally active BA practitioners – a finding that holds regardless of the exact population estimate.

This creates what we term the BA adoption tension: a professional community has widely adopted a transformative technology while methodological guidance and evidence from practice remain limited. The speed and scale of LLM adoption make this tension particularly acute in the BA context. The BABOK Guide [14] does not address LLM-assisted practices. IIBA has published practitioner-oriented material on AI in business analysis [97], but no standard-level guidance, and the first syllabus on AI in requirements engineering by IREB, which also targets business analysts, appeared only as a beta version in November 2025 and in its final version in March 2026, after our data collection [98]. As a result, business analysis professionals who incorporate AI into their workflows often do so without institutional support – a situation that calls for both empirical research and the development of practical recommendations.

Table 12: Multivariable logistic regression by country: odds ratios [95% CI]

Predictor	Regular usage	ChatGPT	Gemini	MS Copilot
Panel A. Ukraine: logistic regression				
Company size	0.83°	0.72°	0.89	**1.33**
	[0.69, 0.99]	[0.56, 0.93]	[0.74, 1.07]	[1.08, 1.63]
Company type: Product	1.73	0.97	0.95	0.60
	[0.93, 3.22]	[0.46, 2.07]	[0.50, 1.80]	[0.29, 1.22]
Company type: Service	1.60	1.55	0.76	1.18
	[0.91, 2.82]	[0.78, 3.08]	[0.43, 1.36]	[0.64, 2.17]
BA experience	0.78°	0.84	1.11	1.07
	[0.63, 0.97]	[0.64, 1.12]	[0.89, 1.38]	[0.84, 1.37]
Development approach	0.87	1.14	1.03	**1.67**
	[0.64, 1.16]	[0.77, 1.68]	[0.77, 1.40]	[1.20, 2.32]
Industry: Finance	1.30	0.50	1.30	0.57
	[0.70, 2.44]	[0.24, 1.03]	[0.68, 2.52]	[0.28, 1.17]
Industry: Regulated other	0.83	3.71°	0.46°	1.32
	[0.41, 1.66]	[1.14, 12.05]	[0.23, 0.93]	[0.62, 2.79]
Industry: Commercial	1.19	1.90	0.54°	0.94
	[0.66, 2.13]	[0.87, 4.16]	[0.30, 0.99]	[0.49, 1.80]
BA activity breadth	1.08	1.07	1.04	**1.31**
	[0.93, 1.26]	[0.87, 1.30]	[0.89, 1.22]	[1.09, 1.57]
n (outcome present)	375 (162)	344 (268)	344 (183)	344 (116)
Smaller outcome class per model term	18.0	8.4	17.9	12.9
McFadden pseudo-R^2 / AUC	0.041 / 0.63	0.123 / 0.73	0.035 / 0.63	0.095 / 0.71
Likelihood-ratio $\chi^2(9)$ p	0.012	< 0.001	0.056	< 0.001
Panel B. Poland: Firth-penalized logistic regression (direction check of Panel A effects)				
Company size	0.65°	–	–	1.69°
	[0.45, 0.93]			[1.09, 2.81]
BA experience	0.98	–	–	–
	[0.66, 1.48]			
Development approach	–	–	–	1.44
				[0.61, 3.48]
BA activity breadth	–	–	–	1.46°
				[1.01, 2.19]
n (outcome present)	94 (46)	not modeled	not modeled	77 (29)

Note. Panel A: maximum-likelihood logistic regression; company size (0–3), BA experience (0–3), development approach (0 Agile, 1 Hybrid, 2 Plan-driven) and BA activity breadth (0–6) entered as linear scores; reference categories: Non-IT company, "Other" industry. Panel B: Firth-penalized models with the predictors nominally significant in Panel A (– = not included) and profile penalized-likelihood CIs; ChatGPT and Gemini are not modeled for Poland (Section 3.4). Odds ratios in **bold**: significant after Benjamini–Hochberg correction within the model ($p_{\text{adj(BH)}} < 0.05$); none remains significant after correction across all 36 coefficients of Panel A. ° Nominal $p < 0.05$ not significant after correction; such coefficients are treated as exploratory. n (outcome present): all respondents for regular usage, respondents with LLM experience for the tool models. Smaller outcome class per model term: the less frequent outcome class (users or non-users) divided by the nine model terms (events per variable [93]).

5.2 The Distribution of LLM Use Across BA Tasks

The task-level adoption data reveal a pattern that seems to map coherently onto the nature of LLM capabilities. Requirements drafting (72.2%) and knowledge search (71.3%) represent tasks where LLMs function as text generation and information synthesis engines – precisely the capabilities that current models excel at. Stakeholder communication (62.7%) and document analysis (56.3%) involve tasks where LLMs assist with text transformation: summarizing meeting notes, restructuring email drafts, and extracting key points from contracts or RFPs. Brainstorming, which is used less frequently (43.9%), also benefits from the information synthesis capabilities of LLMs. Visual modeling support (33.0%), Jira/Confluence generation (27.1%), and wireframe creation (24.7%) involve tasks that require either visual output or integration with specific toolchains (Jira, Confluence). At the time of data collection, support for such tasks was less mature and less often available in the general-purpose chat interfaces that most respondents used, which may partly explain the lower rates; our data, however, do not show why these tasks are supported less often.

This ranking maps onto the BABOK knowledge areas [14]. Requirements Analysis and Design Definition and Elicitation and Collaboration emerge as the most LLM-augmented activity domains; they are, however, also the two knowledge areas in which most respondents work (87–94%; Table 4). The task list did not include activities specific to Solution Evaluation, Business Analysis Planning and Monitoring or Strategy Analysis, so our data cannot show whether LLMs are used less in these areas. Task-level usage partly mirrors involvement: in Ukraine, requirements drafting is reported by 74% of respondents with LLM experience involved in Requirements Analysis and Design Definition versus 60% of those who are not, and brainstorming elicitation questions by 44% of those involved in Elicitation and Collaboration versus 25% of those who are not (43 and 20 respondents outside these areas). This pattern suggests that LLM adoption in BA follows task–capability fit only in part; it also reflects what BA practitioners do most often.

Some of the related studies asked similar questions. Hess et al. [29] inquired about the RE sub-areas in which respondents use LLMs and about the perceived usefulness or harmfulness of LLM use in particular RE tasks. Due to a different structure of questions and available responses, a direct comparison between their findings and ours is not possible, but some observations can be made. Their results on the RE sub-areas supported by LLMs varied from requirements specification (66%), requirements elicitation (58%) and "RE in general" (58%) to requirements management (14%). Since the tasks most frequently supported by LLMs in our study can be assigned to BABOK areas that roughly correspond to requirements elicitation and specification, the findings of both studies corroborate each other. Also, the tasks for which LLM use was perceived as most useful by the participants of Hess et al. [29] included the preparation of surveys and interviews and of workshops and focus groups (corresponding to "Brainstorming elicitation questions or techniques" in our study), as well as creating document structures, drafting specifications and creating use case descriptions (corresponding to "Drafting requirements (user stories, use cases, specs)"). However, the respondents of Hess et al. [29] also rated the creation of diagrams as highly useful, whereas "Visual modeling support" is much less frequent in our study. Santos et al. [27]

report that 75% of their study participants use LLMs to generate user stories, and mention support for creation and brainstorming as another LLM use case (without a reported share). These findings correspond to the requirements drafting and brainstorming use cases in our study. Further comparisons are not possible, as Santos et al. [27] focus on applications of LLMs to user stories and report other LLM-supported tasks based on a small number of open-text answers. The remaining related studies do not address LLM-supported tasks.

5.3 Accuracy Concerns Coexist with Use

Accuracy and hallucinations are the most frequently selected concern (82.7% of respondents with LLM experience). The concern is reported at similar rates by respondents who have tried LLMs but do not use them regularly, occasional users and regular users (78.6%, 80.1% and 85.6%; Kendall's $\tau_b = 0.07$, $p = 0.12$), and respondents who report it rate LLM effectiveness similarly to those who do not (mean 3.64 vs. 3.58). Concern and use therefore coexist. Our data cannot show whether accuracy concerns deter adoption: the 48 respondents without LLM experience, among whom such an effect would be most likely, were not asked about concerns, and the high prevalence of the concern leaves limited room for differences between groups. We interpret this coexistence as possible evidence of adaptive behavior, although our data cannot test it directly. Practitioners may have developed informal verification heuristics – cross-checking LLM outputs against domain knowledge, treating generated text as a draft rather than a deliverable, and applying professional judgment as a post-generation filter. This interpretation aligns with Barke et al.'s [7] finding that developers interact with Copilot in an "acceleration" mode, using AI-generated suggestions as a starting point rather than a finished product. It is worth noting that our respondents do not expect LLMs to fully take the place of human BA practitioners, but only to take over some (probably simpler) tasks. Thus, a human-in-the-loop setting, where final decisions are made by humans who critically consider LLMs' input (among other things), seems to be a necessary approach. This also corresponds with Pir Muhammad et al.'s [26] finding that practitioners use LLMs as an aid rather than an authority for RE tasks, but consider LLM input useful in their own decision-making.

The concerns about LLM use voiced by our respondents show similarities to those reported in related work, but their perceived significance among surveyed practitioners is not necessarily the same. The study by Hess et al. [29] includes a list of limitations and threats related to LLMs, and the ones selected as relevant by most respondents were: hallucinations, over-reliance on AI, ethical and legal concerns, lack of transparency and explainability, loss of human intuition and judgment, and confidentiality. The top concern is thus the same as in our findings; among the others, ethical and legal concerns and confidentiality can be mapped to our results, but their relative positions are different. Santos et al. [27] mention only two concerns ("negative impacts") related to LLM use specifically in the task of user story generation: the difficulty LLMs have in dealing with specific project context and the risks associated with excessive dependence of humans on AI. The first corresponds to our "Lack of domain knowledge in LLM", while the second can be considered part of "Ethical concerns", in particular role replacement. Rani et al. [28] report six challenges of using AI in RE,

assessed by study participants. Sorted from the most to the least frequently selected, they are: knowledge and cognitive limitations, technical and implementation barriers, quality and accuracy concerns, governance, security and compliance challenges, communication and human interaction challenges, and data and methodological challenges. It should be noted that in their study the general perception of challenges is rather low: the most frequent challenge was confirmed by only 15 of 55 participants. The three top challenges correspond, at least to some extent, to our concerns related to lack of domain knowledge in LLM, limited integration with existing BA tools, and accuracy/hallucinations, respectively. Thus, once again there is a partial overlap, with differences in the relative positions of the challenges. As no standardized list of concerns, challenges or threats was used in the studies discussed here, direct comparisons are difficult. Often a one-to-one mapping is not possible; for example, "Ethical concerns (role replacement, bias)" in our study includes items listed as separate concerns by others (such as excessive dependence of humans on AI). It is also quite possible that the exact phrasing of a concern influences its perceived significance. This, among other factors such as different populations, could explain the differences in the evaluations of concerns reported by particular studies, but at the moment it is only a hypothesis that would require dedicated research.

5.4 The Organizational Readiness Gap

Among respondents with LLM experience, willingness to learn AI tools (UA 93.6%, PL 92.2%) is reported far more often than strong management support (UA 54.7%, PL 55.8%) or BA-specific company AI solutions (UA 21.2%, PL 20.8%). Because these indicators measure different constructs on different response scales, the comparison of the three shares is descriptive and depends on where the scales are cut: strong or moderate support is reported by 82.6% in Ukraine and 76.6% in Poland, and by some company-provided AI tool by 51.7% and 57.1%, respectively. The paired comparison is more direct: 41.0% of Ukrainian and 39.0% of Polish respondents with LLM experience are willing to learn without reporting strong management support, whereas only 2.0% and 2.6% report the reverse (exact McNemar test, $p < 0.001$ in both countries).

At the individual level, the high willingness to learn, with 44.8% of Ukrainian and 48.1% of Polish respondents with LLM experience already dedicating time to it, suggests that the adoption bottleneck does not reside in practitioner resistance. At the organizational level, more than half of respondents with LLM experience report strong management support, but this drops sharply when measured by concrete investment: besides the BA-specific AI solutions reported by about one in five respondents with LLM experience, 30.5% in Ukraine and 36.4% in Poland have AI tools that are not BA-specific (e.g., general-purpose Copilot licenses), while the remaining respondents either plan to adopt such solutions (16.0% and 6.5%), have no implementation plans (22.1% and 26.0%), or are uncertain (10.2% and 10.4%).

The studies we discussed as the most closely related work do not extensively investigate organizational readiness. Hess et al. [29] asked their respondents about the reasons preventing the use of AI in RE, including "Lack of support / tools", "Cultural resistance and organizational inertia" and "Using AI is forbidden by my organization". These reasons were selected relatively rarely: the first by 23 of 57 respondents,

the remaining ones by 13 and 14, respectively. There was, however, no explicit question about the level of organizational support. Santos et al. [27] do not pursue this issue either. Their paper only mentions that the barriers preventing professionals from using LLMs in RE include "concerns about privacy, information security, and compliance with the LGPD, factors especially relevant in organizational environments with stricter institutional policies". This statement is not backed by numerical data, so no conclusions about organizational readiness can be drawn from their results. Rani et al. [28] asked their participants about responsible AI practices in their organizations and found that these mostly consist of individual actions, such as reviews of LLM-generated content, which "indicates reactive rather than proactive governance, with practitioners prioritizing immediate oversight over systematic risk management and auditability". They conclude that this governance gap represents a significant organizational risk that should be addressed at the organizational level.

In Ukraine and Poland, the gap may be compounded by the high share of service company models (outsource and outstaff: 46.9% of respondents in Ukraine and 67.0% in Poland), where client confidentiality constraints can limit the use of external LLM tools. A stronger GDPR-related effect could be expected in Poland, but confidentiality concerns are reported at the same rate in both countries (66.3% and 67.5%), so our data do not support this explanation. Beyond the GDPR, Article 4 of the AI Act, applicable since 2 February 2025, requires providers and deployers of AI systems in the EU to ensure a sufficient level of AI literacy among their staff [31]. This would lead one to expect more organizational training and support among Polish employers, yet strong management support was reported at almost identical rates in both countries (55.8% vs. 54.7%). At the same time, Ukrainian service companies working for EU clients are often bound by GDPR obligations through data processing agreements, which weakens the regulatory contrast between the two samples.

In Ukraine, company size is associated with the availability of company AI solutions ($V = 0.149$, $p_{\text{adj(BH)}} = 0.009$), whereas management support showed no association after correction. In the Ukrainian multivariable models, larger companies show higher odds of Copilot use and, at the nominal level only, lower odds of regular LLM use and of ChatGPT use, which is consistent with smaller companies relying more on individual initiative and general-purpose tools, while larger ones provide tools through corporate licensing.

5.5 Context and AI Tool Choice

The 25 Ukrainian associations and the one Polish association that remained significant after Benjamini–Hochberg correction within outcome families (out of 450 pairs per country) suggest that some aspects of LLM tool choice and use vary with project and organizational context. However, the associations are few and small (in Ukraine, Cramér's $V \leq 0.30$ and $|\tau_b| \leq 0.201$; the Polish association reaches $\tau_b = 0.33$), and the measured context explains little of the variation.

Three context–tool pairs illustrate this structure most clearly. In Ukraine, ChatGPT use is lowest in Finance (57.6% of respondents with LLM experience, vs. 78.7–93.0% in the other industry groups); in the multivariable model, the Finance coefficient points in the same direction but is not significant (OR = 0.50 relative to

the "Other" category, $p = 0.061$). Respondents in other regulated sectors have higher odds of using ChatGPT (OR = 3.71, 95% CI $[1.14, 12.05]$), but this coefficient does not survive correction.

In the Ukrainian models, larger company size is associated with lower odds of ChatGPT use (OR = 0.72) and higher odds of Copilot use (OR = 1.33), and Copilot use is also more common in plan-driven environments (OR = 1.67). These Copilot coefficients remain significant after correction within the model, and the company-size effect on Copilot points in the same direction in Poland (OR = 1.69, $p = 0.018$).

Together, these patterns suggest that tool choice is associated with the organizational setting in which a project operates. Enterprise licensing (e.g., Microsoft 365 including Copilot) is a plausible mechanism: in Ukraine, Copilot use is higher where the company provides generic AI tools (OR = 1.85, $p = 0.021$), but adding this indicator hardly changes the company-size effect (OR $1.33 \rightarrow 1.29$), so licensing does not explain why Copilot use is higher in larger companies. However, the modest pseudo-R^2 values (0.035–0.123) for the Ukrainian regression models indicate that the contextual predictors we measured explain only a limited share of the variance, suggesting that unmeasured individual- and team-level factors also play a role. Whether such contextual differences matter for the quality of BA work remains to be tested.

As none of the related studies applied an analytical approach similar to ours to test associations between variables, it is not possible to position our results relative to those studies.

5.6 Connecting to Technology Adoption Theory

Although the study was not designed around a technology acceptance model, two findings can be related to UTAUT [60]. Our management-support item combines two aspects that UTAUT separates: encouragement by management, which UTAUT measures as part of Social Influence (e.g., "The senior management of this business has been helpful in the use of the system"), and the provision of resources, which corresponds to Facilitating Conditions. Company AI solutions correspond more clearly to Facilitating Conditions. BA-specific facilitating conditions are rare in our sample (21.1% of respondents with LLM experience report BA-specific company AI solutions), although about half have access to some company-provided AI tools. At the same time, LLM use is widespread: for consumer-grade tools such as ChatGPT and Gemini, individual access may bypass the organizational infrastructure that usually precedes adoption. This pattern resembles the "bring your own AI" phenomenon reported for knowledge workers in general: in the 2024 Work Trend Index (31,000 respondents in 31 countries), 75% of knowledge workers used AI at work and 78% of AI users brought their own AI tools to work [99]. Our instrument did not ask whether the tools used were provided or approved by the employer, so this interpretation remains a hypothesis. Testing these relationships would require validated TAM/UTAUT scales.

5.7 Limitations

In addition to the threats to validity discussed in Section 3.5, several other limitations should be mentioned. The cross-sectional design captures a snapshot of a rapidly

evolving area: the landscape of LLMs changes substantially every 6–12 months. Consequently, the adoption patterns observed in early 2026 may shift as new models, tools, and organizational policies emerge. At the time of data collection, ChatGPT offered GPT-5.2 (Instant, Thinking, and Pro), Gemini provided Gemini 3 Flash and Gemini 3 Pro, Microsoft 365 Copilot Chat supported GPT-5.2, and Jira/Confluence offered AI-assisted search, content generation, and work-item management through Atlassian Rovo [100–103]. However, access to these capabilities varied by subscription, organizational configuration, and rollout stage. Because product names such as ChatGPT or Copilot denote different models and features at different times and for different users, the meaning of our tool items is time-bound. Consequently, the reported adoption of a named product should not be interpreted as evidence of respondents' access to, or use of, any particular model version or feature. This also limits comparisons with surveys conducted at other times. The geographical area (Ukraine and Poland) limits the generalizability of the results: these markets share common characteristics, e.g., a developed IT outsourcing sector, high English proficiency, and distributed teams, that may reinforce certain adoption patterns compared to, for instance, Western European markets. The convenience sampling strategy likely introduces upward bias in adoption rates, as practitioners active in professional communities may be early adopters. The multivariable models are consistent in direction with several bivariate associations, but only three coefficients remain significant after correction within the model, and their modest explanatory power (McFadden pseudo-R^2 0.035–0.123 in the Ukrainian models) indicates that the contextual variables we measured account for only a small portion of the variance in adoption. Unmeasured factors – including individual AI literacy, team-level social norms, client-imposed restrictions, and personal technology preferences – likely play substantial roles that our current design cannot capture. Respondents without LLM experience ($n = 48$) were not asked about attitudes, management support or company AI solutions. Because practitioners in organizations that discourage or prohibit AI use are likely to be overrepresented among them, the reported levels of organizational support should be read as upper bounds that describe respondents with LLM experience only. The associative and regression analyses used in this study allow us to identify and assess statistical relationships but cannot establish causality. Other context factors not included in our models may mediate or confound the observed associations between context and tool choice.

6 Conclusions

6.1 Summary of Findings

This study presents one of the first large-scale, quantitative surveys of LLM adoption in business analysis practice. A total of 469 respondents were surveyed, with 375 from Ukraine and 94 from Poland. The study provides empirical answers to four research questions:

RQ1: What is the current level and pattern of LLM adoption among BA practitioners in Ukraine and Poland? LLM use is high among the surveyed practitioners: 80.8% of respondents actively use LLMs (UA 82.4%, PL 74.5%), 44.3%

regularly (UA 43.2%, PL 48.9%) and 36.5% occasionally (UA 39.2%, PL 25.5%); only 3.8% reject LLM use entirely (UA 3.2%, PL 6.4%). Among respondents with LLM experience, ChatGPT dominates with 79.1% usage (UA 77.9%, PL 84.4%), followed by Google Gemini with 50.6% usage (UA 53.2%, PL 39.0%) and Microsoft Copilot with 34.4% usage (UA 33.7%, PL 37.7%). Tool preferences vary by context: in Ukraine, ChatGPT use is lower in larger companies and is lowest in Finance sector, while Copilot use is higher in larger companies. Gemini is used less often in Poland, but this difference is tentative.

RQ2: Which BA tasks are LLMs used for most and least, and what concerns about their use do practitioners report? In both countries, the tasks most often selected for LLM use are requirements drafting (UA 72.4%, PL 71.4% of respondents with LLM experience), and knowledge search (69.5%, 79.2%), followed by stakeholder communication (61.0%, 70.1%), document analysis (56.7%, 54.5%) and brainstorming of elicitation questions (42.7%, 49.4%). Visual modeling (33.4%, 31.2%), Jira/Confluence generation (27.3%, 26.0%) and wireframe creation (24.7%, 24.7%) are reported by a minority. Thus, tasks involving producing and transforming text are used most frequently, although the ranking also reflects the activities in which respondents are most often involved. Accuracy and hallucinations are the dominant concerns (UA 82.8%, PL 81.8%), followed by confidentiality and data security (66.3%, 67.5%). The concern shares do not differ significantly between the countries. After correction, no individual concern is associated with project context in either country. However, in Ukraine, the number of reported concerns increases with the breadth of BA activity ($\tau_b = 0.163$, $p_{\text{adj(BH)}} = 0.002$).

RQ3: What attitudes do BA practitioners hold toward AI's role in their profession, and how ready are their organizations for AI adoption? Among respondents with LLM experience, the willingness to learn AI tools (UA 93.6%, PL 92.2%) is reported far more often than strong management support (54.7% UA, 55.8% PL) or BA-specific company AI solutions (21.2% UA, 20.8% PL; a further 30.5% and 36.4% have generic AI tools only). Notably, 41.0% of Ukrainian and 39.0% of Polish respondents with LLM experience are willing to learn without reporting strong management support, whereas only 2.0% and 2.6% report the reverse. The imbalance between these discordant groups was statistically significant in both countries according to the two-sided exact McNemar tests, both $p < 0.001$, indicating that practitioners are more ready than their organizations for AI adoption. Regarding AI's future role, most respondents with LLM experience believe that AI will partially replace BA tasks (UA 60.8%, PL 57.1%). About a third view it as a support tool only (UA 33.4%, PL 32.5%), and few expect AI to replace the BA role entirely (UA 4.1%, PL 9.1%).

RQ4: Which project and organizational context factors are associated with variations in LLM adoption and usage patterns? In Ukraine, 25 of the 450 context–outcome pairs remained significant after Benjamini–Hochberg correction within outcome families (17 of these pairs were also significant after the correction

across all pairs). In Poland, only one pair remained significant, and none after the correction across all pairs. Most associations are small: the largest associations are between industry and ChatGPT use ($V = 0.300$ in Ukraine) and between BA activity breadth and the number of LLM-supported tasks ($\tau_b = 0.201$ in Ukraine and 0.334 in Poland). The latter is the only association found in both countries. In Ukrainian multivariable models, Copilot use is associated with larger companies, plan-driven approaches, and broader BA activity breadth. These three effects remain significant after correction within the model. The associations of ChatGPT use with smaller companies and with other regulated sectors point towards the same direction as the bivariate results, but they do not survive correction. In Poland, the penalized Copilot model shows coefficients of the same sign for company size, development approach and BA activity breadth, but the estimates are imprecise and only the coefficients for company size and BA activity breadth reach $p < 0.05$. The modest explanatory power of the Ukrainian models (McFadden pseudo-R^2 0.035–0.123) indicates that unmeasured individual- and team-level factors also play a role.

6.2 Implications for Research and Practice

The fact that 82.7% of respondents with LLM experience express concerns about accuracy, at all levels of use, suggests that systematic verification procedures may be needed. This is especially important for requirements artifacts because errors in them can propagate to subsequent stages of the software development process. Since the effectiveness of such procedures was not evaluated in this study, the recommendations in this section are proposals for practice derived from the survey findings rather than empirically validated solutions.

The contrast between practitioners' willingness to learn AI tools (93.3% of respondents with LLM experience) and the level of strong management support (54.9%), with approximately 40% of respondents with LLM experience in each country willing to learn without strong support, suggests room for organizational support. Formalized AI adoption policies, clear guidelines on data confidentiality when using external LLM tools, and investments in BA-specific AI solutions (currently reported by only 21.1% of respondents with LLM experience) could help bridge the readiness gap.

For tool developers and vendors, our data offer limited but concrete signals. Limited integration with existing BA tools is reported as a concern by about a third of respondents with LLM experience (UA 30.5%, PL 33.8%), and the tasks tied to specific toolchains or visual output (visual modeling support, Jira/Confluence generation and wireframe creation) are the least frequently supported by LLMs. Integrating LLM support into the tools that business analysts already use, rather than relying on stand-alone chat interfaces, may therefore address a reported need. Tool choice is also associated with organizational context, mainly industry and company size, but these associations are small and may partly reflect the policies of individual employers, so they do not by themselves justify context-specific product strategies.

For researchers, the main implication follows from the weak explanatory power of the measured context: the associations are few and small, and the Ukrainian multivariable models explain only a small share of the variation in adoption (McFadden pseudo-R^2 0.035–0.123). Studies of LLM adoption in BA should therefore measure

the individual- and team-level factors that our instrument did not capture, such as AI literacy, team norms, client-imposed restrictions and personal technology preferences, and should treat the contextual associations reported here as hypotheses to be tested rather than as established determinants (Section 6.3).

6.3 Future Research

This work suggests several directions. First, our survey shows what practitioners do with LLMs, but not how or why. Qualitative studies, such as interviews or diary studies with BA practitioners, should examine the mechanisms the survey only signals: how practitioners verify LLM output before it enters requirements artifacts, how client and employer policies constrain tool choice, and how often consumer tools are used outside organizational provision ('shadow AI'). Such studies would also show whether the proposed verification procedures match practitioners' actual needs.

Second, the associations identified in RQ4 should be tested in a confirmatory study, preferably preregistered (e.g., as a registered report), with the expected directions specified in advance and an a priori power analysis based on the effect sizes observed here. Candidate hypotheses include: (H1) respondents involved in more BA knowledge areas use LLMs for more BA tasks, the only association found in both countries; (H2) Copilot use is more common in larger companies and (H3) in more plan-driven projects; (H4) respondents whose key stakeholders are located abroad use LLMs more often for stakeholder communication and for brainstorming elicitation questions; and (H5) larger companies more often provide company AI solutions.

A confirmatory study should also improve the measures used here. Adoption determinants should be measured with validated multi-item scales, for example, performance expectancy, social influence and facilitating conditions from UTAUT [60], or with constructs organized by the Technology–Organization–Environment framework applied by Neumann et al. [65], instead of the single items of our instrument. Concerns should be rated for intensity rather than only selected from a list, and the questions on concerns and organizational support should also be presented to respondents without LLM experience, among whom organizational barriers are most likely. Verification practices should be measured directly: how often LLM output is reviewed before it enters requirements artifacts, by whom, and against which sources. This would allow the coexistence of accuracy concerns and use reported here to be explained rather than only described. Tool questions should also distinguish between consumer and enterprise versions of the tools and ask whether each tool is provided or approved by the employer, thereby allowing the bring-your-own-AI interpretation discussed in Section 5.6 to be tested.

Third, a longitudinal study tracking the same population over 12–24 months would reveal how adoption patterns evolve as LLM capabilities advance and organizational policies mature. Fourth, quality-impact studies should measure whether LLM-assisted BA work produces better requirements artifacts, using established metrics such as ambiguity rates, completeness, and defect density. Controlled experiments comparing LLM-assisted and traditional task performance would provide the causal evidence that our correlational design cannot. Finally, geographic replication, which would extend the survey to Western European (e.g., DACH and the Nordics), North American,

and Asian BA communities, would test the generalizability of our context-dependency findings.

Acknowledgements. The contents of this section will be added after the review due to the rules of the double-blind review process.

Declaration of AI Assistance. We acknowledge that we used Generative AI for the following purposes, in addition to *AI-assisted copy editing*: development of Python data analysis scripts, drafting the LaTeX code of the figures included in the article, translation of tables into LaTeX format. All AI-generated suggestions were reviewed and edited by the authors, who take full responsibility for the final content of this article.

References

[1] Dwivedi, Y.K., Kshetri, N., Hughes, L., Slade, E.L., Jeyaraj, A., Kar, A.K., Baabdullah, A.M., Koohang, A., Raghavan, V., Ahuja, M., *et al.*: Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management **71**, 102642 (2023) https://doi.org/10.1016/j.ijinfomgt.2023.102642

[2] Bendig, D., Bräunche, A.: The role of artificial intelligence algorithms in information systems research: a conceptual overview and avenues for research. Management Review Quarterly, 2863–2908 (2024) https://doi.org/10.1007/s11301-024-00451-y

[3] Russo, D.: Navigating the complexity of generative AI adoption in software engineering. ACM Transactions on Software Engineering Methodology **33**(5) (2024) https://doi.org/10.1145/3652154

[4] Storey, V.C., Yue, W.T., Zhao, J.L., Lukyanenko, R.: Generative Artificial Intelligence: Evolving Technology, Growing Societal Impact, and Opportunities for Information Systems Research. Information Systems Frontiers **27**, 2081–2102 (2025) https://doi.org/10.1007/s10796-025-10581-7

[5] Bracken, A., Reilly, C., Feeley, A., Sheehan, E., Merghani, K., Feeley, I.: Artificial Intelligence (AI) – Powered Documentation Systems in Healthcare: A Systematic Review. Journal of Medical Systems **49**(1), 28 (2025) https://doi.org/10.1007/s10916-025-02157-4

[6] Triando, Simaremare, M., Wang, X., Prasad, A.S.R.: The use of generative AI tools in the inception stage of software startups. In: Papatheocharous, E., Farshidi, S., Jansen, S., Hyrynsalmi, S. (eds.) Software Business, pp. 439–453. Springer, Cham (2025)

[7] Barke, S., James, M.B., Polikarpova, N.: Grounded Copilot: How programmers interact with code-generating models. Proceedings of the ACM on Programming Languages **7**(OOPSLA1), 85–111 (2023)

[8] Feldt, R., Kang, S., Yoon, J., Yoo, S.: Towards autonomous testing agents via conversational large language models. In: 2023 38th IEEE/ACM International Conference on Automated Software Engineering (ASE), pp. 1688–1693 (2023). IEEE

[9] Wang, J., Huang, Y., Chen, C., Liu, Z., Wang, S., Wang, Q.: Software testing with large language models: Survey, landscape, and vision. IEEE Transactions on Software Engineering **50**(4), 911–936 (2024)

[10] Sami, M.A., Zhang, Z., Waseem, M., Kemell, K.-K., Rasheed, Z., Herda, T., Abrahamsson, P.: Bridging humans and LLMs: Investigating Human-AI collaboration in multi-agent requirements analysis for organizational AI adoption. e-Informatica Software Engineering Journal **20**(1), 260103 (2026) https://doi.org/10.37190/e-Inf260103

[11] Holzmann, V., Zitter, D., Peshkess, S.: The expectations of project managers from artificial intelligence: A delphi study. Project Management Journal **53**(5), 438–455 (2022)

[12] Nguyen-Duc, A., Cabrero-Daniel, B., Przybylek, A., Arora, C., Khanna, D., Herda, T., Rafiq, U., Melegati, J., Guerra, E., Kemell, K.-K., *et al.*: Generative artificial intelligence for software engineering—a research agenda. Software: Practice and Experience **55**(11), 1806–1843 (2025) https://doi.org/10.1002/spe.70005

[13] Jarzębowicz, A., Przybyłek, A., Estima, J., Ng, Y.Y., Swacha, J., Zielosko, B., Madeyski, L., Carroll, N., Kemell, K.-K., Marcinkowski, B., Silva, A., Stray, V., Iivari, N., Nguyen-Duc, A., Melegati, J., Delibašić, B., Insfran, E.: The landscape of generative AI in information systems: A synthesis of secondary reviews and research agendas. e-Informatica Software Engineering Journal **21**(1), 270101 (2027) https://doi.org/10.37190/e-inf270101

[14] International Institute of Business Analysis: A Guide to the Business Analysis Body of Knowledge (BABOK Guide), Version 3 edn. International Institute of Business Analysis (IIBA), Toronto, ON (2015)

[15] Görer, B., Aydemir, F.B.: GPT-powered elicitation interview script generator for requirements engineering training. In: 2024 IEEE 32nd International Requirements Engineering Conference (RE), pp. 372–379 (2024). IEEE

[16] Zhang, J., Chen, Y., Liu, C., Niu, N., Wang, Y.: Empirical evaluation of ChatGPT on requirements information retrieval under zero-shot setting. In: 2023 International Conference on Intelligent Computing and Next Generation

Networks (ICNGN), pp. 1–6 (2023). IEEE

[17] Ronanki, K., Cabrero-Daniel, B., Horkoff, J., Berger, C.: Requirements engineering using Generative AI: Prompts and prompting patterns. In: Generative AI for Effective Software Development, pp. 109–127. Springer, Cham (2024)

[18] Sami, M.A., Rasheed, Z., Waseem, M., Zhang, Z., Herda, T., Abrahamsson, P.: Prioritizing software requirements using Large Language Models. arXiv preprint arXiv:2405.01564 (2024)

[19] Norheim, J.J., Rebentisch, E.: Structuring natural language requirements with Large Language Models. In: 2024 IEEE 32nd International Requirements Engineering Conference Workshops (REW), pp. 68–71 (2024). https://doi.org/10.1109/REW61692.2024.00013 . IEEE

[20] Wang, B., Wang, C., Liang, P., Li, B., Zeng, C.: How LLMs aid in UML modeling: An exploratory study with novice analysts. In: 2024 IEEE International Conference on Software Services Engineering (SSE), pp. 249–257 (2024). IEEE

[21] Krishna, M., Gaur, B., Verma, A., Jalote, P.: Using LLMs in software requirements specifications: An empirical evaluation. In: 2024 IEEE 32nd International Requirements Engineering Conference (RE), pp. 475–483 (2024). IEEE

[22] Abualhaija, S., Ceci, M., Sannier, N., Bianculli, D., Briand, L.C., Zetzsche, D., Bodellini, M.: AI-enabled regulatory change analysis of legal requirements. In: 2024 IEEE 32nd International Requirements Engineering Conference (RE), pp. 5–17 (2024). IEEE

[23] Fantechi, A., Gnesi, S., Passaro, L., Semini, L.: Inconsistency detection in natural language requirements using ChatGPT: a preliminary evaluation. In: 2023 IEEE 31st International Requirements Engineering Conference (RE), pp. 335–340 (2023). IEEE

[24] White, J., Hays, S., Fu, Q., Spencer-Smith, J., Schmidt, D.C.: ChatGPT prompt patterns for improving code quality, refactoring, requirements elicitation, and software design. In: Generative AI for Effective Software Development, pp. 71–108. Springer, Cham (2024)

[25] Ronanki, K., Berger, C., Horkoff, J.: Investigating ChatGPT's potential to assist in requirements elicitation processes. In: 2023 49th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), pp. 354–361 (2023). IEEE

[26] Pir Muhammad, A., Berntsson Svensson, R., Inayat, I.: Influence of LLM prioritizations on human decisions in requirements engineering. In: International Conference on Product-Focused Software Process Improvement, pp. 369–384 (2025). Springer

[27] Santos, R., Steinmacher, I., Conte, T., Oran, A., Gadelha, B.: Adoption of LLMs in requirements engineering: What practitioners are worried about? In: Anais do XXIV Simpósio Brasileiro de Qualidade de Software, pp. 248–258. SBC, Porto Alegre, RS, Brasil (2025). https://doi.org/10.5753/sbqs.2025.15089

[28] Rani, L.M., Svensson, R.B., Feldt, R.: AI for requirements engineering: Industry adoption and practitioner perspectives. In: 2025 40th IEEE/ACM International Conference on Automated Software Engineering Workshops (ASEW), pp. 244–251 (2025). IEEE

[29] Hess, A., Vogelsang, A., Franch, X., Herrmann, A., Kopczyńska, S., Rachmann, A.: Opportunities and limitations of GenAI in RE: Viewpoints from practice. In: International Working Conference on Requirements Engineering: Foundation for Software Quality, pp. 320–335 (2026). Springer

[30] European Parliament and Council of the European Union: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Official Journal of the European Union, L 119, pp. 1–88. General Data Protection Regulation (GDPR) (2016). https://eur-lex.europa.eu/eli/reg/2016/679/oj

[31] European Parliament: Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union (2024). https://eur-lex.europa.eu/eli/reg/2024/1689/oj

[32] Wiegers, K., Beatty, J.: Software Requirements, 3rd edn. Microsoft Press, Redmond, WA (2013)

[33] Washizaki, H.: Guide to the Software Engineering Body of Knowledge (SWEBOK) v. 4.0. IEEE Computer Society (2024)

[34] Jain, C., Anish, P.R., Singh, A., Ghaisas, S.: A transformer-based approach for abstractive summarization of requirements from obligations in software engineering contracts. In: 2023 IEEE 31st International Requirements Engineering Conference (RE), pp. 169–179 (2023). https://doi.org/10.1109/RE57278.2023.00025

[35] Ren, S., Nakagawa, H., Tsuchiya, T.: Combining prompts with examples to enhance LLM-based requirement elicitation. In: 2024 IEEE 48th Annual Computers, Software, and Applications Conference (COMPSAC), pp. 1376–1381 (2024). https://doi.org/10.1109/COMPSAC61105.2024.00181

[36] Ataei, M., Cheong, H., Grandi, D., Wang, Y., Morris, N., Tessier, A.: Elicitron: An LLM agent-based simulation framework for design requirements elicitation.

arXiv preprint arXiv:2404.16045 (2024)

[37] Arora, C., Grundy, J., Abdelrazek, M.: Advancing requirements engineering through generative ai: Assessing the role of llms. In: Generative AI for Effective Software Development, pp. 129–148. Springer, Cham (2024)

[38] Korn, A., Gorsch, S., Vogelsang, A.: Llmrei: Automating requirements elicitation interviews with LLMs. In: 2025 IEEE 33rd International Requirements Engineering Conference (RE), pp. 19–30 (2025). IEEE

[39] Oswal, J.U., Kanakia, H.T., Suktel, D.: Transforming software requirements into user stories with GPT-3.5-: an AI-powered approach. In: 2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), pp. 913–920 (2024). IEEE

[40] Binkhonain, M., Alfayez, R.: Are prompts all you need? Evaluating prompt-based Large Language Models (LLM)s for software requirements classification. Requirements Engineering **30**(4), 423–443 (2025) https://doi.org/10.1007/s00766-025-00451-8

[41] Fazelnia, M., Koscinski, V., Herzog, S., Mirakhorli, M.: Lessons from the use of natural language inference (NLI) in requirements engineering tasks. In: 2024 IEEE 32nd International Requirements Engineering Conference (RE), pp. 103–115 (2024). IEEE

[42] Peer, J., Mordecai, Y., Reich, Y.: NLP4ReF: Requirements classification and forecasting: From model-based design to large language models. In: 2024 IEEE Aerospace Conference, pp. 1–16 (2024). IEEE

[43] Bashir, S., Ferrari, A., Khan, A., Strandberg, P.E., Haider, Z., Saadatmand, M., Bohlin, M.: Requirements ambiguity detection and explanation with LLMs: An industrial study. In: 2025 IEEE International Conference on Software Maintenance and Evolution (ICSME), pp. 620–631 (2025). IEEE

[44] Lubos, S., Felfernig, A., Tran, T.N.T., Garber, D., El Mansi, M., Erdeniz, S.P., Le, V.-M.: Leveraging LLMs for the quality assurance of software requirements. In: 2024 IEEE 32nd International Requirements Engineering Conference (RE), pp. 389–397 (2024). IEEE

[45] Karimi, M.A., Tiwari, S.: Ucgen: Leveraging LLMs for automated generation of use case textual descriptions from requirements specification. In: Proceedings of the 19th Innovations in Software Engineering Conference, pp. 1–11 (2026)

[46] Zhang, Z., Rayhan, M., Herda, T., Goisauf, M., Abrahamsson, P.: LLM-based agents for automating the enhancement of user story quality: An early report. In: International Conference on Agile Software Development, pp. 117–126. Springer, Cham (2024)

[47] Ferrari, A., Abualhaija, S., Arora, C.: Model generation with LLMs: From requirements to UML sequence diagrams. In: 2024 IEEE 32nd International Requirements Engineering Conference Workshops (REW), pp. 291–300 (2024). IEEE

[48] Kinderen, S., Winter, K.: Towards taming large language models with prompt templates for legal grl modeling. In: International Conference on Business Process Modeling, Development and Support, pp. 213–228 (2024). Springer

[49] Cosler, M., Hahn, C., Mendoza, D., Schmitt, F., Trippel, C.: nl2spec: Interactively translating unstructured natural language to temporal logics with large language models. In: International Conference on Computer Aided Verification, pp. 383–396 (2023). Springer

[50] Cheng, H., Husen, J.H., Lu, Y., Racharak, T., Yoshioka, N., Ubayashi, N., Washizaki, H.: Generative AI for requirements engineering: A systematic literature review. Software: Practice and Experience **56**(2), 141–170 (2026)

[51] Hasso, H., Fischer-Starcke, B., Geppert, H.: Quest-RE QUestion generation and exploration STrategy for Requirements Engineering. In: 2024 IEEE 32nd International Requirements Engineering Conference Workshops (REW), pp. 1–9 (2024). IEEE

[52] Mircea, M., Gevrek, E., Schmid, E., Schneider, K.: Supporting stakeholder requirements expression with LLM revisions: An empirical evaluation. In: International Working Conference on Requirements Engineering: Foundation for Software Quality, pp. 303–319 (2026). Springer

[53] Hassine, J.: An LLM-based approach to recover traceability links between security requirements and goal models. In: Proceedings of the 28th International Conference on Evaluation and Assessment in Software Engineering, pp. 643–651 (2024)

[54] Niu, F., Pan, R., Briand, L.C., Hu, H.: Tvr: Automotive system requirement traceability validation and recovery through retrieval-augmented generation. arXiv preprint arXiv:2504.15427 (2025)

[55] Zadenoori, M.A., Dąbrowski, J., Alhoshan, W., Zhao, L., Ferrari, A.: Large language models (LLMs) for requirements engineering (RE): A systematic literature review. arXiv preprint arXiv:2509.11446 (2025)

[56] Hemmat, A., Sharbaf, M., Kolahdouz-Rahimi, S., Lano, K., Tehrani, S.Y.: Research directions for using LLM in software requirement engineering: A systematic review. Frontiers in Computer Science **7**, 1519437 (2025)

[57] De Michele, W., Armas Cervantes, A., Frermann, L.: Automated business process analysis: An LLM-based approach to value assessment. In: International

Conference on Advanced Information Systems Engineering, pp. 62–69 (2025). Springer

[58] Brath, R., Bradley, A., Jonker, D.: Strategic management analysis: from data to strategy diagram by LLM. In: IEEE 2024 NLVIZ (NLVIZ: Exploring Research Opportunities for Natural Language, Text, and Data Visualization) Workshop (2024)

[59] Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly **13**(3), 319–340 (1989)

[60] Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: Toward a unified view. MIS quarterly **27**(3), 425–478 (2003)

[61] Simaremare, M., Edison, H.: The state of generative AI adoption from software practitioners' perspective: An empirical study. In: 2024 50th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), pp. 106–113 (2024). IEEE

[62] Ge, H., Wu, Y.: An empirical study of adoption of ChatGPT for bug fixing among professional developers. Innovation & Technology Advances **1**(1), 21–29 (2023)

[63] Xia, Y., Chen, Y.: Driving factors of generative AI adoption in new product development teams from a UTAUT perspective. International Journal of Human–Computer Interaction **41**(10), 6067–6088 (2025) https://doi.org/10.1080/10447318.2024.2375686

[64] Lambiase, S., Catolino, G., Palomba, F., Ferrucci, F., Russo, D.: Investigating the role of cultural values in adopting large language models for software engineering. ACM Transactions on Software Engineering Methodology **35**(1) (2025) https://doi.org/10.1145/3725529

[65] Neumann, M., Bischof, L., Hinz, N.E., Altun, A., Stockmann, L., Schrader, D., Ahaus, A.C., Demirci, E.C., Gabel, B., Rauschenberger, M., *et al.*: Between policy and practice: GenAI adoption in agile software development teams. In: International Conference on Agile Software Development, pp. 290–308 (2026). Springer

[66] Stray, V., Barbala, A., Wivestad, V.T.: Human-AI collaboration in software development: A mixed-methods study of developers' use of GitHub Copilot and ChatGPT. In: Proceedings of the 33rd ACM International Conference on the Foundations of Software Engineering, pp. 1325–1332 (2025)

[67] Shao, D., Ishengoma, F.: Empirical analysis of generative AI tool adoption in software development. Information and Software Technology, 108036 (2026)

[68] Hall, T., Beecham, S., Rainer, A.: Requirements problems in twelve software companies: an empirical analysis. IEE Proceedings-Software **149**(5), 153–160 (2002)

[69] Solemon, B., Sahibuddin, S., Ghani, A.A.A.: Requirements engineering problems and practices in software companies: An industrial survey. In: Advances in Software Engineering, pp. 70–77. Springer, Berlin, Heidelberg (2009). https://doi.org/10.1007/978-3-642-10619-4_9

[70] Liu, L., Li, T., Peng, F.: Why requirements engineering fails: A survey report from China. In: 2010 18th IEEE International Requirements Engineering Conference, pp. 317–322 (2010). IEEE

[71] Jarzębowicz, A., Ślesiński, W.: What is troubling IT analysts? A survey report from Poland on requirements-related problems. In: Engineering Software Systems: Research and Praxis, pp. 3–19. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-99617-2_1

[72] Gobov, D., Titlova, O.: Towards identifying challenges in business analysis on IT projects–a practical study. Radioelectronic and Computer Systems **2023**(2), 193–206 (2023)

[73] Kassab, M., Laplante, P.: The current and evolving landscape of requirements engineering in practice. IEEE Software **39**(5), 76–83 (2022) https://doi.org/10.1109/MS.2022.3147692

[74] Ozkaya, M., Akdur, D., Toptani, E.C., Kocak, B., Kardas, G.: Practitioners' perspectives towards requirements engineering: A survey. Systems **11**(2), 65 (2023) https://doi.org/10.3390/systems11020065

[75] Ochodek, M., Kopczyńska, S.: Perceived importance of agile requirements engineering practices – a survey. Journal of Systems and Software **143**, 29–43 (2018) https://doi.org/10.1016/j.jss.2018.05.012

[76] Vogelsang, A., Eckhardt, J., Mendez, D., Berger, M.: Views on quality requirements in academia and practice: commonalities, differences, and context-dependent grey areas. Information and Software Technology **121**, 106253 (2020)

[77] Gobov, D., Sokolovskiy, N.: Association rule mining for requirement elicitation techniques in IT projects. In: 2023 18th Conference on Computer Science and Intelligence Systems (FedCSIS), pp. 983–987 (2023). IEEE

[78] Jarzębowicz, A., Połocka, K.: Selecting requirements documentation techniques for software projects: a survey study. In: 2017 Federated Conference on Computer Science and Information Systems (FedCSIS), pp. 1189–1198 (2017). IEEE

[79] Moreira, I., Adolfo, L., Melegati, J., Choma, J., Guerra, E., Zaina, L.: Exploring documentation strategies for NFR in agile software development. In: International Conference on Agile Software Development, pp. 229–244 (2025). Springer Nature Switzerland Cham

[80] Méndez Fernández, D., Wagner, S.: Naming the pain in requirements engineering: A design for a global family of surveys and first results from germany. Information and Software Technology **57**, 616–643 (2015) https://doi.org/10.1016/j.infsof.2014.05.008

[81] Wagner, S., Fernández, D.M., Felderer, M., Vetrò, A., Kalinowski, M., Wieringa, R., Pfahl, D., Conte, T., Christiansson, M.-T., Greer, D., *et al.*: Status quo in requirements engineering: A theory and a global family of surveys. ACM Transactions on Software Engineering and Methodology (TOSEM) **28**(2), 1–48 (2019)

[82] Wagner, S., et al.: Status quo in requirements engineering: An updated theory from a decade of longitudinal evidence. unpublished (2026)

[83] Gobov, D., Zuieva, O., Shashko, V.: Context-oriented framework for determining requirements change documentation approaches. International Journal of Information Technology and Computer Science **18**(3), 96–108 (2026) https://doi.org/10.5815/ijitcs.2026.03.07

[84] Gobov, D., Sokolovskiy, N.: An association rule mining for selection requirement elicitation and analysis techniques in it projects. In: Conference on Practical Aspects of and Solutions for Software Engineering, pp. 82–96 (2023). Springer

[85] Wilson, E.B.: Probable inference, the law of succession, and statistical inference. Journal of the American Statistical Association **22**(158), 209–212 (1927) https://doi.org/10.1080/01621459.1927.10502953

[86] Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological) **57**(1), 289–300 (1995)

[87] Agresti, A.: Categorical Data Analysis, 3rd edn. Wiley Series in Probability and Statistics. Wiley, Hoboken, NJ (2013)

[88] Wohlin, C., Runeson, P., Höst, M., Ohlsson, M., Regnell, B., Wesslén, A.: Experimentation in Software Engineering. Springer, Berlin, Heidelberg (2024). https://doi.org/10.1007/978-3-662-69306-3

[89] Tomczak, M., Tomczak, E.: The need to report effect size estimates revisited. An overview of some recommended measures of effect size. Trends in Sport Sciences **21**(1), 19–25 (2014)

[90] Bergsma, W.: A bias-correction for Cramér's V and Tschuprow's T. Journal of the Korean Statistical Society **42**(3), 323–328 (2013) https://doi.org/10.1016/j.jkss.2012.10.002

[91] Cohen, J.: Statistical Power Analysis for the Behavioral Sciences, 2nd edn. Lawrence Erlbaum Associates, Hillsdale, NJ (1988)

[92] Romano, J., Kromrey, J.D., Coraggio, J., Skowronek, J.: Appropriate statistics for ordinal level data: Should we really be using t-test and Cohen's d for evaluating group differences on the NSSE and other surveys? In: Annual Meeting of the Florida Association of Institutional Research (2006)

[93] Peduzzi, P., Concato, J., Kemper, E., Holford, T.R., Feinstein, A.R.: A simulation study of the number of events per variable in logistic regression analysis. Journal of clinical epidemiology **49**(12), 1373–1379 (1996)

[94] Firth, D.: Bias reduction of maximum likelihood estimates. Biometrika **80**(1), 27–38 (1993)

[95] Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y., Podsakoff, N.P.: Common method biases in behavioral research: A critical review of the literature and recommended remedies. Journal of Applied Psychology **88**(5), 879–903 (2003) https://doi.org/10.1037/0021-9010.88.5.879

[96] Gelman, A., Stern, H.: The difference between "significant" and "not significant" is not itself statistically significant. The American Statistician **60**(4), 328–331 (2006)

[97] Wick, A., Coventry, T.: Futureproof: Delivering Value with Business Analysis and AI. International Institute of Business Analysis, https://www.iiba.org/career-resources/business-analysis-webinars/2024/futureproof-delivering-value-with-business-analysis-and-ai/. Accessed 2 November 2026 (2024)

[98] Harde, G., Herbert, M., Jimenez, S., Kreiner, D., Mey, M., Tesar, M., Zehentner, F.: AI4RE Micro-Credential: Syllabus and Study Guide, Version 1.0.0. International Requirements Engineering Board (IREB), https://www.ireb.org (2026)

[99] Microsoft and LinkedIn: 2024 Work Trend Index Annual Report: AI at Work Is Here. Now Comes the Hard Part. https://blogs.microsoft.com/blog/2024/05/08/microsoft-and-linkedin-release-the-2024-work-trend-index-on-the-state-of-ai-at-work/. Accessed 2 November 2026 (2024)

[100] OpenAI: Introducing GPT-5.2 (2025). https://openai.com/index/introducing-gpt-5-2/

[101] Google: Gemini 3 Flash Comes to the Gemini App (2025). https://blog.google/products-and-platforms/products/gemini/gemini-3-flash-gemini-app/

[102] Microsoft: Microsoft 365 Copilot Release Notes. Update dated January 27, 2026 (2026). https://learn.microsoft.com/en-us/microsoft-365/copilot/release-notes

[103] Atlassian: Rovo Chat: AI That Works Where You Work (2025). https://www.atlassian.com/blog/announcements/rovo-chat-july-2025-updates

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementary.zip