Plan Overview

A Data Management Plan created using DMPonline

Title: Seeing is Believing: Behavioural Profiling and AI Previews to Build Trust in Pest Management Tools

Creator: Oluseye Oludoye

Principal Investigator: Oluseye Oludoye

Contributor: Nasamu Musa, Mark Ramsden, Holly Alpren, Chidimma Opara

Affiliation: Teesside University

Funder: Economic and Social Research Council (ESRC)

Template: ESRC Template

Project abstract:

Although digital tools exist to help farmers reduce pesticide use, many UK smallholders still avoid them. This project focuses on one such tool: *IPM Decisions*, a pan-European platform offering real-time pest forecasts to support Integrated Pest Management (IPM). Despite its scientific value, smallholder farmers often find these tools hard to trust, confusing to explore, or irrelevant to their day-to-day needs. In collaboration with ADAS and DEFRA, this project asks: Can we make decision-support tools more appealing by showing farmers what they *missed*? We will co-design a simple “preview tool” that uses open-access pest data to show what the past season’s pest risks looked like for a given crop and location. The tool will also let users “play” with basic settings, like adjusting spray dates, to explore how different decisions might have changed outcomes. This interactive feature is designed to build confidence, curiosity, and trust without needing a full account or training. Working closely with smallholders, we will test whether this low-effort, gamified insight improves trust, understanding, and willingness to use the full platform. A hands-on workshop will test the prototype with different types of users (e.g. tech-sceptical, cautious, overwhelmed), and the findings will inform policy and design improvements. By combining behavioural science and AI-driven personalisation, this project will help bridge the gap between innovation and adoption, making sustainable farming tools more useful, usable, and used by the people who need them most.

ID: 207987

Start date: 01-07-2026

End date: 31-07-2026

Last modified: 10-07-2026

Seeing is Believing: Behavioural Profiling and AI Previews to Build Trust in Pest Management Tools

Assessment of existing data

Provide an explanation of the existing data sources that will be used by the research project, with references

The project will use existing secondary data from the IPM Decisions platform and its associated model/API infrastructure. IPM Decisions is a pan-European online platform developed to give farmers and advisers access to a range of decision-support systems for Integrated Pest Management, including pest, weed and disease risk models. The platform was developed through the EU-funded IPM Decisions project and is now maintained through the IPM Decisions Initiative coordinated by ADAS.

The existing data sources to be used will be limited to publicly available or open-access pest-risk/model outputs and related platform/API information. These may include historical pest-risk outputs, crop/pest model information, location-based risk summaries, and related weather/model inputs where these are available through the IPM Decisions platform or its permitted API routes. The platform brings together decision-support models and meteorological data to support pest-risk forecasting across crops and regions.

These data will be used only to develop illustrative retrospective examples for the proposed preview tool. For example, the project may use previous-season pest-risk information to show what a farmer might have seen for a selected crop and location. The purpose is not to produce live agronomic recommendations, but to test whether retrospective pest-risk previews can make decision-support tools feel more understandable, relevant and trustworthy to farmers.

The project may also use technical documentation or published information about the IPM Decisions platform and its APIs to guide prototype design. Published descriptions of the platform explain that it provides access to multiple IPM decision-support systems and associated resources, and recent work describes the development of decision-support and weather APIs linked to the IPM Decisions infrastructure.

No identifiable farmer-level, user-level or farm-level secondary data will be used. The project will not analyse individual platform accounts, identifiable usage histories, commercial farm records, or personal data from IPM Decisions users. Existing data will be used only in anonymised, open-access, aggregate, or illustrative form to support prototype development and behavioural evaluation.

Provide an analysis of the gaps identified between the currently available and required data for the research

The existing IPM Decisions data and platform resources provide a strong technical foundation for the project, but they do not provide all the data required to answer the project’s behavioural research questions. The existing data are primarily designed to support pest-risk forecasting and Integrated Pest Management decision-making. They can provide historical or retrospective information on pest-risk patterns for selected crops, pests, locations and seasons. These data are therefore suitable for developing illustrative examples for the proposed preview tool, such as showing what pest-risk information might have looked like in a previous growing season.

However, there are several important gaps between the currently available data and the data required for this project:

1. **Limited behavioural data on non-users and hesitant users**

Existing IPM Decisions data can show what the platform offers and, where available, broad patterns of use. However, it does not explain why many farmers do not engage with the platform in the first place. It does not capture farmers' awareness, trust, perceived relevance, digital confidence, reliance on advisors, time constraints or concerns about complexity. These behavioural data must be collected through the project survey.

2. **Limited data on smallholder and small-scale farmer perspectives**
   The project specifically focuses on smallholder, small-scale and small/medium arable farmers. Existing platform data are unlikely to provide sufficient insight into how this group perceives and evaluates decision-support tools. The primary survey is therefore needed to understand the barriers and motivations affecting this target group.
3. **Limited evidence on pre-consultation barriers**
   The project focuses on the stage before farmers log in, select a model or use the full platform. Existing technical data do not show what would persuade a hesitant farmer to try a decision-support tool for the first time. The survey and workshop will therefore collect data on what kinds of previews, explanations, trusted messengers and support routes may reduce this initial barrier.
4. **Limited data on trust, interpretation and usability**
   Historical pest-risk outputs can be used to show what a decision-support tool might have provided in a previous season, but they do not indicate whether farmers find those outputs understandable, credible or useful. The project will collect primary data on how farmers interpret retrospective risk summaries, whether they trust them, and what design features would improve usability.
5. **Limited evidence on response to the proposed preview tool**
   The existing data can support the content of a prototype preview, but it cannot show whether the preview concept improves farmers' curiosity, perceived usefulness or willingness to try IPM Decisions. The project therefore needs primary data from survey questions and a co-design/prototype-testing workshop to evaluate farmer responses to the preview tool, including interactive "what-if" features.
6. **Limitations in farm-level specificity**
   The retrospective examples may rely on available model outputs, default assumptions or broader location-based information. They may not fully reflect individual farm conditions, management histories or field-level variation. The project will address this limitation by presenting the preview tool as an illustrative onboarding concept rather than a live agronomic recommendation tool, and by clearly communicating assumptions and limitations to participants.

Information on new data

Provide information on the data that will be produced or accessed by the research project

The project will produce and access several categories of data.

1. Primary survey data

The project will produce primary quantitative and qualitative survey data from up to 300 UK smallholder, small-scale, and small/medium arable farmers or growers involved in pest management decisions. The survey will be administered online through Jisc Surveys.

The survey data will include:

- eligibility and consent responses;
- farm background information, such as broad location, crop types, approximate arable area, farm scale and farming experience;
- current pest-management decision-making practices;

- sources of pest-management information and advice;
- awareness and use of IPM Decisions or similar decision-support tools;
- reasons for non-use or limited use of decision-support tools;
- Likert-scale responses on awareness, relevance, ease of use, trust, credibility, digital capability, time constraints, motivation and innovation style;
- responses to the proposed retrospective preview-tool concept;
- intentions to try a preview tool or IPM Decisions;
- preferred support routes and trusted messengers;
- optional open-text responses about barriers, concerns and suggestions;
- optional demographic and access information, such as age group, gender and internet access.

The main survey dataset will be anonymised and will not include names, contact details or identifiable farm-level information.

2. Administrative contact and prize-draw data

The project will also collect limited contact details from participants who choose to enter the survey prize draw and/or express interest in follow-up workshop participation. These data may include name, email address, phone number and preferred contact method. These details will be collected or stored separately from the anonymous survey responses and will be used only for prize-draw administration and/or workshop recruitment.

3. Workshop/co-design data

The project will produce qualitative data from a co-design/prototype-testing workshop with approximately 15 farmer participants. The workshop will take place at the ADAS office in Cambridge. Data produced may include:

- workshop consent records;
- written notes from discussions;
- participant feedback on the preview-tool prototype;
- responses to structured workshop activities;
- comments on usability, trust, relevance, clarity and accessibility;
- feedback on interactive “what-if” features, such as adjusting timing of intervention or risk thresholds;
- audio recordings and transcripts, where consent is given.

Workshop data will be anonymised before analysis. Names, farm names and other direct identifiers will be removed from transcripts and reports.

4. Secondary data accessed from IPM Decisions

The project will access existing secondary data from the **IPM Decisions** platform and related API/model resources. These will consist of publicly available or open-access pest-risk/model data, such as historical pest-risk outputs or model-based information for selected crops, pests, locations and seasons.

These data will be used to develop illustrative retrospective examples for the preview-tool prototype. They will not be used to make live agronomic recommendations, assess individual farms, or evaluate farmer performance. No identifiable farmer-level, user-level or farm-level secondary data will be requested, accessed or analysed.

5. Derived research data and outputs

The project will also produce derived data and analytical outputs, including:

- cleaned and anonymised survey datasets;
- coded open-text responses;
- summary statistics and cross-tabulations;
- behavioural segmentation/persona outputs;

- anonymised workshop transcripts or summaries;
- prototype logic notes and illustrative preview-tool outputs;
- policy/design recommendations;
- reports, presentations, briefings and academic outputs.

All data produced or accessed will be used only for the stated research purposes: understanding behavioural barriers to IPM decision-support tool uptake, evaluating the proposed retrospective preview-tool concept, and developing evidence-based recommendations for improving the accessibility, trustworthiness and uptake of digital pest-management tools.

Quality assurance of data

Describe the procedures for quality assurance that will be carried out on the data collected at the time of data collection, data entry, digitisation and data checking.

For the online survey, the questionnaire will be reviewed by the research team and ADAS before launch to ensure that questions are clear, relevant to the target participants and aligned with the project aims. The survey will be piloted internally before distribution to check routing, consent questions, response options, completion time and whether the wording is understandable. Jisc Surveys will be used to reduce manual data-entry errors, as responses will be entered directly by participants into the online form.

At the point of data collection, eligibility screening questions will be used to confirm that respondents are aged 18 or over, UK-based, involved in pest-management decisions, and connected to arable farming. The survey will use structured response formats where appropriate, including fixed categories and Likert-scale items, to support consistency and reduce ambiguity. Open-text questions will be optional and used only where more detailed explanation is needed.

After data export from Jisc Surveys, the dataset will be checked for completeness, eligibility and quality. The research team will screen for duplicate or near-duplicate responses, very short completion times, straight-line responses across Likert-scale items, inconsistent answers, missing data patterns, irrelevant open-text responses and responses suggesting automated or low-quality completion. Responses that are clearly ineligible, fraudulent, bot-generated or unusable will be excluded from analysis, with exclusion decisions documented.

For workshop data, consent will be confirmed before data collection. Where audio recording is used, recordings will be checked after the session to ensure they are complete and audible. Transcripts will be checked against the audio or workshop notes for accuracy. Any identifying details, such as names, farm names or specific locations, will be removed during anonymisation. Thematic coding of workshop and open-text survey data will be checked by the PI and discussed within the research team to ensure that interpretations are grounded in the data.

For secondary IPM Decisions data, only relevant publicly available or open-access model/pest-risk data will be accessed. The research team will document the source, date of access, crop/pest/location parameters, and any assumptions or limitations attached to the data. ADAS will provide technical sense-checking to ensure that the data are interpreted appropriately and that any retrospective examples used in the preview prototype are realistic and not presented as live agronomic advice.

All cleaned datasets, coding files, and analysis outputs will be version-controlled using clear file-naming conventions and stored securely on Teesside University OneDrive. Key cleaning and analysis decisions will be documented to support transparency, reproducibility and auditability.

Backup and security of data

Describe the data security and backup procedures you will adopt to ensure the data and metadata are securely stored during the lifetime of the project.

The online survey will be administered through Jisc Surveys, which provides an encrypted, University-approved platform for questionnaire data collection. Access to the survey account and exported data will be restricted to authorised members of the research team. After data export, research files will be stored on secure University storage, such as Teesside University OneDrive. Data will not be stored on personal cloud accounts.

Direct identifiers will be kept separate from research data. Contact details for prize draw administration, workshop recruitment, payment or travel reimbursement will be stored in a separate restricted-access folder from the anonymised survey and workshop datasets. Workshop audio recordings will be transferred to secure University storage as soon as possible after the session and deleted from the recording device after successful transfer. Transcripts will be anonymised before analysis.

Access permissions will be managed by the PI. Only authorised project team members who need access for data collection, administration, analysis or reporting will be granted access. Files will be protected through University login credentials, password protection, and institutional security arrangements. Any transfer of data between Teesside University and project partners will use secure methods, such as password-protected files and encrypted transfer. No identifiable participant data will be shared with ADAS or DEFRA.

Data and metadata will be backed up through the University’s U Drive. Version control will be maintained through clear file naming conventions, dated folders, and separate storage of raw, cleaned and analysed datasets. A data log will record key files, dates of collection/export, cleaning decisions, analysis versions, and any derived outputs.

All data will be retained, archived or securely deleted in accordance with Teesside University data management and retention policies, the approved ethics application, and any relevant licensing conditions attached to IPM Decisions data.

Management and curation of data

Outline your plans for preparing, organising and documenting data.

Survey data will be collected through Jisc Surveys and exported into standard formats such as Excel/CSV for cleaning and analysis. Raw survey exports will be saved separately from cleaned and analysed versions. The research team will use clear file naming conventions, including project name, data type, version number and date, for example: BRUK_IPM_survey_raw_v1_date, BRUK_IPM_survey_cleaned_v1_date, and BRUK_IPM_workshop_transcript_anonymised_v1_date.

A structured data folder will be created on Teesside University-approved secure storage. Files will be organised into separate folders for: raw survey data, cleaned survey data, consent records, contact/payment administration, workshop notes/audio/transcripts, secondary IPM Decisions data, analysis files, outputs, and documentation. Identifiable data, such as contact details for prize draw administration, workshop recruitment and travel reimbursement, will be stored separately from anonymised research data.

Survey variables will be documented in a data dictionary/codebook. This will include variable names,

question wording, response options, coding decisions, missing-value codes, reverse-coded items, derived variables and behavioural index/persona construction. Any changes made during data cleaning, such as removing ineligible responses or recoding categories, will be recorded in a cleaning log.

Workshop data will be organised by date and data type. Audio recordings, where consent is given, will be transcribed and anonymised. Transcripts will be given participant codes rather than names. A separate linkage file will be kept only where necessary and stored securely with restricted access. Thematic coding frameworks and coding notes will be documented to support transparency in qualitative analysis.

Secondary IPM Decisions data used for the prototype will be documented with source, date accessed, crop/pest/location parameters, model/output type and any assumptions or limitations. Derived preview-tool examples will be labelled clearly as illustrative research outputs, not live agronomic recommendations.

All datasets and metadata will be version-controlled. Key decisions on cleaning, anonymisation, exclusion criteria, coding and analysis will be documented so that the research process is transparent and reproducible.

Difficulties in data sharing and measures to overcome these

Identify any potential obstacles to sharing your data, explain which and the possible measures you can apply to overcome these.

There are several potential obstacles to sharing the full project dataset, but these can be managed through anonymisation, selective sharing and clear documentation.

The main obstacle is participant confidentiality. Although the online survey will be anonymous, some responses may include indirect identifiers, such as region, crop type, farm scale, unusual farming practices or detailed open-text comments. Workshop data may also contain contextual details that could make individual farmers, farms or businesses identifiable, particularly because the workshop sample will be small. To address this, all data will be reviewed and anonymised before any sharing. Direct identifiers will be removed, and indirect identifiers will be generalised or suppressed where necessary. Open-text responses and workshop quotations will be redacted or paraphrased if they could identify a participant, farm or business.

A second obstacle is that contact, payment, prize draw and travel reimbursement data cannot be shared. These records contain personal information and are collected only for administration. They will be stored separately from the research dataset and will not be included in any shared data.

A third obstacle relates to commercial or professional sensitivity. Farmers may provide views about pesticide use, pest management practices, confidence in decision-making, use of advisors or perceptions of digital tools. Even where these are not directly identifiable, sharing could risk reputational or commercial harm. To manage this, only anonymised and aggregated datasets will be considered for sharing, and no individual farm-level performance or commercially sensitive information will be released.

A fourth obstacle concerns workshop data. Because workshop discussions involve a small number of participants and group interaction, full transcripts may be difficult to anonymise completely. The project will therefore not share full raw workshop transcripts openly. Instead, anonymised extracts, coded themes, summary tables or derived qualitative findings may be shared where this can be done without compromising confidentiality.

A fifth obstacle relates to secondary IPM Decisions data and API-derived outputs. These data may be subject to platform terms of use, licensing conditions or restrictions set by data/model providers. The project will not redistribute secondary data unless permitted. Instead, the project will document data

sources, dates accessed, model/output types and parameters used. Where possible, links or references to publicly available sources will be provided rather than re-sharing restricted data.

The project will therefore apply a managed sharing approach. An anonymised survey dataset, codebook, questionnaire, and summary documentation may be shared where ethical, legal and licensing conditions allow. Data that cannot be safely anonymised, such as identifiable contact records, payment records, raw workshop transcripts, or restricted secondary data, will not be shared. This approach balances transparency and reuse with confidentiality, data protection and participant trust.

Consent, anonymisation and strategies to enable further re-use of data

Make explicit mention of the planned procedures to handle consent for data sharing for data obtained from human participants, and/or how to anonymise data, to make sure that data can be made available and accessible for future scientific research.

Consent for data sharing will be built into the participant information and consent process.

For the online survey, participants will be informed before taking part that their anonymised responses may be used in research reports, policy briefings, academic outputs and, where appropriate, future scientific research. The survey will be anonymous, and participants will not be asked to provide their name or direct identifying details as part of the main questionnaire. Any contact details provided for prize draw administration or follow-up workshop recruitment will be collected separately or stored separately from the survey dataset and will not form part of any shared research data.

For the workshop, participants will receive a separate Participant Information Sheet and Consent Form. This will explain how workshop data will be used, including anonymised reporting, use of anonymised quotations or summaries, and potential future reuse of anonymised findings. Participants will also be asked separately for consent to audio recording. If a participant does not consent to audio recording, written notes may be used instead where appropriate.

Before any data are made available for future research, the research team will anonymise the data. This will include removing names, contact details, farm names, exact locations, email addresses, phone numbers and any other direct identifiers. Indirect identifiers, such as very specific crop combinations, unusual farm characteristics, detailed locations, or open-text comments that could identify a participant or farm, will be generalised, redacted or paraphrased. Survey open-text responses and workshop quotations will be checked carefully before use or sharing.

The project will not share contact details, consent forms, prize draw records, payment records, travel reimbursement records, raw audio recordings or any data that cannot be safely anonymised. Full raw workshop transcripts are unlikely to be shared openly because the small number of participants and contextual detail may make complete anonymisation difficult. Instead, anonymised excerpts, coded themes, summary tables or derived qualitative findings may be shared where this can be done without compromising confidentiality.

Where appropriate, the project may make available an anonymised survey dataset, questionnaire, codebook, derived variables, and documentation explaining the data structure and coding. These shared materials will contain no direct identifiers and will be checked for disclosure risk before release. This approach will support future scientific research while protecting participant confidentiality, complying with data protection requirements, and respecting the consent given by participants.

Copyright and intellectual property ownership

State who will own the copyright and IPR of any new data that you will generate.

The copyright, database rights and other intellectual property rights in new data generated by the project will be held by Teesside University, as the Recipient institution and lead organisation for the funded project, subject to the terms of the Grant Funding Agreement.

New data generated by the project will include anonymised survey data, anonymised workshop/co-design data, derived behavioural segmentation outputs, analytical summaries, prototype logic documentation, and project reports or design/policy outputs. These will be treated as project “Results” under the funding agreement, which defines Results as information, data, techniques, know-how, results, software and materials developed in the course of the project.

Any pre-existing data, software, models, materials or intellectual property brought into the project by Teesside University, ADAS, DEFRA, IPM Decisions or any other third party will remain the property of the party that contributed it. The funding agreement makes clear that background intellectual property remains with the party that contributed it, or its licensors.

The project will not claim ownership over the underlying IPM Decisions platform, pest-risk models, APIs, software, or third-party model/weather data used to support the illustrative preview-tool prototype. These will remain owned by their existing rights holders. The project will only generate and own the new research data, analysis, documentation and project outputs created during the funded work.

Although Teesside University will hold rights in the new data generated by the project, these rights will be managed in line with the ESRC open science requirements. The agreement requires project Results to be made available on an open access basis and requires all data created or repurposed during the project to be made available for reuse or archiving within three months of the end of the project period.

Only data that can be ethically and legally shared will be made available. Identifiable participant data, contact details, consent records, payment/prize draw records, travel reimbursement records, raw audio recordings and any data that cannot be safely anonymised will not be shared openly. Anonymised datasets, codebooks, questionnaires, summary tables and derived research outputs may be shared or archived where this is consistent with participant consent, Teesside University policy, data protection law and any relevant third-party licensing conditions.

Responsibilities

Outline responsibilities for data management within research teams at all partner institutions

Teesside University will have overall responsibility for data management because it is the lead organisation and recipient of the BR-UK award. The project will be managed by **Dr Oluseye Oludoye**, who is identified in the funding agreement as the Project Lead responsible for managing the project on behalf of Teesside University. The funding agreement also requires Teesside University to obtain and maintain the necessary ethical approvals, keep accurate research records, and comply with relevant laws, regulations and funder requirements.

The **Principal Investigator, Dr Oluseye Oludoye**, will be responsible for ensuring that project data are collected, stored, analysed, anonymised, retained and shared in accordance with the ethics approval, this Data Management Plan, Teesside University policy, UK GDPR/Data Protection Act 2018 requirements, and BR-UK/ESRC open science expectations. He will oversee the survey dataset, workshop data, consent records, anonymisation procedures, file organisation, access permissions,

data documentation, and preparation of any anonymised data for archiving or reuse.

Teesside University research team members, including Dr Chidimma Opara, will support data cleaning, preparation, analysis and documentation. Access to raw or potentially identifiable data will be limited to authorised Teesside research team members only. Contact details for prize draw administration, workshop recruitment, payment and travel reimbursement will be stored separately from the anonymised research dataset.

ADAS will act as a subcontracted project partner. ADAS will support recruitment by circulating the survey invitation through its farmer/advisor networks, provide technical advice on IPM Decisions, support the workshop at the ADAS office in Cambridge, and help sense-check the prototype and interpretation of findings. ADAS will not be responsible for storing or managing the main research dataset. ADAS will not receive identifiable survey responses or identifiable participant contact details from Teesside University. ADAS may contribute to interpretation of anonymised or aggregated findings only. Where ADAS staff are involved in workshops, they will be expected to treat participant contributions confidentially and will not use any information outside the project purpose.

DEFRA, through Dr Holly Alpren, will provide policy advisory input and help ensure that outputs are relevant to current IPM and pesticide policy priorities. DEFRA will not collect, store or manage participant-level data. DEFRA will receive only anonymised, aggregated or summary-level findings for interpretation, policy relevance and dissemination purposes.

The University of Edinburgh/BR-UK is the funder/institution administering the award and may receive reports, results and project outputs in accordance with the funding agreement. It will not receive identifiable participant-level research data. The agreement requires project results and data created or repurposed during the project to be made available for reuse or archiving within the required timeframe, but this will be done only with anonymised data and where ethical, legal and licensing conditions allow.

All partners will follow the principle that identifiable participant data will be accessed only by those who need it for approved project administration or research purposes. Anonymised and aggregated data will be used for partner discussion, reporting, dissemination, archiving and future reuse wherever possible.

Preparation of data for sharing and archiving

Are the plans for preparing and documenting data for sharing and archiving with the UK Data Service appropriate?

Yes. The plans for preparing and documenting data for sharing and archiving with the UK Data Service are appropriate, provided that only anonymised and non-disclosive data are deposited.

The project will prepare a shareable version of the survey dataset by removing all direct identifiers and excluding any administrative data, such as names, email addresses, phone numbers, prize draw records, workshop recruitment details, payment records and travel reimbursement information. The survey is designed to collect mostly structured responses on farm background, pest-management decision-making, awareness and use of IPM Decisions, behavioural barriers, trust, motivation, reaction to the proposed preview tool, and intention to use decision-support tools. These data are suitable for anonymised secondary analysis if properly documented and disclosure-checked.

The project will also prepare supporting documentation for deposit, including the final questionnaire, variable labels, response codes, missing-value codes, derived variables, behavioural index/persona construction, data-cleaning decisions, and a short methodological note describing recruitment, eligibility criteria, survey mode, sample characteristics and limitations. This will make the dataset understandable and reusable by future researchers.

Qualitative open-text responses and workshop-derived materials will be reviewed carefully before any

sharing. Open-text responses will be anonymised, redacted or paraphrased where necessary to remove references that could identify individuals, farms, businesses, exact locations or unusual farming circumstances. Full raw workshop transcripts are unlikely to be suitable for open sharing because the workshop sample will be small and participants may be indirectly identifiable. Instead, anonymised excerpts, coded themes, summary tables or derived qualitative findings may be shared where this can be done safely.

Secondary IPM Decisions data or API-derived material will not be deposited unless licensing and access conditions permit this. Where these data cannot be shared directly, the project will document the source, date accessed, model/output type, crop/pest/location parameters and assumptions used, so that the basis of the prototype examples is transparent.

These plans align with the funding agreement, which states that data created or repurposed during the ESRC grant must be made available for reuse or archiving within three months of the end of the grant and refers to UK Data Service guidance on data management. The project will therefore deposit or archive the anonymised survey dataset and supporting documentation with the UK Data Service, or another appropriate repository if advised, while withholding any data that cannot be ethically or legally shared.

Is there evidence that data will be well documented during research to provide highquality contextual information and/or structured metadata for secondary users?

Yes. The project has clear procedures to ensure that data are well documented during the research and that sufficient contextual information and structured metadata are available for secondary users.

First, the survey instrument itself provides a clear structure for documentation. It records the project title, project focus, target respondents, estimated completion time, data-collection mode, consent process, eligibility criteria and question sections. These sections cover farm background, current pest-management decision-making, awareness and use of IPM Decisions, behavioural barriers, trust, motivation, reaction to the proposed preview tool, intention to use, follow-up interest, and optional demographic/access information.

Second, the dataset will be accompanied by a codebook. This will document each variable name, full question wording, response options, coding scheme, missing-value codes, derived variables, reverse-coded items and behavioural index construction. For example, the questionnaire already groups items into analytically meaningful sections such as awareness/relevance, ease/effort barriers, trust/credibility, capability/opportunity, motivation/innovation style, preview appeal and intention to use. The researcher notes also specify planned indices and suggested behavioural personas, such as low-trust sceptics, overwhelmed but open users, advisor-dependent users, pragmatic experimenters and digitally confident early adopters.

Third, contextual metadata will be recorded during the research process. This will include survey launch and closure dates, recruitment route, eligibility criteria, survey mode, sample size, inclusion/exclusion decisions, data-cleaning decisions, and any changes made to the questionnaire before launch. A cleaning and analysis log will record decisions such as removing ineligible responses, duplicate or low-quality responses, recoding variables, anonymising open-text responses and creating derived variables.

Fourth, workshop/co-design data will be documented with date, location, participant category, workshop activity type, prototype version tested, facilitation notes, consent status, recording/transcription status and anonymisation decisions. Any qualitative coding framework used for open-text or workshop data will be documented so that secondary users can understand how themes were developed.

Fifth, any secondary IPM Decisions data used for the retrospective preview prototype will be documented with source, date accessed, model or output type, crop/pest/location parameters, assumptions, limitations and any licensing conditions. This is important because the prototype will use illustrative retrospective pest-risk outputs and will not be presented as live agronomic advice.

Finally, the funding agreement requires project results to be made available on an open access basis and data created or repurposed during the project to be made available for reuse or archiving within three months of the end of the project period. The project’s documentation procedures are therefore designed to ensure that any anonymised data shared or archived are understandable, reusable and appropriately contextualised for future research.