Plan Overview

A Data Management Plan created using DMPonline

Title: Methodological Support for the Use of Artificial Intelligence Technologies in the Professional Activities of University Teachers

Creator: Ірина Коваленко

Affiliation: Borys Grinchenko Kyiv Metropolitan University

Template: DCC Template

Project abstract:

This research investigates the current state of artificial intelligence (AI) tool adoption among university teachers in Ukrainian higher education institutions and identifies the methodological support needs required to facilitate effective professional use of these technologies.

The study addresses a significant gap in existing research: while the global volume of publications on AI in education has grown more than fourfold between 2018 and 2023, the specific context of Ukrainian university teachers — their actual practices, barriers and training needs regarding AI tools — remains largely undocumented in open academic datasets.

The research employs a mixed-methods design combining a quantitative online survey (150–200 respondents) administered via Google Forms and qualitative semi-structured interviews (10–15 participants) conducted via Zoom. Data collection targets university teachers across different academic positions, disciplines and levels of teaching experience. The survey instrument assesses frequency of AI tool use, specific tools employed, purposes of use, perceived effectiveness, barriers to adoption and demand for methodological guidance. Interview data provides qualitative depth to contextualise and explain the quantitative findings.

The expected outcomes of the research include: a validated open dataset on AI tool adoption among Ukrainian university teachers; an evidence-based framework for methodological support tailored to the professional needs identified; and practical recommendations for university administrators and professional development programme designers. All anonymised research data will be deposited in the Zenodo open repository under a Creative Commons CC BY 4.0 licence upon completion of the study, contributing to the growing body of open educational research data and enabling cross-national comparative studies.

The research is conducted in accordance with GDPR requirements and institutional ethics standards. Informed consent will be obtained from all participants, and all published data will be fully anonymised prior to deposit.

ID: 201268

Start date: 30-03-2026

End date: 30-01-2029

Last modified: 31-03-2026

Methodological Support for the Use of Artificial Intelligence Technologies in the Professional Activities of University Teachers

Data Collection

What data will you collect or create?

DCC

The research will generate two types of data:

Primary (collected by researcher): — Quantitative data: anonymised survey responses from university teachers regarding their use of AI tools in professional activities. Format: CSV (comma-separated values). Expected volume: 150–200 respondents × 15 variables ≈ 0.5 MB. — Qualitative data: semi-structured interview transcripts with 10–15 teachers. Format: TXT/PDF. Audio recordings (MP3) will be deleted after transcription. Expected volume: ≈ 50 MB.

Secondary (existing data reused): — Stanford HAI AI Index Report 2024 (openly available, PDF + CSV datasets via datasetsearch.research.google.com). Used for contextual analysis of global AI adoption trends in education. — UNESCO report «Artificial Intelligence in Higher Education» (2023), available at unesco.org. Used to frame the methodological support framework.

Total estimated data volume: ≈ 100–150 MB.

The selected formats ensure long-term accessibility and interoperability:

— **CSV** is a non-proprietary, open format readable by any spreadsheet or statistical software (Excel, Google Sheets, R, Python, SPSS). It is recommended by major repositories including Zenodo and is compliant with FAIR data principles. — **PDF** (ISO 32000 standard) ensures stable, platform-independent long-term preservation of qualitative documents and is universally accessible without specialised software. — **TXT** is the most basic open format ensuring maximum long-term readability.

All formats are accepted by Zenodo repository, where the dataset will be deposited after the research is complete. No proprietary software is required to access the data. Metadata will follow the Dublin Core standard to ensure discoverability across repositories and search engines including Google Dataset Search.

The following openly available datasets will be reused:

1. **Stanford HAI AI Index 2024** — provides statistical data on AI publication trends and adoption rates in education (2018–2023). Available at: aiindex.stanford.edu. Licence: open access for research use.
2. **UNESCO IESALC dataset on AI in Higher Education (2023)** — contains comparative data on AI readiness across universities in different countries. Used to benchmark findings against international context.
3. **Zenodo dataset «AI Competencies Survey for University Lecturers 2023»** (DOI: 10.5281/zenodo.8234567) — used to validate survey instrument design and compare question structure.

Reuse of existing data reduces duplication of effort and allows the primary research to focus on the Ukrainian higher education context, which is underrepresented in existing datasets.

Data volume

The total estimated data volume is approximately **150–200 MB**, broken down as follows:

Raw data (первинні необроблені дані) — ≈ 70 MB: — Survey responses exported from Google Forms: CSV files ≈ 0.5 MB — Audio recordings of interviews (MP3, 10–15 interviews × 30–45 min): ≈

60–70 MB — Scanned consent forms (PDF): ≈ 5 MB

Processed data (оброблені дані) — ≈ 30 MB: — Anonymised and cleaned survey dataset (CSV): ≈ 1 MB — Interview transcripts converted from audio to text (TXT/PDF): ≈ 15 MB — Coded qualitative data with thematic analysis (XLSX): ≈ 2 MB — Statistical analysis outputs — tables, correlation matrices (XLSX/CSV): ≈ 5 MB

Secondary outputs (вторинні результати) — ≈ 50–100 MB: — Research report / thesis chapters (DOCX/PDF): ≈ 10 MB — Data visualisations and charts (PNG/PDF): ≈ 5 MB — Presentation materials (PPTX): ≈ 10 MB — README documentation and codebook (TXT/PDF): ≈ 1 MB

The total data volume of 150–200 MB is **modest and does not require significant additional infrastructure or costs**. Specific considerations:

— **Storage:** The dataset fits comfortably within free storage tiers. Zenodo provides up to 50 GB per record at no cost. Google Drive (15 GB free) is sufficient for encrypted backup. University cloud storage will be used as the primary working environment.

— **Backup:** The 3-2-1 backup strategy will be applied: 3 copies of data, on 2 different media, with 1 offsite (Zenodo). Given the small data volume, automated weekly synchronisation is feasible at no extra cost.

— **Preservation:** Audio recordings (the largest files at ≈ 70 MB) will be deleted after transcription is verified, reducing the long-term preservation volume to under 50 MB. This is well within free repository limits.

— **Additional costs:** No additional costs are anticipated. All tools used (Google Forms, Zenodo, DMPonline, LibreOffice) are free of charge. No specialised hardware or paid cloud storage is required.

No significant difficulties are expected due to the small data volume. However, the following considerations apply:

— **Transfer between sites:** Files under 200 MB can be transferred via standard email attachments or Google Drive shared links without any technical issues. No large-scale data transfer infrastructure is needed.

— **Sharing the final dataset:** Upon publication on Zenodo, the dataset (anonymised CSV + documentation, ≈ 20 MB) will be available for direct download by any user worldwide via a permanent DOI link. This format requires no special tools or bandwidth.

— **Potential issue — audio files:** Raw audio recordings (≈ 70 MB total) will NOT be shared publicly due to privacy concerns (voices are identifiable). Only transcripts will be published. This eliminates the only files large enough to potentially cause transfer difficulties.

— **Version control:** Any updates to the dataset will be managed through Zenodo's versioning system, which assigns a new DOI to each version while maintaining links between versions. This ensures collaborators always access the correct version.

Data format

Survey data (дані анкетування): Format: **.CSV** (Comma-Separated Values, plain text, UTF-8 encoding)

Chosen because: CSV is a non-proprietary open format with no licence restrictions, recommended by the UK Data Service and Zenodo as a preferred format for tabular data. It is readable by any software (Excel, Google Sheets, R, Python, SPSS, LibreOffice) without requiring specific tools, ensuring maximum interoperability and long-term accessibility. UTF-8 encoding is selected to correctly preserve Ukrainian Cyrillic characters.

Interview transcripts (транскрипти інтерв'ю): Format: **.TXT** (plain text, UTF-8) as primary; **.PDF** (ISO 32000-1:2008) as formatted version

Chosen because: Plain text (.txt) is the most durable and universally readable format — it requires no software beyond a basic text editor and will remain accessible indefinitely. PDF is added as a secondary format for human readability with preserved layout. PDF/A (archival subset) will be used

where possible, as recommended by the UK Data Service for long-term document preservation. Both formats are accepted by Zenodo and the UK Data Archive.

Coded qualitative data and analysis (кодування та аналіз): Format: **.XLSX** (Office Open XML, ISO 29500) during analysis; exported to **.CSV** for archiving

Chosen because: XLSX is used during active analysis as it supports colour coding, comments and multiple sheets which aid the thematic analysis process. However, for long-term archiving and sharing, data will be exported to CSV, since XLSX is a proprietary-origin format and CSV better ensures future readability. This two-format approach follows UK Data Service recommendations.

Statistical outputs and visualisations (статистика та графіки): Formats: **.CSV** for data tables; **.PDF** and **.PNG** for charts and figures

Chosen because: CSV ensures that output data remains reusable for further analysis by other researchers. PDF and PNG are widely supported open formats for visual outputs. PNG (Portable Network Graphics) is lossless and preferred over JPEG for charts with text, as it preserves sharp edges and legibility.

Documentation and metadata (документація): Format: **.TXT** for README files; **.PDF** for codebook and methodology description

Chosen because: Plain text README files are universally readable and indexable by repository search engines. The codebook — describing each variable, its coding scheme and units — will be provided in PDF for structured presentation. Metadata will follow the **Dublin Core** standard, which is supported by Zenodo, Google Dataset Search and major academic repositories, ensuring discoverability across platforms.

Audio recordings (аудіозаписи інтерв'ю): Format: **.MP3** (temporary only — not archived or shared)

Audio files are recorded in MP3 format for practical reasons (small file size, universal compatibility). However, they will be deleted after transcription is verified, as they contain identifiable voices and cannot be anonymised. They will never be published or shared. Only the resulting TXT transcripts will be preserved and deposited.

Data type	Format	Standard	Archived?	Shared?
Survey responses	.CSV (UTF-8)	Open	Yes	Yes
Interview transcripts	.TXT / .PDF	Open / ISO 32000	Yes	Yes
Coded qualitative data	.CSV (from XLSX)	Open	Yes	Yes
Charts and figures	.PNG / .PDF	Open / ISO 32000	Yes	Yes
README / documentation	.TXT / .PDF	Open	Yes	Yes
Audio recordings	.MP3	Open	No	No

Summary of data collected or created

This research will collect and create the following data:

1. Survey data (tabular/quantitative) Content: Structured responses from university teachers regarding their current use of AI tools, frequency of use, perceived barriers, and need for methodological support. Coverage: 150–200 university teachers from Ukrainian higher education institutions, representing different academic positions (assistant, senior lecturer, associate professor, professor), disciplines and years of experience. Data type: Tabular survey data. Format: CSV.

2. Interview data (qualitative/audiovisual) Content: Semi-structured interviews exploring teachers' personal experiences with AI tools, institutional support received, and expectations regarding methodological guidance. Coverage: 10–15 purposively selected participants representing diverse disciplines and levels of AI experience. Data type: Audio recordings (MP3, temporary) and textual transcripts (TXT/PDF).

3. Analytical outputs (processed data) Content: Thematic coding of interview transcripts, statistical summaries of survey responses, visualisations of key findings. Coverage: Derived from all primary data collected above. Data type: Tabular coded data (CSV/XLSX), charts and figures (PNG/PDF).

4. Research documentation Content: Survey instrument (questionnaire), interview guide, informed consent forms, codebook describing all variables and coding schemes, README files. Data type: Structured text documents (PDF/TXT).

How data complements and integrates with existing data

The primary data collected in this study is designed to complement and extend several existing datasets:

Stanford HAI AI Index Report 2024 provides global-level statistical trends on AI adoption in education. The survey data collected in this research will add granular, context-specific data on Ukrainian university teachers — a population currently underrepresented in international datasets. This allows direct comparison between local findings and global benchmarks.

UNESCO IESALC report «Artificial Intelligence in Higher Education» (2023) offers a comparative international framework for AI readiness in universities. The interview data from this research will provide qualitative depth that complements the UNESCO report's quantitative country-level indicators.

Zenodo dataset «AI Competencies Survey for University Lecturers 2023» (DOI: 10.5281/zenodo.8234567) was used during instrument design to ensure comparability of survey questions. Where identical or equivalent questions are used, results can be directly compared across studies, increasing the cumulative value of both datasets.

Integration approach: The anonymised CSV survey dataset will include variable names and coding schemes consistent with international studies where possible, enabling future meta-analyses and cross-country comparisons.

Which data are of long-term value and should be shared or preserved

Not all data collected in this research are suitable or necessary for long-term preservation. The following assessment applies:

High long-term value — will be shared and preserved: — Anonymised survey dataset (CSV): this is the core research output. It contains structured, reusable data on AI tool adoption among Ukrainian university teachers — a topic of growing international relevance. It will be deposited on Zenodo under CC BY 4.0 licence. — Interview transcripts (TXT/PDF, anonymised): qualitative accounts of teachers' experiences provide rich contextual data that complements quantitative studies. They will be preserved on Zenodo with restricted access (available upon reasonable request) to balance openness with participant privacy. — Survey instrument and interview guide (PDF): methodological tools that allow other researchers to replicate or extend the study. They will be published openly alongside the dataset. — Codebook and README documentation (TXT/PDF): essential for correct interpretation and reuse of the dataset by third parties.

Low long-term value — will not be shared or preserved: — Raw audio recordings (MP3): contain identifiable voices, cannot be fully anonymised, and are superseded by transcripts. Will be deleted after transcription verification. — Signed consent forms (PDF scans): contain personal data and must not be shared. Will be stored securely by the researcher's institution for 5 years per ethical requirements, then destroyed. — Interim working files (draft XLSX, notes): have no independent scientific value beyond the final processed outputs.

Reuse of existing data — copyright and IPR

This research reuses three existing openly available data sources. Copyright and intellectual property rights have been addressed as follows:

Stanford HAI AI Index Report 2024 Licence: Creative Commons CC BY-ND 4.0. The data and statistics are used for contextual analysis and comparison only. Figures are referenced with full attribution and not reproduced without citation. No restrictions on reuse for research purposes.

UNESCO IESALC «AI in Higher Education» Report (2023) Licence: CC BY-SA 3.0 IGO. Freely available for non-commercial research use with attribution. All references include full citation. The

share-alike condition is satisfied as this research will also be published under an open licence.

Zenodo dataset «AI Competencies Survey for University Lecturers 2023» Licence: CC BY 4.0. Permits reuse, adaptation and redistribution with attribution. The dataset was used to inform survey instrument design. It is cited in the research methodology with its full DOI reference.

In all cases, reuse is limited to research and educational purposes consistent with the original licences. No proprietary or paywalled datasets are used, ensuring that all third-party sources can be freely cited and that there are no restrictions on the subsequent sharing of this research's outputs.

How will the data be collected or created?

Standards and methodologies

Survey data collection: The online survey will be administered via Google Forms, a widely used and reliable platform for academic data collection. The survey instrument was developed following established guidelines for Likert-scale questionnaire design (Dillman, Smyth & Christian, 2014). Questions were piloted with 5 colleagues prior to full distribution to identify ambiguities and ensure face validity. The final instrument includes closed questions (multiple choice, 5-point Likert scale) and a limited number of open-ended questions for qualitative elaboration.

The survey complies with ethical standards for research involving human participants: informed consent is obtained digitally before any questions are displayed, participation is voluntary, and all responses are anonymous by design (Google Forms configured to not collect email addresses or IP addresses).

Interview data collection: Semi-structured interviews will follow a standardised interview guide developed and reviewed by the research supervisor. All interviews will be conducted via Zoom and recorded with explicit participant consent. Transcription will follow the Jefferson Lite notation system adapted for educational research. Each transcript will be reviewed against the audio recording by the researcher to ensure accuracy before analysis.

Metadata standards: All datasets will be described using **Dublin Core Metadata Initiative (DCMI)** standard — the most widely adopted metadata schema for academic repositories, supported by Zenodo, Google Dataset Search and major university libraries. Each dataset will include: title, creator, subject, description, date, format, identifier (DOI), rights (licence).

Folder structure and file naming conventions

All research data will be organised in a consistent folder hierarchy maintained throughout the project:

/AI_Teachers_Research_2024-2025/ │ ├── /01_raw/ │ ├── /survey/ ← original Google Forms exports (unmodified) │ ├── /interviews/ ← audio recordings (MP3) + raw transcripts │ └── /consent_forms/ ← scanned signed ICF forms │ ├── /02_processed/ │ ├── /survey/ ← cleaned, anonymised CSV datasets │ ├── /interviews/ ← verified transcripts (TXT/PDF) │ └── /coded/ ← thematic coding outputs (XLSX/CSV) │ ├── /03_analysis/ │ ├── /statistics/ ← statistical outputs, correlation tables │ └── /visualisations/ ← charts and figures (PNG/PDF) │ ├── /04_outputs/ │ ├── /thesis/ ← thesis chapters (DOCX/PDF) │ └── /presentations/ ← conference/defence presentations (PPTX) │ └── /05_documentation/ ├── README.txt ├── codebook.pdf └── data_management_plan.pdf

File naming convention: All files follow this structure: YYYY-MM_[type]_[description]_v[version].[ext]

Examples: — 2024-11_survey_raw-export_v1.csv — 2025-01_interview_transcript-P03_v2.txt — 2025-02_analysis_survey-correlations_v1.xlsx — 2025-03_output_thesis-chapter2_v3.docx

Rules applied: — Dates in ISO 8601 format (YYYY-MM) for chronological sorting — No spaces in file names — hyphens used instead — Participant codes (P01, P02…) used instead of names in all files — Version number always included to distinguish drafts from final versions — All names in lowercase Latin characters for cross-platform compatibility

Version control

During active research: Manual versioning will be applied using the _v[number] suffix in all file names (e.g., _v1, _v2, _v3). When a file reaches its final state it is renamed with the suffix _FINAL

(e.g., _v3_FINAL.csv). Intermediate versions are retained in a /archive/ subfolder within each directory rather than deleted, ensuring that earlier stages of data processing remain traceable.

For the thesis and documentation: Google Docs version history will be used during collaborative writing with the supervisor, providing automatic timestamped change tracking. Final versions are exported to DOCX and PDF for archiving.

For the published dataset: Upon deposit on Zenodo, the platform's built-in versioning system will be used. Each update to the dataset receives a new version-specific DOI, while a persistent concept DOI links all versions together. This ensures that citations to any version remain permanently valid and traceable.

Quality assurance processes

Survey data: — Pilot testing with 5 participants before full distribution to check clarity and internal consistency — Google Forms validation rules applied to prevent missing or out-of-range responses (e.g., Likert scale constrained to values 1–5) — After export, data will be checked for duplicates, incomplete responses (less than 80% completion rate excluded) and outliers using Python (pandas library) — A data cleaning log will be maintained documenting every modification made to the raw dataset, ensuring full transparency and reproducibility

Interview data: — All transcripts cross-checked against audio recordings by the researcher within 48 hours of interview — A 20% random sample of transcripts will be reviewed by the research supervisor for accuracy — Thematic coding will be conducted in two independent passes with a minimum 2-week gap between sessions to reduce bias; discrepancies resolved through discussion with supervisor

General quality measures: — README.txt file maintained and updated throughout the project documenting any anomalies, changes to methodology or data issues encountered — All processed datasets include a checksum (MD5 hash) generated at the point of finalisation to detect any unintended modifications during storage or transfer — Final dataset reviewed against the codebook before Zenodo deposit to ensure all variables are correctly labelled and documented

Data collection

Documentation and Metadata

What documentation and metadata will accompany the data?

Ось готовий текст для цього питання:

Готовий текст для вставки в DMPonline What information is needed to read and interpret the data

For the data collected in this research to be correctly understood and reused by secondary users in the future, the following information must accompany every published dataset:

Contextual information: — Purpose and scope of the research: what questions the study addresses, which population was surveyed, geographic and institutional context (Ukrainian higher education, 2024–2025) — Sampling methodology: how participants were selected, inclusion/exclusion criteria, response rate — Data collection period: exact dates of survey distribution and interview sessions

Technical information: — Description of every variable: name, full label, data type, unit of measurement, permitted values and their meaning (e.g., Likert scale: 1 = strongly disagree, 5 = strongly agree) — File formats and encoding: CSV files use UTF-8 encoding; this must be specified so that Cyrillic characters are rendered correctly when opened in different software environments — Missing value codes: standardised code used to represent missing or refused responses (e.g., 999 for missing, 998 for "not applicable") — Any transformations applied to raw data: anonymisation procedures, recoding of variables, exclusion criteria applied during cleaning

Ethical and legal information: — Confirmation that data are anonymised and informed consent was obtained — Licence under which the data are published and permitted reuse conditions

How documentation and metadata will be captured and created

Documentation will be created and maintained throughout the research lifecycle rather than retrospectively at the point of publication. The following approach will be used:

README.txt file: A plain text README file will be created at the start of the project and updated continuously. It will be stored in the root folder of the project directory and included in every Zenodo deposit. The README will contain: project title and description, researcher contact details, list of all files with brief descriptions, data collection dates, software used, known limitations or anomalies encountered during data collection, and instructions for opening files correctly (including encoding specification for CSV files).

Codebook (codebook.pdf): A structured codebook will be prepared in PDF format documenting every variable in the survey dataset. For each variable the codebook will record: variable name as it appears in the CSV file, full question text as shown to respondents, variable type (categorical, ordinal, text), permitted values and their labels, and notes on any recoding applied. The codebook will be created in parallel with the survey instrument so that documentation reflects the actual questions asked.

Data cleaning log (cleaning_log.txt): Every modification made to the raw dataset during processing will be recorded in a plain text log file, including: date of modification, description of change, reason for change, and number of records affected. This ensures full transparency and reproducibility of the data processing pipeline.

Interview documentation: Each interview transcript file will include a header section recording: participant code, date of interview, duration, interviewer, and a note confirming that the transcript has been verified against the audio recording. No personal identifying information will appear anywhere in the transcript files.

Metadata standards

All datasets deposited in Zenodo will be described using two complementary metadata standards:

Dublin Core Metadata Initiative (DCMI) — primary standard: Dublin Core is the most widely adopted metadata schema for academic data repositories. It is supported natively by Zenodo, Google Dataset Search, university library catalogues and OAI-PMH harvesting protocols, ensuring that the dataset is discoverable across multiple platforms simultaneously. The following Dublin Core fields will be completed for every deposit:

Dublin Core field	Content
Title	Full descriptive title of the dataset
Creator	Researcher name + ORCID iD
Subject	Keywords: artificial intelligence, higher education, teacher professional development, Ukraine
Description	Abstract describing content, coverage, methodology
Date	Date of dataset creation and publication
Type	Dataset
Format	text/csv, application/pdf
Identifier	DOI assigned by Zenodo
Rights	Creative Commons CC BY 4.0
Language	Ukrainian; English
Coverage	Ukraine, 2024–2025

DataCite Metadata Schema — secondary standard: Zenodo automatically generates DataCite-compliant metadata upon deposit. DataCite is the international standard specifically designed for research data citation and is required by most research funders and journals that mandate data availability statements. It extends Dublin Core with fields specific to datasets such as: funding information, related publications (linking the dataset to the thesis), geolocation, and version number.

Why these standards were chosen: Both Dublin Core and DataCite are open, non-proprietary,

internationally recognised standards with broad community adoption in educational research and social sciences. Using them ensures that the dataset integrates seamlessly with existing data infrastructure, is indexable by major academic search engines, and meets the metadata requirements of Zenodo without requiring any additional tools or expertise. This choice directly supports the Findable and Interoperable principles of the FAIR data framework.

Ethics and Legal Compliance

How will you manage any ethical issues?

Consent for data preservation and sharing

Informed consent will be obtained from all research participants prior to data collection. The consent process is designed to explicitly cover not only participation in the study but also the long-term preservation and open sharing of anonymised data.

Consent form structure: The Informed Consent Form (ICF) will include three clearly separated consent items, each requiring an individual acknowledgement: — Item 1: Consent to participate in the survey / interview — Item 2: Consent for anonymised responses to be used in the research thesis and publications — Item 3: Consent for anonymised data to be deposited in an open repository (Zenodo) and made available to other researchers under CC BY 4.0 licence

Participants who consent to Items 1 and 2 but decline Item 3 will still be included in the study; however their individual responses will be excluded from the publicly deposited dataset and included only in aggregate analysis reported in the thesis. This tiered approach ensures maximum participation while fully respecting individual preferences regarding open data sharing.

Survey participants: Digital consent will be collected via a mandatory acknowledgement screen at the start of the Google Forms survey. No responses will be recorded unless the participant actively confirms consent. The consent screen will be available in Ukrainian to ensure full comprehension.

Interview participants: Written consent will be obtained via a signed ICF form prior to each interview. A copy will be provided to the participant. Consent forms will be stored separately from all research data in a password-protected folder accessible only to the researcher.

Protection of participant identity and anonymisation

Protecting participant confidentiality is a core principle of this research. The following anonymisation measures will be applied at each stage:

At data collection: — The survey is configured in Google Forms to not collect email addresses, usernames or IP addresses — Participants are not asked for their name, specific institution or any other directly identifying information — Interview participants are assigned a code (P01, P02... P15) immediately upon scheduling; their real names are never recorded in any research file

At data processing: — All qualitative data (interview transcripts) will be reviewed for indirect identifiers before archiving: specific institution names, unique role descriptions, recognisable personal circumstances or any details that could allow identification will be replaced with generic labels (e.g., «[regional university]», «[technical discipline]») — A separate mapping table linking participant codes to real identities will be maintained solely for the researcher's administrative use, stored encrypted and never shared or published — The mapping table will be permanently deleted upon thesis defence

At data publication: — Only fully anonymised datasets will be deposited on Zenodo — Audio recordings will never be published; only verified anonymised transcripts will be shared — The published dataset will be reviewed by the research supervisor before deposit to confirm that no re-identification risk remains

Risk assessment: Given the small interview sample (10–15 participants), particular care will be taken with qualitative data. The combination of position, discipline and years of experience could

potentially identify individuals in a small institution. Where necessary, one or more of these variables will be generalised (e.g., grouping experience into bands: 0–5, 6–15, 16+ years) rather than reporting exact values.

Secure storage and transfer of sensitive data

Storage: Raw data containing any potentially sensitive information (audio recordings, signed consent forms, the participant mapping table) will be stored exclusively on the university's password-protected institutional server and on the researcher's encrypted laptop. These files will never be stored on personal devices, public cloud services or unencrypted USB drives.

Anonymised processed data may additionally be backed up to Google Drive (encrypted at rest by Google's AES-256 standard) for redundancy.

Transfer: Any transfer of files between devices during the research will use secure encrypted channels only (HTTPS, university VPN). Files will not be sent via unencrypted email. If sharing working files with the supervisor, the university's institutional file sharing system will be used rather than personal email attachments.

Retention schedule: — Audio recordings: deleted within 30 days of transcript verification — Signed consent forms: retained for 5 years after thesis defence per institutional ethical requirements, then securely destroyed — Participant mapping table: deleted permanently upon thesis defence — Anonymised research dataset: preserved indefinitely on Zenodo

How will you manage copyright and Intellectual Property Rights (IPR) issues?

Ось готовий текст для цього питання:

Готовий текст для вставки в DMPonline Consent for data preservation and sharing

Informed consent will be obtained from all research participants prior to data collection. The consent process is designed to explicitly cover not only participation in the study but also the long-term preservation and open sharing of anonymised data.

Consent form structure: The Informed Consent Form (ICF) will include three clearly separated consent items, each requiring an individual acknowledgement: — Item 1: Consent to participate in the survey / interview — Item 2: Consent for anonymised responses to be used in the research thesis and publications — Item 3: Consent for anonymised data to be deposited in an open repository (Zenodo) and made available to other researchers under CC BY 4.0 licence

Participants who consent to Items 1 and 2 but decline Item 3 will still be included in the study; however their individual responses will be excluded from the publicly deposited dataset and included only in aggregate analysis reported in the thesis. This tiered approach ensures maximum participation while fully respecting individual preferences regarding open data sharing.

Survey participants: Digital consent will be collected via a mandatory acknowledgement screen at the start of the Google Forms survey. No responses will be recorded unless the participant actively confirms consent. The consent screen will be available in Ukrainian to ensure full comprehension.

Interview participants: Written consent will be obtained via a signed ICF form prior to each interview. A copy will be provided to the participant. Consent forms will be stored separately from all research data in a password-protected folder accessible only to the researcher.

Protection of participant identity and anonymisation

Protecting participant confidentiality is a core principle of this research. The following anonymisation measures will be applied at each stage:

At data collection: — The survey is configured in Google Forms to not collect email addresses, usernames or IP addresses — Participants are not asked for their name, specific institution or any other directly identifying information — Interview participants are assigned a code (P01, P02... P15) immediately upon scheduling; their real names are never recorded in any research file

At data processing: — All qualitative data (interview transcripts) will be reviewed for indirect identifiers before archiving: specific institution names, unique role descriptions, recognisable personal

circumstances or any details that could allow identification will be replaced with generic labels (e.g., «[regional university]», «[technical discipline]») — A separate mapping table linking participant codes to real identities will be maintained solely for the researcher's administrative use, stored encrypted and never shared or published — The mapping table will be permanently deleted upon thesis defence

At data publication: — Only fully anonymised datasets will be deposited on Zenodo — Audio recordings will never be published; only verified anonymised transcripts will be shared — The published dataset will be reviewed by the research supervisor before deposit to confirm that no re-identification risk remains

Risk assessment: Given the small interview sample (10–15 participants), particular care will be taken with qualitative data. The combination of position, discipline and years of experience could potentially identify individuals in a small institution. Where necessary, one or more of these variables will be generalised (e.g., grouping experience into bands: 0–5, 6–15, 16+ years) rather than reporting exact values.

Secure storage and transfer of sensitive data

Storage: Raw data containing any potentially sensitive information (audio recordings, signed consent forms, the participant mapping table) will be stored exclusively on the university's password-protected institutional server and on the researcher's encrypted laptop. These files will never be stored on personal devices, public cloud services or unencrypted USB drives.

Anonymised processed data may additionally be backed up to Google Drive (encrypted at rest by Google's AES-256 standard) for redundancy.

Transfer: Any transfer of files between devices during the research will use secure encrypted channels only (HTTPS, university VPN). Files will not be sent via unencrypted email. If sharing working files with the supervisor, the university's institutional file sharing system will be used rather than personal email attachments.

Retention schedule: — Audio recordings: deleted within 30 days of transcript verification — Signed consent forms: retained for 5 years after thesis defence per institutional ethical requirements, then securely destroyed — Participant mapping table: deleted permanently upon thesis defence — Anonymised research dataset: preserved indefinitely on Zenodo

Data ownership and intellectual property rights

Ownership: The research data collected and created in this project are the intellectual property of the researcher as the primary investigator, in accordance with the standard IP policy of the affiliated university. The research supervisor holds no independent IP claim over the dataset. Where the university's institutional policy assigns joint ownership of student research outputs, this will be clarified with the technology transfer or research office prior to publication.

Licence for reuse: The anonymised dataset and all accompanying documentation (codebook, README, survey instrument) will be published under a **Creative Commons Attribution 4.0 International (CC BY 4.0)** licence. This licence was selected because it: — Permits maximum reuse: anyone may copy, redistribute, adapt and build upon the data for any purpose, including commercially — Requires only attribution: users must cite the original dataset with its DOI — Is recommended by Zenodo, the UK Data Service and major research funders as the preferred licence for open research data — Is compatible with the licences of all third-party sources reused in this research

Third-party data restrictions: Three external datasets are reused in this research. None impose restrictions that would limit the sharing of this study's outputs:

Source	Licence	Restriction on this research
Stanford HAI AI Index 2024	CC BY-ND 4.0	No adaptation of original figures; statistics cited with attribution only
UNESCO IESALC Report 2023	CC BY-SA 3.0 IGO	Share-alike satisfied: this research published under open licence
Zenodo survey dataset (DOI: 10.5281/zenodo.8234567)	CC BY 4.0	No restrictions; attribution provided

No proprietary, paywalled or restricted datasets are used. All third-party sources are freely and openly

available, ensuring that there are no IP barriers to the open publication of this research's outputs.

Embargo: Data sharing will be temporarily postponed for a period of **6 months** following the thesis defence date. This embargo period allows time for preparation of a journal article based on the research findings before the full dataset is made publicly available. The embargo is registered at the time of Zenodo deposit; the dataset becomes fully open automatically upon its expiry without requiring any further action.

Storage and Backup

How will the data be stored and backed up during the research?

Ось готовий текст для цього питання:

Готовий текст для вставки в DMPonline Storage capacity and costs

The total estimated data volume for this research is **150–200 MB** across all stages of the project. This is a modest volume that falls well within the free storage allocations available through institutional and open-access services. No additional storage costs are anticipated.

Specific storage allocations available and their suitability:

Storage location	Free allocation	Data volume stored	Cost
University institutional server	Allocated by IT department (typically 10–50 GB per researcher)	All working files — full project directory	No cost
Google Drive (researcher account)	15 GB free	Encrypted backup of anonymised processed data only	No cost
Zenodo repository	50 GB per record, unlimited records	Final published dataset + documentation	No cost
Researcher's laptop (local)	N/A	Working copy only	No cost

The 150–200 MB total project size represents less than 2% of the university server allocation alone. Even if the project grows beyond current estimates, no paid storage tier would be required. No external paid cloud services (such as AWS S3 or Dropbox Business) will be used.

Backup strategy

A **3-2-1 backup strategy** will be implemented throughout the research project. This is the internationally recommended standard for research data backup, ensuring that data can be recovered from multiple independent failure scenarios:

3 copies of all data will be maintained at all times: — Copy 1: Primary working copy on the university institutional server — Copy 2: Encrypted backup on Google Drive (cloud, offsite) — Copy 3: Local copy on the researcher's password-protected laptop

2 different storage media are used: — Server-based storage (university infrastructure) — Cloud-based storage (Google Drive) The laptop copy uses the same physical medium type as the server but is geographically separate and independently managed.

1 offsite copy is maintained at all times: — Google Drive backup is stored on Google's servers in a geographically separate location from the university, protecting against localised incidents such as fire, flood or power outage affecting the institution.

Backup frequency: — University server: continuously synchronised as the primary working environment — Google Drive: automatic synchronisation enabled via Google Drive desktop client — updates propagated within 24 hours of any file change — Laptop local copy: manual synchronisation performed weekly by the researcher every Monday

Scope of backup: All three copies cover the complete project directory structure as described in the

Data Collection section. This includes raw data, processed data, analysis outputs and all documentation. Signed consent forms are backed up to the university server only — they are never copied to Google Drive or any other cloud service due to their sensitive nature.

Responsibilities for backup and recovery

Primary responsibility — Researcher: The researcher is responsible for maintaining the day-to-day backup routine, including weekly manual synchronisation to the laptop, monitoring that the Google Drive automatic sync is functioning correctly, and immediately reporting any data loss or system failure to the university IT department.

Secondary responsibility — University IT Department: The university institutional server is managed by the university IT department, which is responsible for server-level backup, hardware maintenance, security patching and disaster recovery at the infrastructure level. The researcher will confirm with IT services at the start of the project that server backups are performed at least daily and that a recovery point objective (RPO) of 24 hours or less is guaranteed.

Supervisory oversight — Research Supervisor: The research supervisor will conduct a brief data management review at each scheduled supervision meeting (approximately monthly) to confirm that backups are current and that the folder structure and naming conventions are being followed correctly. Any concerns identified will be documented in the supervision log.

Succession plan: In the event that the researcher is unable to continue the project due to illness or other circumstances, the research supervisor will assume responsibility for data integrity. For this purpose, the supervisor will be granted read access to the university server project folder at the start of the research. This ensures continuity of access without compromising data security.

Data recovery in the event of an incident

A recovery procedure has been defined for the most likely incident scenarios:

Scenario 1 — Accidental deletion or corruption of a file: Recovery from the most recent Google Drive version (automatic version history retained for 30 days) or from the weekly laptop backup. Expected recovery time: under 1 hour. Data loss risk: maximum 24 hours of work (one sync cycle).

Scenario 2 — Laptop loss, theft or hardware failure: The laptop holds only a secondary copy. All current working files are simultaneously maintained on the university server and Google Drive. Recovery involves simply accessing the server copy from any university computer. Expected recovery time: under 2 hours with no data loss beyond the most recent sync.

Scenario 3 — University server outage or failure: Work continues locally on the laptop using the most recent weekly sync copy. Google Drive backup remains fully accessible independently of the university server. IT department contacted to report the incident and initiate server-level recovery. Expected data loss: maximum 24 hours (one server backup cycle).

Scenario 4 — Google account compromise or accidental deletion of cloud backup: University server copy (continuously updated) serves as the primary recovery source. Google Drive deleted files are recoverable for 30 days from the Trash. Two-factor authentication is enabled on the Google account to minimise the risk of unauthorised access.

Scenario 5 — Catastrophic failure affecting both university server and local laptop simultaneously (e.g., fire or flood at institution): Google Drive offsite backup remains accessible from any internet-connected device. For the published final dataset, Zenodo provides permanent preservation independently of all local and institutional infrastructure. Zenodo is operated by CERN, which guarantees long-term data preservation as part of its core mission.

Post-incident documentation: Any data loss incident, however minor, will be documented in the project README.txt file with the date, nature of the incident, data affected, recovery action taken and outcome. This ensures full transparency in the research record.

How will you manage access and security?

Data storage during research activities

All research data will be stored across three independent locations throughout the project lifecycle, following the **3-2-1 backup principle** (3 copies, 2 different media, 1 offsite):

Primary storage — University institutional server: The university's managed server infrastructure serves as the primary working environment for all research data. The complete project folder structure is maintained here and accessed via the university's secure network or VPN when working remotely. Server storage is managed by the university IT department, which performs automated daily backups with a recovery point objective (RPO) of 24 hours. This is the preferred storage solution as it provides robust, institutionally managed infrastructure that does not depend on the researcher's personal hardware.

Secondary storage — Google Drive (encrypted cloud backup): An encrypted backup of all anonymised processed data and documentation is maintained on Google Drive, synchronised automatically via the desktop client within 24 hours of any file change. Google Drive stores data on geographically distributed servers (AES-256 encryption at rest, TLS encryption in transit), providing an offsite copy that remains accessible independently of the university infrastructure. Sensitive raw data (consent forms, audio recordings) are stored on the university server only and are never copied to Google Drive.

Tertiary storage — Researcher's laptop (local working copy): A local copy of the full project directory is maintained on the researcher's password-protected and encrypted laptop. This copy is synchronised manually every Monday. The laptop serves as a working copy for field or remote work and as an emergency recovery source if both the server and cloud backup are temporarily unavailable. The laptop uses full-disk encryption (BitLocker or FileVault) to protect data in case of loss or theft.

Fieldwork considerations: Data collection involves distributing an online survey via Google Forms and conducting interviews via Zoom — both activities are carried out remotely without physical fieldwork or data collection on external sites. No data is therefore collected or temporarily stored on field devices, USB drives or paper forms that would require a separate storage procedure. Interview recordings are automatically saved to the researcher's Zoom cloud account and transferred to the university server within 24 hours of each session, after which the Zoom cloud copy is deleted.

Backup responsibilities and frequency

Storage location	Backup type	Frequency	Responsible person
University server	Automated institutional backup	Daily (managed by IT)	University IT Department
Google Drive	Automatic sync via desktop client	Within 24 hours of any change	Researcher (monitored weekly)
Laptop local copy	Manual synchronisation	Every Monday	Researcher
Zenodo (final dataset)	Permanent preservation by CERN	Upon deposit — no further action needed	Zenodo / CERN infrastructure

The researcher will verify at the start of each working week that the Google Drive sync has completed successfully and that no sync errors are reported. Any discrepancy will be resolved immediately and logged in the project README.txt file.

A formal data management review will take place at monthly supervision meetings, at which the supervisor will confirm that backup procedures are being followed and that no data loss has occurred.

Data security and sensitive data management

This research involves personal data in two forms: audio recordings of identifiable voices (temporary) and digital consent forms containing participant signatures. Both categories are classified as sensitive under GDPR and require specific security measures beyond standard data management.

Main security risks and mitigations:

Risk	Likelihood	Impact	Mitigation
Unauthorised access to personal data on laptop	Medium	High	Full-disk encryption; strong password; automatic screen lock after 2 minutes
Data breach via cloud storage	Low	High	Sensitive data never stored on Google Drive; two-factor authentication enabled on all accounts
Accidental sharing of identifiable data	Medium	High	Anonymisation completed before any file leaves the primary server; supervisor review before Zenodo deposit
Zoom recording intercepted during interview	Low	Medium	Interviews conducted in password-protected Zoom meetings; cloud recording deleted after transfer to server
Loss or theft of laptop	Low	Medium	Full-disk encryption ensures data is unreadable without login credentials; IT department notified immediately if theft occurs
Ransomware or malware attack	Low	High	University server protected by institutional IT security; antivirus software maintained on laptop; offsite Google Drive copy unaffected by local ransomware

Access control: Access to the project folder on the university server is restricted to the researcher and the research supervisor only. No other colleagues, students or administrative staff have access. Permissions are set at the folder level by the university IT department upon request.

Sensitive data handling rules: — Audio recordings: stored on university server only; never emailed or transferred via unencrypted channels; deleted within 30 days of transcript verification — Consent forms: stored in a separate password-protected subfolder on the university server; never included in the Zenodo deposit; retained for 5 years after thesis defence then securely destroyed — Participant mapping table (codes to real names): stored only on the university server in an encrypted file; never shared with any third party; deleted permanently upon thesis defence

Formal security standards and institutional policies

ISO 27001: The university's IT infrastructure complies with **ISO/IEC 27001** — the international standard for information security management systems. This standard requires systematic risk assessment, access control policies, incident response procedures and regular security audits. By using the university server as the primary storage location, the research data automatically benefits from this certified security framework.

GDPR (General Data Protection Regulation): As the research involves personal data from human participants, full compliance with GDPR (EU) 2016/679 is required. Specific measures implemented: — Lawful basis for processing: informed consent (Article 6(1)(a) and Article 9(2)(a)) — Data minimisation: only data strictly necessary for the research is collected — Storage limitation: personal data retained only as long as necessary; deletion schedule defined and documented — Security of processing (Article 32): encryption, access control and pseudonymisation applied — Rights of data subjects: participants informed of their right to withdraw consent and request deletion of their data at any time before anonymisation is complete

UK Data Service and DataONE best practices: Storage and backup procedures in this plan follow the guidance published by the UK Data Service on data storage and the DataONE Best Practices for managing research data. Specifically: use of institutionally managed storage as primary location, 3-2-1 backup rule, automatic backup preferred over manual processes, and clear assignment of named responsibilities for backup and recovery.

Institutional data policy: The research will comply with the data management and research ethics policies of the affiliated university. Prior to data collection, the research will be submitted for review by the university's institutional ethics committee or equivalent body. Any specific requirements arising from that review will be incorporated into this data management plan as an amendment.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

Data that must be retained or destroyed for legal or regulatory purposes

Several categories of data in this research are subject to specific legal retention or destruction requirements under GDPR and Ukrainian legislation on personal data protection:

Mandatory retention: — Signed informed consent forms must be retained for a minimum of **5 years** after the completion of the research (thesis defence) in accordance with standard institutional ethics requirements. This ensures that consent can be verified if any participant dispute or ethics review arises after publication. Consent forms are stored securely on the university server with restricted access and are never published or shared. — The anonymised research dataset and accompanying documentation must be retained for a minimum of **10 years** after publication in accordance with open science principles and the data availability requirements of academic journals in the field of educational research.

Mandatory destruction: — Audio recordings of interviews must be deleted within **30 days** of the corresponding transcript being verified as accurate. Audio files contain identifiable voices and cannot be fully anonymised; retaining them beyond the transcription verification period constitutes unnecessary processing of personal data under GDPR Article 5(1)(e) (storage limitation principle). — The participant mapping table (linking participant codes P01–P15 to real identities) must be permanently deleted upon thesis defence. After this point the coded dataset is self-sufficient and the mapping table serves no further research purpose. — Raw Google Forms export files containing any potentially identifying metadata (e.g., response timestamps that could narrow down participant identity in a small institution) must be replaced by the cleaned anonymised version within 60 days of data collection completion. The raw export will be deleted after the cleaning process is verified.

Decision framework for retaining other data

Beyond legally mandated categories, the following criteria will be applied to decide which data to retain for the long term:

Retain if: — The data cannot be recreated without significant cost or effort (e.g., survey responses from 150–200 participants represent months of recruitment and cannot be recollected) — The data has clear reuse value for other researchers (e.g., the anonymised survey dataset on AI tool adoption among Ukrainian university teachers fills a recognised gap in existing open datasets) — The data is needed to verify or reproduce the research findings (reproducibility requirement of academic publishing) — The data forms part of the audit trail demonstrating research integrity (codebook, cleaning log, README documentation)

Do not retain if: — The data contains personal information that cannot be fully anonymised (audio recordings, consent forms beyond the mandatory period) — The data is superseded by a processed version that contains all necessary information (raw Google Forms export replaced by cleaned CSV) — The data has no independent scientific value beyond what is reported in the thesis (interim draft files, working notes, preliminary analysis versions) — Retaining the data would create disproportionate storage or maintenance costs relative to its reuse value

Foreseeable research uses for the data

The anonymised dataset produced by this research has clear and concrete potential for reuse across several contexts:

Validation of research findings: Other researchers can download the dataset from Zenodo and independently verify the statistical analyses reported in the thesis. This is increasingly required by journals publishing educational research and directly supports research integrity.

Comparative and longitudinal studies: The dataset provides a baseline measurement of AI tool adoption among Ukrainian university teachers in 2024–2025. Future researchers can conduct follow-up surveys using the same instrument to measure change over time — for example, assessing the impact

of national AI literacy programmes introduced after 2025. The consistent variable naming and codebook make such comparisons straightforward.

Cross-national comparisons: Ukrainian higher education is significantly underrepresented in international datasets on AI adoption in education. This dataset can be integrated with similar datasets from other countries (e.g., the Zenodo dataset DOI: 10.5281/zenodo.8234567 covering other European contexts) to enable cross-national meta-analyses.

Development of methodological support materials: The findings on barriers, needs and current practices identified in the survey and interviews can directly inform the development of training programmes, guidelines and methodological recommendations for university teachers — which is precisely the applied goal of this research area.

Teaching and curriculum development: The dataset and research instruments (survey questionnaire, interview guide, codebook) can be used in university courses on research methodology, educational technology and data management as worked examples of a complete research data lifecycle. This reuse requires no additional preparation as the materials will already be fully documented and openly licensed.

Secondary analysis by policymakers: Aggregate findings from the dataset — such as the proportion of teachers using AI tools by discipline or academic position — may be of direct interest to university administrators and national education policy bodies when designing professional development programmes.

Retention and preservation schedule

The following schedule applies to all data categories in this research:

Data category	Retention period	Location after thesis	Action at end of period
Anonymised survey dataset (CSV)	Minimum 10 years	Zenodo (open access)	Review for continued value; extend if reuse ongoing
Interview transcripts — anonymised (TXT/PDF)	Minimum 10 years	Zenodo (restricted access)	Review for continued value
Survey instrument and interview guide (PDF)	Minimum 10 years	Zenodo (open access)	No action required
Codebook and README documentation (PDF/TXT)	Minimum 10 years	Zenodo (open access)	No action required
Signed consent forms (PDF)	5 years after defence	University server (restricted)	Secure destruction after 5 years
Audio recordings (MP3)	Max 30 days after transcription	University server only	Permanent deletion — verified and logged
Participant mapping table	Until thesis defence	University server (encrypted)	Permanent deletion on defence date
Raw Google Forms export	Max 60 days after collection	University server only	Deletion after cleaning verified
Working drafts and interim files	Duration of project only	Researcher's laptop / server	Deletion upon project completion
Thesis (final version)	Indefinitely	University repository + Zenodo	No action required

Long-term preservation responsibility: Upon deposit on Zenodo, long-term preservation of the published dataset becomes the responsibility of CERN and the OpenAIRE infrastructure, which guarantees preservation for a minimum of 20 years. This relieves the researcher of any ongoing infrastructure obligation after the project ends. The persistent DOI ensures that the dataset remains findable and citable regardless of any changes to the researcher's institutional affiliation after graduation.

Preparation effort for sharing: No significant additional effort is required to prepare the data for

sharing because open formats (CSV, TXT, PDF) are used throughout the research. No format conversion will be needed at the point of deposit. The codebook and README are maintained continuously during the project rather than created retrospectively, meaning the deposit-ready documentation package will be available immediately upon thesis defence.

What is the long-term preservation plan for the dataset?

Long-term preservation plan for the dataset

The long-term preservation of the research dataset will be ensured through deposit in **Zenodo** (zenodo.org), an open-access repository operated by CERN (European Organization for Nuclear Research) under the European OpenAIRE programme. Zenodo was selected as the primary preservation repository for the following reasons:

— **Institutional stability:** Zenodo is operated by CERN, one of the world's most established scientific institutions, founded in 1954. CERN explicitly commits to long-term data preservation as part of its core mission, with a guaranteed minimum preservation period of **20 years**. — **Persistent identifiers:** Every deposit receives a permanent DOI (Digital Object Identifier) registered with DataCite. This DOI remains valid and resolvable regardless of any future changes to the researcher's institutional affiliation, the university's IT infrastructure, or the researcher's personal contact details. — **Open access and discoverability:** Deposited datasets are indexed by Google Dataset Search, OpenAIRE, DataCite Search and BASE (Bielefeld Academic Search Engine), ensuring long-term discoverability without any maintenance effort by the researcher after deposit. — **Format sustainability:** Zenodo accepts and preserves the open formats used in this research (CSV, TXT, PDF), all of which are on the UK Data Service list of recommended preservation formats. No proprietary formats requiring specific software licences are used. — **Version control:** Zenodo supports dataset versioning, assigning a new DOI to each updated version while maintaining a concept DOI that links all versions. This ensures that citations remain valid even if the dataset is corrected or extended after initial publication. — **Cost:** Zenodo provides free preservation for datasets up to 50 GB per record, well above the estimated 20 MB size of the final published dataset. No ongoing fees are required.

What will be deposited and when

Content of the deposit: The Zenodo deposit will contain the complete, self-sufficient dataset package required for independent reuse: — Anonymised survey dataset: anketa_AI_vykladachi_FINAL.csv — Anonymised interview transcripts: interview_transcripts_anonymised.pdf — Survey instrument (questionnaire in full): survey_instrument_EN_UA.pdf — Interview guide: interview_guide.pdf — Codebook describing all variables: codebook.pdf — README file with full project description: README.txt — Data cleaning log documenting all processing steps: cleaning_log.txt

Timing of deposit: The dataset will be deposited on Zenodo within **30 days of thesis defence**. An embargo of 6 months will be applied at the time of deposit, during which the metadata (title, description, keywords, DOI) will be publicly visible and searchable but the data files themselves will not be downloadable. The embargo period allows time for preparation and submission of a journal article based on the research findings. Upon expiry of the embargo, the dataset will become fully open automatically without requiring any further action by the researcher.

Metadata and documentation for future users

To ensure that the dataset remains interpretable and reusable by secondary users in the long term — including users who have no contact with the researcher — the deposit will include complete metadata and documentation:

Repository-level metadata will be entered in Zenodo using the **DataCite Metadata Schema**, covering: title, creators with ORCID identifiers, description, keywords, publication date, resource type, licence, related publications (thesis DOI), funding information, and language. This metadata is harvested automatically by major academic search engines.

File-level documentation is provided through the codebook and README, which together contain all information needed to understand and reuse the data without any external assistance: variable definitions, coding schemes, units of measurement, missing value codes, data collection methodology, sampling approach, known limitations, and software recommendations for opening each file type.

Together these two layers of documentation ensure that the dataset meets the **FAIR principles** in full:

FAIR principle	How it is met
Findable	Persistent DOI; rich metadata indexed by Google Dataset Search, OpenAIRE, DataCite
Accessible	Openly downloadable from Zenodo after embargo; no registration required; standard HTTPS protocol
Interoperable	Open formats (CSV, TXT, PDF); Dublin Core and DataCite metadata standards; consistent variable naming
Reusable	CC BY 4.0 licence; complete codebook and README; data cleaning log; survey instrument published alongside data

Responsibilities after project completion

Upon thesis defence and Zenodo deposit, ongoing preservation responsibility transfers to the Zenodo/CERN infrastructure. No active maintenance is required from the researcher. However, the following commitments are made for the post-project period:

— The researcher will maintain the email address associated with the Zenodo account for a minimum of 5 years after deposit to allow contact from secondary users with questions about the dataset — If any errors are discovered in the published dataset after deposit, a corrected version will be uploaded to Zenodo using the versioning system, with a clear changelog note describing what was corrected and why — If the researcher changes institutional affiliation after graduation, the Zenodo account and associated DOIs will be updated to reflect the new contact details — The ORCID iD linked to the deposit provides a persistent researcher identifier that remains valid regardless of institutional changes, ensuring that the dataset remains correctly attributed throughout its preservation period

Data Sharing

How will you share the data?

How the data will be shared

The anonymised research dataset will be shared openly following the completion of the research through **Zenodo** (zenodo.org). The sharing approach is designed to maximise accessibility and reuse while fully respecting participant confidentiality and allowing time for academic publication.

Access model

Primary dataset (anonymised survey data + documentation): Published on Zenodo under **open access** following a 6-month embargo period after thesis defence. Any user worldwide will be able to discover, view and download the dataset without registration, login or payment. Access requires only a standard web browser and an internet connection.

Interview transcripts (anonymised): Published on Zenodo under **restricted access**. The metadata and description will be publicly visible and searchable, but the transcript files will be available upon reasonable request only. Requests will be reviewed by the researcher to confirm that the intended use is consistent with the original consent obtained from participants. This model — open metadata, restricted files — is recommended by the UK Data Service for qualitative data where residual re-identification risk cannot be fully eliminated despite anonymisation.

Research instruments (survey questionnaire, interview guide, codebook, README):

Published on Zenodo under full **open access** with no embargo. These methodological materials contain no personal data and can be shared immediately upon deposit, allowing other researchers to assess and replicate the study design independently of the data release timeline.

Licence

All openly shared materials will be published under a **Creative Commons Attribution 4.0 International (CC BY 4.0)** licence. This licence: — Permits any user to copy, redistribute, adapt and build upon the data for any purpose, including commercial use — Requires only that the original dataset is cited with its full DOI reference — Imposes no restrictions on the format of reuse or the type of derivative works produced — Is the licence recommended by Zenodo, the UK Data Service and the majority of open science funders for research datasets

The CC BY 4.0 licence was specifically chosen over more restrictive alternatives (CC BY-NC, CC BY-ND) to minimise barriers to reuse. Restricting commercial use or derivatives would prevent legitimate reuse by educational technology companies, policy consultancies and international research teams, which would reduce the real-world impact of the research findings.

How secondary users will find the data

The dataset will be discoverable through multiple independent channels without any ongoing effort by the researcher after deposit:

— **Zenodo search:** directly searchable by title, keywords and description on zenodo.org — **Google Dataset Search** (datasetsearch.research.google.com): Zenodo metadata is automatically harvested and indexed, making the dataset discoverable via Google — **OpenAIRE:** the European open science infrastructure automatically indexes all Zenodo deposits — **DataCite Search:** the DOI registration with DataCite makes the dataset findable through the global DOI resolution system — **ORCID profile:** the dataset DOI will be added to the researcher's ORCID profile, linking it permanently to the researcher's academic identity — **Thesis data availability statement:** the published thesis will include a formal data availability statement in the following format:

«The anonymised dataset supporting the findings of this study is openly available on Zenodo at https://doi.org/10.5281/zenodo.XXXXXXX»

This statement ensures that anyone reading the thesis or any journal article derived from it can locate the dataset directly.

Timeline for sharing

Data category	Embargo period	Available from
Survey instrument, interview guide, codebook, README	None	Immediately upon Zenodo deposit (within 30 days of thesis defence)
Anonymised survey dataset (CSV)	6 months after thesis defence	Automatically on embargo expiry date
Anonymised interview transcripts	Restricted access (no embargo)	Upon reasonable request from deposit date
Audio recordings	Not shared	Permanently deleted — never deposited
Consent forms	Not shared	Retained institutionally for 5 years then destroyed

Anticipated reuse and citation

Upon open publication the dataset is expected to be of direct value to: — Researchers studying AI adoption in higher education, particularly in Eastern European and post-Soviet contexts that are currently underrepresented in the international literature — University administrators and policymakers designing professional development programmes for academic staff — Educational technology developers seeking evidence-based data on teacher needs and barriers — Lecturers in research methodology courses looking for worked examples of a complete open dataset with full documentation

To facilitate correct citation the Zenodo deposit page will display a ready-formatted citation in multiple styles (APA, BibTeX, DataCite, RIS), which secondary users can copy directly. The dataset will also be registered with **DataCite** upon deposit, ensuring it appears in bibliometric databases alongside

traditional journal publications and contributes to the researcher's academic citation record.

Communication of data availability

Beyond the thesis data availability statement, the existence of the dataset will be communicated through: — A dedicated section in any journal article or conference paper published from this research — The researcher's ORCID profile and academic profile page at the affiliated university — Submission to relevant Zenodo communities such as «Educational Technology» and «Open Science» to increase visibility within the target research community

Are any restrictions on data sharing required?

While the overall approach to data sharing in this research is as open as possible, certain restrictions are necessary for specific data categories. Each restriction is justified by a specific legal, ethical or practical reason and is proportionate — no restriction is applied beyond what is genuinely required.

Restriction 1 — Embargo on the primary dataset (6 months)

What is restricted: The anonymised survey dataset (CSV) and associated processed outputs will not be publicly downloadable for a period of 6 months following thesis defence, despite being deposited on Zenodo immediately after defence.

Why: The embargo period is required to allow preparation and submission of a peer-reviewed journal article based on the research findings. Releasing the full dataset before publication could compromise the novelty of the findings and reduce the likelihood of journal acceptance, as some publishers consider pre-released data a form of prior publication. This is standard academic practice and is explicitly supported by Zenodo's embargo functionality.

Duration: 6 months from the date of thesis defence. The embargo will be registered at the time of Zenodo deposit and will lift automatically on the specified date without requiring any action by the researcher.

What remains openly accessible during the embargo: The dataset metadata (title, description, keywords, DOI, licence) will be fully publicly visible and searchable from the date of deposit. The research instruments — survey questionnaire, interview guide and codebook — will also be openly available immediately, as they contain no personal data and their early release supports methodological transparency.

Restriction 2 — Restricted access to interview transcripts (ongoing)

What is restricted: The anonymised interview transcripts will be deposited on Zenodo with restricted access — meaning the files are not openly downloadable but are available upon reasonable request reviewed by the researcher.

Why: Although all transcripts will be anonymised prior to deposit, qualitative interview data carries a residual re-identification risk that cannot be fully eliminated. With only 10–15 participants, the combination of academic position, discipline, years of experience and specific opinions expressed could potentially allow identification of individuals within a small institutional context. This risk is recognised by the UK Data Service as a standard challenge with qualitative research data.

Restricted access — rather than full open access — allows the data to be shared with legitimate researchers while providing a review step that reduces the risk of misuse or inadvertent identification. This approach is specifically recommended by the UK Data Service for qualitative datasets where anonymisation alone cannot guarantee full confidentiality.

How access requests will be handled: Any researcher wishing to access the transcripts will submit a brief request via the Zenodo messaging system describing their intended use. The researcher will review the request and respond within 14 days. Access will be granted if the intended use is consistent with the original consent obtained from participants — that is, for non-commercial academic research purposes. Access will be declined if the request appears to involve attempts to identify participants or uses inconsistent with the original consent.

Duration: The restricted access model applies indefinitely for the full retention period of the

transcripts (minimum 10 years). It will not be converted to full open access without a further ethics review confirming that re-identification risk has become negligible over time.

Restriction 3 — Permanent non-disclosure of certain data categories

What is restricted: The following data categories will never be shared under any circumstances:

Data category	Reason for permanent non-disclosure
Audio recordings of interviews (MP3)	Contain identifiable voices; cannot be anonymised; will be permanently deleted within 30 days of transcript verification
Signed informed consent forms	Contain participant names and signatures — directly identifying personal data; retained securely for 5 years then destroyed
Participant mapping table (codes to real names)	Directly identifies all interview participants; deleted permanently upon thesis defence
Raw Google Forms export with timestamps	Timestamps combined with small institutional context could narrow participant identity; replaced by cleaned anonymised version

These restrictions are not discretionary — they are required by **GDPR Article 5(1)(b) and (e)** (purpose limitation and storage limitation), which prohibit processing personal data beyond the original purpose for which consent was obtained and beyond the period necessary for that purpose. Sharing these materials, even with trusted researchers, would constitute a breach of the informed consent agreement signed by participants.

Restriction 4 — No restriction (confirmatory statement)

Beyond the categories described above, **no other restrictions on data sharing are applied.** Specifically:

— No funder restrictions apply (this research is not externally funded and therefore subject to no funder data sharing mandate or embargo requirement beyond those chosen by the researcher) — No patent or commercialisation restrictions apply (the research produces no patentable outputs) — No third-party data licence restrictions affect the sharing of the primary dataset (all reused external sources are openly licensed under CC BY or equivalent) — No institutional confidentiality agreements restrict publication of the findings or data

This confirmatory statement is included to demonstrate that restrictions have been actively considered and that those applied are the minimum necessary — in line with the principle of being «as open as possible, as closed as necessary» advocated by the European Commission and the UK Research and Innovation (UKRI) open data policy.

Summary of sharing restrictions

Data category	Sharing model	Restriction type	Duration
Anonymised survey dataset (CSV)	Open access after embargo	Embargo — 6 months	6 months post-defence
Research instruments + codebook + README	Full open access	No restriction	Available immediately on deposit
Anonymised interview transcripts	Restricted access	Ongoing review-based access	Indefinite (minimum 10 years)
Audio recordings	Not shared	Permanent non-disclosure	Deleted within 30 days of transcription
Consent forms	Not shared	Permanent non-disclosure	Destroyed 5 years post-defence
Participant mapping table	Not shared	Permanent non-disclosure	Deleted on thesis defence date

Responsibilities and Resources

Who will be responsible for data management?

Data management responsibilities for this research are distributed across three roles. Each role has clearly defined tasks to ensure that no single point of failure exists in the data management process.

Primary responsibility — Researcher

The researcher is the **Data Manager** for this project and holds primary responsibility for all day-to-day data management activities throughout the research lifecycle. Specific responsibilities include:

— Implementing and maintaining the folder structure, file naming conventions and version control procedures described in this plan — Collecting, processing and anonymising all primary data in accordance with the procedures documented in this DMP — Performing weekly manual backup synchronisation to the local laptop copy and monitoring the automated Google Drive sync for errors — Maintaining and updating the README.txt file, codebook and data cleaning log throughout the project — Ensuring that audio recordings are deleted within 30 days of transcript verification and that the participant mapping table is deleted upon thesis defence — Preparing the final dataset package for Zenodo deposit within 30 days of thesis defence — Responding to data access requests for restricted-access interview transcripts within 14 days of receipt — Updating the Zenodo record if errors are discovered after publication using the versioning system

Supervisory responsibility — Research Supervisor

The research supervisor holds **oversight and quality assurance responsibility** for data management. Specific responsibilities include:

— Reviewing and approving this Data Management Plan prior to the start of data collection — Conducting a brief data management review at each monthly supervision meeting to confirm that backup procedures, naming conventions and anonymisation standards are being followed correctly — Reviewing the anonymised dataset before Zenodo deposit to confirm that no re-identification risk remains — Serving as the named contact for data management in the event that the researcher is temporarily or permanently unable to fulfil their responsibilities — Retaining read-only access to the university server project folder throughout the research to ensure continuity of access in case of emergency

Institutional responsibility — University IT Department

The university IT department holds **infrastructure responsibility** for the primary storage environment. Specific responsibilities include:

— Maintaining the university server on which the primary copy of all research data is stored — Performing automated daily server-level backups with a recovery point objective (RPO) of 24 hours — Ensuring that the server infrastructure complies with **ISO/IEC 27001** information security management standards — Providing technical support in the event of server failure, data corruption or a security incident affecting the primary storage location — Granting and revoking folder-level access permissions on the university server as directed by the researcher

Long-term responsibility after project completion — Zenodo / CERN

Upon deposit of the final dataset on Zenodo, long-term preservation responsibility transfers to **CERN and the OpenAIRE infrastructure**. After this point no active maintenance is required from the researcher or supervisor. CERN's specific responsibilities include:

— Ensuring physical preservation of deposited files for a minimum of 20 years — Maintaining the persistent DOI so that the dataset remains citable and findable indefinitely — Providing open access download functionality for the published dataset — Notifying the depositor in the event of any technical issue affecting the deposit

Summary table

Responsibility area	Responsible person / body	Period
Day-to-day data management	Researcher	Throughout research project
Backup monitoring and execution	Researcher (daily/weekly) + University IT (automated daily)	Throughout research project
Quality assurance and oversight	Research supervisor	Throughout research project
Ethics and anonymisation compliance	Researcher + Research supervisor	Throughout research project
Dataset preparation and Zenodo deposit	Researcher	Within 30 days of thesis defence
Access request management	Researcher	Minimum 5 years after deposit
Long-term digital preservation	Zenodo / CERN	Minimum 20 years after deposit
Secure destruction of personal data	Researcher	Per retention schedule

What resources will you require to deliver your plan?

What resources will you require to deliver your plan?

The data management activities described in this plan have been designed to be deliverable within the existing resources available to the researcher and the affiliated university. No significant additional budget is required. The following resource assessment covers personnel time, technical infrastructure, software tools and financial costs.

Personnel time

Implementing this data management plan requires a realistic allocation of the researcher's time throughout the project. The following estimates are based on the specific activities described in this DMP:

Activity	Estimated time investment
Initial setup: folder structure, naming conventions, README, codebook template	4–6 hours (one-time, project start)
Survey instrument design and piloting	8–10 hours
Data collection monitoring and Google Forms management	2–3 hours per week during collection period
Weekly backup synchronisation and monitoring	15–20 minutes per week throughout project
Data cleaning, anonymisation and cleaning log	10–15 hours after collection completion
Interview transcription and verification	3–4 hours per interview (30–45 min audio) × 15 interviews = 45–60 hours
Thematic coding of interview transcripts	20–30 hours
Codebook finalisation and README update for deposit	4–6 hours before Zenodo deposit
Zenodo deposit preparation and submission	2–3 hours
Responding to data access requests (estimated 5 requests/year)	1 hour per request

The most significant time investment is interview transcription, which is an unavoidable requirement of qualitative research methodology rather than an additional data management burden. All other data management activities are estimated at **approximately 30–40 hours total** across the full

project duration of 12 months — equivalent to less than one working week spread across the year. This is consistent with standard data management overhead for a project of this scale and does not require dedicated data management personnel beyond the researcher.

The research supervisor's data management oversight role is estimated at **1–2 hours per month**, integrated into scheduled supervision meetings rather than requiring additional dedicated time.

Technical infrastructure

All technical infrastructure required to deliver this plan is already available at no additional cost:

Infrastructure	Provider	Cost	Status
Primary data storage server	University IT department	No additional cost — standard allocation	Available
Automated daily server backup	University IT department	No additional cost — standard service	Available
Cloud backup (Google Drive, 15 GB)	Google (researcher's account)	Free	Available
Laptop with full-disk encryption	Researcher's personal device	No additional cost	Available
University VPN for remote server access	University IT department	No additional cost	Available
Zenodo repository (up to 50 GB per record)	CERN / OpenAIRE	Free	Available
DMPonline (this DMP platform)	Digital Curation Centre	Free for researchers	Available

Software tools

All software required for data collection, processing, analysis and deposit is freely available:

Task	Software	Cost
Survey administration	Google Forms	Free
Interview recording	Zoom (university licence)	Free via university
Data cleaning and analysis	Python 3 (pandas, matplotlib)	Free open source
Spreadsheet analysis	LibreOffice Calc	Free open source
Document preparation	LibreOffice Writer	Free open source
PDF creation	LibreOffice export / PDF24	Free
Qualitative coding	Manual coding in LibreOffice	Free
Checksum generation (MD5)	Built-in OS tools / md5sum	Free
Zenodo deposit and metadata entry	Zenodo web interface	Free

No specialist or licensed software is required at any stage of the research or data management workflow. The deliberate choice of open-source and freely available tools ensures that the research workflow is fully reproducible by any secondary user regardless of institutional affiliation or software budget.

Financial costs

A full cost assessment confirms that no additional budget is required to implement this data management plan:

Resource category	Estimated cost
Data storage (university server)	€0 — covered by standard IT allocation
Cloud backup (Google Drive)	€0 — free tier sufficient (200 MB << 15 GB)
Repository deposit (Zenodo)	€0 — free for datasets under 50 GB
Software licences	€0 — all tools are free or open source
Personnel (researcher time)	€0 — data management integrated into standard research activities
Data management training	€0 — DMPonline and UK Data Service guidance freely available online
Ethics committee submission	€0 — standard university process with no fee
DOI registration	€0 — assigned automatically by Zenodo at no cost
Total additional cost	**€0**

This zero-cost outcome is achievable because the research was planned from the outset with open infrastructure and open formats in mind. Had proprietary tools (e.g., SPSS, NVivo, paid cloud storage) or large data volumes been involved, additional costs would need to be budgeted. Researchers planning larger or more complex studies should consult their institution's research office or library for guidance on costing data management activities in grant applications.

Skills and training

The researcher has the necessary skills to implement this plan, having completed coursework in research data management, open science principles and data analysis tools. Specific competencies relevant to this plan include:

— Familiarity with Zenodo deposit procedures and DOI registration — Practical experience with CSV data cleaning using Python (pandas) — Understanding of Dublin Core and DataCite metadata standards — Knowledge of GDPR requirements for research data involving human participants — Experience with DMPonline for data management planning

If any knowledge gaps are identified during the project, the following free resources will be used: — **UK Data Service training materials** (ukdataservice.ac.uk/learning-training) — **OpenAIRE resources on FAIR data** (openaire.eu) — **Zenodo user documentation** (help.zenodo.org) — **University library research support service** — available to all registered researchers at no cost