

AI-Powered Anomaly Detection to Improve SharePoint Data Quality

Paul Griffiths

2410226

Computing (Data and Information
Systems)

Project Management & Planning

Project Scope and Proposed Solution

This project aims to design and evaluate an AI-powered anomaly detection solution that improves the quality of data in SharePoint Lists used for reporting. These Lists often suffer from issues like inconsistency and human error, impacting business decision-making. [1]

The solution will use Microsoft's Power Platform tools, including AI Builder, Power Automate and Power BI. [2] The system will detect unusual or invalid data entries, flag them and log the results for further analysis. Notifications will be sent when anomalies are found and results visualised using Power BI.

The project will use an iterative approach, allowing for incremental improvements. The solution will be piloted, evaluated and refined based on feedback.

Required Technologies and Resources

Key tools and resources include:

- Power Platform Tools
 - Power Automate - Triggers workflows and sends anomaly alerts
 - AI Builder - Trains a prediction model on labelled data to detect outliers
 - Power BI - Visualises anomaly trends to support insights and monitoring
- SharePoint
 - The source of the data
- Data Labelling and Feedback Tools
 - Excel – For manually labelling training data or analysing feedback
 - Microsoft Forms – To collect feedback from pilot users
- Hardware and Access
 - PC with organisational Microsoft 365 licence and SharePoint admin access
 - Access to Power Platform admin centre for usage tracking

The main constraint is licencing - AI Builder has limited free usage under Microsoft 365. [3]

Project Timeline and Schedule

The project spans nine weeks and includes research, development, testing and final reporting. Weekly goals help ensure progress, with time allocated for pilot testing and user feedback. Below is a Gantt Chart showing the planned schedule.

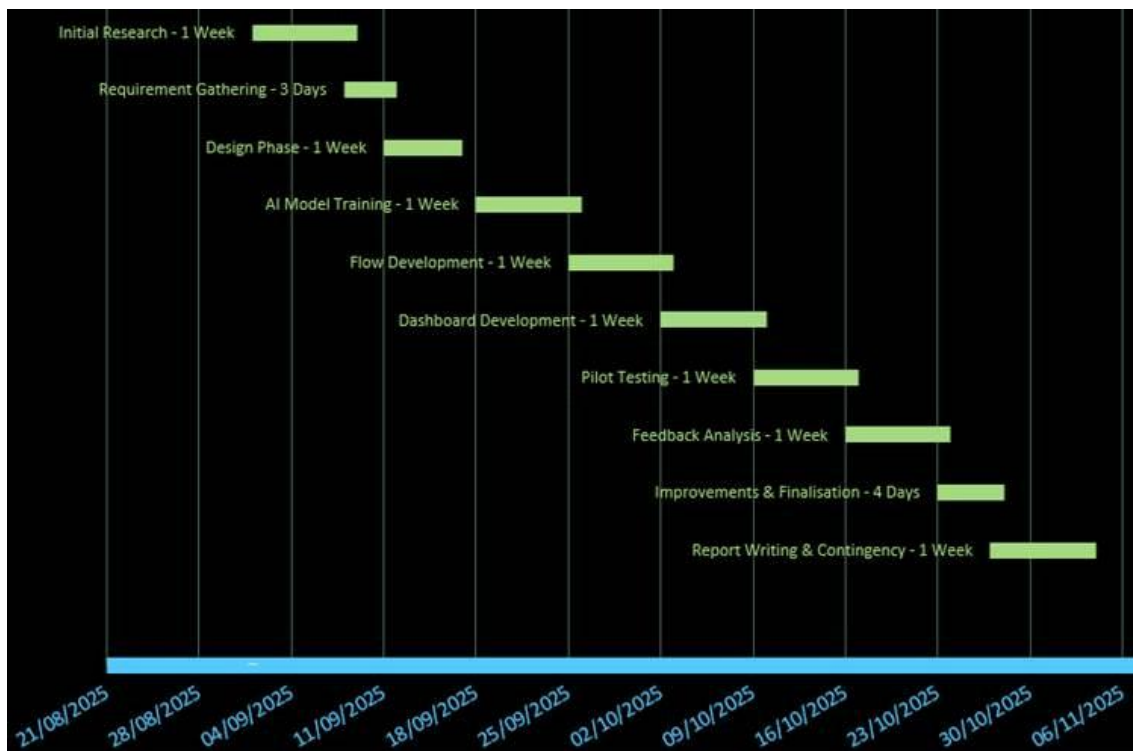


Figure 1 - Gantt Chart showing planned project phases [4]

Phases included are:

- Initial Research - Research SharePoint data issues and anomaly detection tools
- Requirement Gathering - Identifying data quality criteria
- Design Phase - Create system diagrams
- AI Model Training - Label data, train AI Builder model
- Flow Development - Build and test Power Automate flow
- Dashboard Development - Create Power BI dashboard
- Pilot Testing - Deploy solution to 5 users, collect usage data and feedback
- Feedback Analysis - Analyse feedback and performance
- Improvements & Finalisation - Make final changes based on pilot feedback
- Report Writing & Contingency - Write evaluation, reflect on success criteria. Allowance for contingency

Development Method

An incremental prototyping approach with agile thinking is the most suitable development method for this project as this allows for early testing, user feedback and adaptability - important when working with AI tools and changing requirements. [5]

The project will begin with a low-fidelity prototype, tested by a limited user group. This initial version, developed as a Minimum Viable Product (MVP), will help gather useful insights quickly without wasting time on unnecessary features. [6]

Pilot testing will form a feedback loop that enables incremental improvements, increasing accuracy and functionality with each iteration. This reflects agile principles like adaptability and collaboration. [7] Using low-code Power Platform tools also allows for quick updates, ensuring ineffective elements can be revised without disrupting the entire timeline.

By applying this method, the solution will be responsive to the needs of its users while being grounded in structured, iterative development cycles.

Project Design

Requirements Specification

The Requirements Specification below provides a structured foundation for the solution and is divided into functional (what the solution should do) and non-functional (how the solution should perform). [8]

Requirement Type	ID	Requirement Description
Functional	F1	The system should use AI Builder to detect anomalies in data entered into a specific SharePoint List
Functional	F2	The Power Automate flow should trigger automatically when new or modified data is detected in the SharePoint List
Functional	F3	The system should log all detected anomalies, including the ID, timestamp and anomaly reason, into a separate SharePoint List or Excel file
Functional	F4	The system should send notifications via email or teams to designated users when an anomaly is detected
Functional	F5	The Power BI dashboard should display visual representations of anomaly patterns, frequency and trends over time
Functional	F6	The system should allow manual override of detected anomalies by specified individuals
Functional	F7	A feedback form should be provided to gather qualitative insights from pilot users regarding the tool's usefulness and accuracy
Non-Functional	N1	The anomaly detection flow should respond within 2 minutes of a new record being added or updated
Non-Functional	N2	The solution should be implemented using only tools available within licensing already available to users
Non-Functional	N3	All components should adhere to data governance and security policies, including no use of personal or sensitive data
Non-Functional	N4	The design should be scalable to other SharePoint Lists with minimal changes such as changing list references or retraining the AI model
Non-Functional	N5	The solution should be maintainable by non-developers using standard Power Platform interfaces

Figure 2 - Table showing Requirements Specifications

Flow Chart

The flow chart below shows how the completed system will work from start to finish and helps identify where parts connect, how users use it and where feedback comes in.

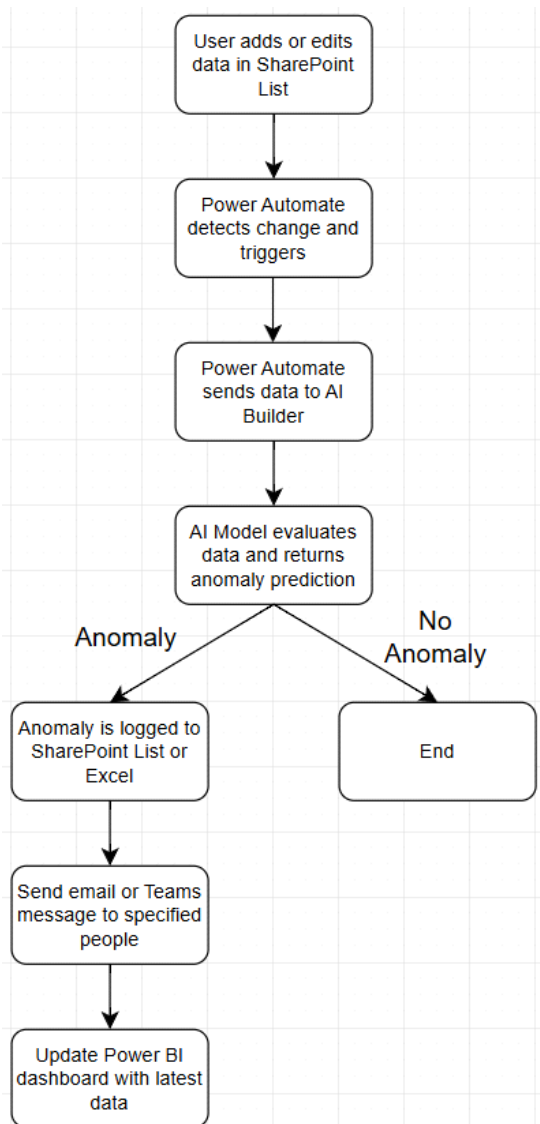


Figure 3 - Flow Chart showing process of final product [9]

Together, these two design components provide effective planning and reduce risks – important principles in successful project management.

References

[1]

S. Thorne, "Spreadsheet errors can have disastrous consequences – yet we keep making the same mistakes," *The Conversation*, Jan. 25, 2024. <https://theconversation.com/spreadsheet-errors-can-have-disastrous-consequences-yet-we-keep-making-the-same-mistakes-219356>

[2]

"AI-Powered Low-Code Tools | Microsoft Power Platform," *Microsoft.com*, 2024. <https://www.microsoft.com/en-gb/power-platform?msocid=385fa11d27e4669a3a9db5d3262f6749> (accessed Aug. 12, 2025).

[3]

Antoine2F, "AI Builder licensing and credit management," *Microsoft.com*, Nov. 02, 2024.
<https://learn.microsoft.com/en-us/ai-builder/credit-management>

[4]

Microsoft, "Gantt Charts Design Templates | Microsoft Create," *create.microsoft.com*, 2023.
<https://create.microsoft.com/en-us/templates/gantt-charts>

[5]

GeeksforGeeks, "Agile Prototyping," *GeeksforGeeks*, Feb. 05, 2024.
<https://www.geeksforgeeks.org/software-engineering/agile-prototyping/>

[6]

N. Raj, "Minimum Viable Product (MVP) - What is it & how to start," *Atlassian*, 2024.
<https://www.atlassian.com/agile/product-management/minimum-viable-product>

[7]

A. M. Carreno, "Building a Continuous Feedback Loop for Real-Time Change Adaptation: Best Practices and Tools," *Zenodo*, vol. 1, no. 1, Nov. 2024, doi:
<http://dx.doi.org/10.5281/zenodo.14051466>.

[8]

Grow Solutions, "Functional and Non-Functional Requirements: The Ultimate Checklist with Examples," *Medium*, Nov. 17, 2023. <https://medium.com/@growsolutions/functional-and-non-functional-requirements-the-ultimate-checklist-with-examples-cde16aba33d7>

[9]

draw.io, "Flowchart Maker & Online Diagram Software," *app.diagrams.net*, 2025.
<https://app.diagrams.net/>

PEST & SWOT Analysis

I have chosen Innovation for my project as of the two options, this is the one that best suits it. Microsoft is a good choice for the analysis as it provides the Power Platform tools - Power Automate, AI Builder, Power BI and SharePoint are all central to the proposed anomaly detection solution. [1] Understanding Microsoft's business can help assess the feasibility and scalability of implementing this solution in a large organisation.

PEST Analysis [2]

- **Political:** Microsoft benefits from strong government relationships and complies with global regulations like GDPR, important for data governance in enterprise solutions. [3]
- **Economic:** Microsoft is a trillion-dollar company and invests heavily in AI and cloud services, ensuring the long-term sustainability of platforms like AI Builder. [3]
- **Social:** There's an increasing need for data transparency and automation, supporting the adoption of AI-driven tools to improve reporting and reduce manual data error.
- **Technological:** Microsoft has done a lot with low-code and AI development, integrating tools like Copilot and Azure AI into Power Platform, making it suitable for this project's AI-driven approach. [3]

SWOT Analysis [4]

- **Strengths:** Trusted brand, enterprise-scale cloud platform, ready integration across tools.
- **Weaknesses:** AI Builder licensing can limit free use, technical documentation may be complex for new users. [5]
- **Opportunities:** Demand for automated data quality assurance, increased AI adoption in workplaces.
- **Threats:** Increasing concerns over AI bias and ethical use, rapid tech shifts could render current tools outdated. [6]

All of this suggests Microsoft is a good choice as its tools fit the project and it's built to scale up if the solution is rolled out across a larger organisation.

Deployment of the Anomaly Detection Project

If deployed in a large organisation, this AI-powered anomaly detection solution would become an important data quality layer across multiple departments that use SharePoint Lists for reporting and operational decisions. The scope would expand beyond a single pilot list to dozens of SharePoint Lists across finance, HR, ITS and operations. While personal or sensitive data would be avoided where possible, GDPR-compliant handling, access controls and audit logs would ensure data protection when unavoidable.

The innovation lies in the combination of low-code AI with business reporting. The idea is to bring smart tools straight into the reporting process - AI Builder will spot anomalous data, Power Automate will handle alerts or follow-up actions and Power BI will help people see the bigger picture through visual dashboards. This supports goals such as improving operational efficiency, reducing human error and promoting data-driven decision-making.

The project combines:

- **AI-powered anomaly detection:** identifying patterns outside the norm
- **Data quality improvement:** reducing reporting errors at the source
- **Low-code automation:** enabling fast and scalable rollout
- **Human-in-the-loop validation:** allowing overrides where AI misclassifies data

For enterprise rollout, tools would include:

- **Power Platform:** scalable through environment segmentation and governance controls [1]
- **Microsoft Dataverse:** for handling larger datasets and more complex models [7]
- **SharePoint Online:** the data entry and workflow hub [8]

Enterprise licensing models would need to factor in AI Builder capacity and premium connector access.

The solution would follow an agile, iterative approach, using feedback loops, versioning, and low-code DevOps practices. Components would be updated through Power Platform's built-in deployment tools.

Maintenance would be streamlined through Power Platform GUIs, allowing trained Power Users or IT Admins to adjust thresholds, retrain models or reconfigure flows. Governance would include role-based permissions, data access policies and usage monitoring dashboards.

Finally, legal and ethical practices would follow Microsoft's Responsible AI Standard, ensuring transparency, accountability and fairness in AI decisions, while maintaining compliance with GDPR, data residency rules and organisational data protection frameworks. [9]

References

[1]

"AI-Powered Low-Code Tools | Microsoft Power Platform," *Microsoft.com*, 2024.

<https://www.microsoft.com/en-gb/power-platform?msocid=385fa11d27e4669a3a9db5d3262f6749> (accessed Aug. 12, 2025).

[2]

W. Kenton, "What Is PEST Analysis? Its Applications and Uses in Business," *Investopedia*, Feb. 26, 2025.

<https://www.investopedia.com/terms/p/pest-analysis.asp>

[3]

"GDPR and Generative AI: A Guide for Public Sector Organizations," *Public Sector Center of Expertise*, Jul. 10, 2024. <https://wwps.microsoft.com/blog/gdpr-genai>

[4]

K. Peterdy, "SWOT analysis," *Corporate Finance Institute*, 2025.

<https://corporatefinanceinstitute.com/resources/management/swot-analysis/>

[5]

Antoine2F, "AI Builder licensing and credit management," *Microsoft.com*, Nov. 02, 2024.

<https://learn.microsoft.com/en-us/ai-builder/credit-management>

[6]

V. Maphosa, "The Rise of Artificial Intelligence and Emerging Ethical and Social Concerns," *AI, Computer*

[7]

“Microsoft Dataverse | Microsoft Power Platform,” *Microsoft.com*, 2025. <https://www.microsoft.com/en-us/power-platform/dataverse?msockid=385fa11d27e4669a3a9db5d3262f6749> (accessed Aug. 12, 2025).

[8]

Microsoft, “Office 365 Login | Microsoft Office,” *Office.com*, 2019. <https://www.office.com/>

[9]

“Responsible AI: Ethical policies and practices | Microsoft AI,” *Microsoft.com*, 2025. <https://www.microsoft.com/en-us/ai/responsible-ai?msockid=385fa11d27e4669a3a9db5d3262f6749> (accessed Aug. 12, 2025).

Project Reflection

Task 3: Project Reflection

Skill Development

This project gives me the opportunity to develop several valuable skills. I will learn to train and evaluate AI Builder prediction models, design workflows using Power Automate and build dashboards in Power BI - skills increasingly in demand with the rise of low-code and AI tools. [1] I will also develop research skills through the use of search strategies, database evaluation and critical review of sources. Additionally, I will strengthen my ability in user-centred design and agile development as I iterate based on user feedback.

These skills will be developed mainly through practical application. I will use Microsoft Learn materials to explore Power Platform tools while also learning from academic sources on AI and data quality. [2] The pilot implementation offers hands-on experience in model training, flow building and stakeholder engagement. I will also develop iterative testing skills by adjusting the system based on feedback.

The project supports my professional growth, especially as I currently work with SharePoint and Power BI. Learning to enhance data quality with AI will improve both my current role and future opportunities in data governance or automation. Personally, it also allows me to explore ethical technology use, responsible innovation and the impact of emerging tools on everyday workflows.

Impact on Society

This project could positively impact organisations and wider society by improving data quality and enabling more accurate reporting. Inaccurate data often leads to poor decisions, wasted time or loss of trust in reporting systems. [3] By automating anomaly detection and reducing manual data-cleaning, the solution improves efficiency, accuracy and accountability. In the long term, better data quality can lead to smarter resource allocation and service delivery, whether in public sector organisations, healthcare or education. The project also shows how AI can be applied ethically to assist rather than replace human decisions.

There are a few risks and challenges that could affect this project. One is AI misclassification - false positives or missed anomalies - which could reduce trust in the tool. Another risk is licensing limitations within Microsoft AI Builder, which could limit usage if scaling is required. [4] There's also a chance that end users may resist the change, either due to a lack of familiarity with AI tools or a lack of confidence in automated suggestions. Lastly, even with non-sensitive data, ethical concerns about data use and transparency should be considered.

To mitigate these, validation will allow users to override AI suggestions. Retraining the model and tuning thresholds based on feedback will improve accuracy. Licensing issues will be managed by staying within available Microsoft 365 limits during the pilot. Clear communication, training and ethical guidelines, including GDPR compliance and anonymised feedback, will ensure responsible deployment.

Addressing these challenges will also improve my learning, helping me build skills in adaptability, stakeholder engagement and ethical design.

Conclusion

This project has shown how AI and automation can tackle real-world problems - in this case, improving data quality in SharePoint Lists. It has helped me develop technical and research skills while also thinking about ethical, organisational and practical challenges. Using AI Builder, Power Automate and Power BI, the solution demonstrates how low-code tools can support innovation without requiring advanced programming knowledge. It also highlights the value of user feedback and iterative design in creating technology that is both effective and usable.

Working on this project has helped me grow both personally and professionally. I now feel more confident in

using AI tools to solve practical problems in a work setting.

References

[1]

N. Damoulakis, "Orases," *Orases*, May 08, 2025. <https://orases.com/blog/the-top-6-ai-data-automation-tools-2025/> (accessed Aug. 12, 2025).

[2]

Microsoft Learn, "Training," *learn.microsoft.com*, 2024. <https://learn.microsoft.com/en-us/training/>

[3]

K. D. Foote, "The impact of poor data quality (and how to fix it)," *Dataiversity*, Mar. 01, 2024. <https://www.dataiversity.net/the-impact-of-poor-data-quality-and-how-to-fix-it/>

[4]

Antoine2F, "AI Builder licensing and credit management," *Microsoft.com*, Nov. 02, 2024. <https://learn.microsoft.com/en-us/ai-builder/credit-management>

Appendix

Unit 1 Topic 1

According to the Cambridge Dictionary, “innovation is a new idea or method, or the use of new ideas and methods.”

5 things that were not around 10 years ago:

- Chat GPT
- Remote Working being normalised
- Google Assistant
- Google Wallet
- Food Delivery Apps

These have all changed my life for the better in terms of making day-to-day things a little easier but none of them have made a major difference.

I would say that the invention of Agriculture has had the biggest impact on society as we know it today. This is because without agriculture, humans would still be spread out in tribes, unable to settle down in one place. Settling down enabled many other advancements but also allowed population numbers to rise.

In general, innovation has changed things for the better. Deaths from natural disasters has halved over the last century while the percentage of people living in extreme poverty has also halved. People are living longer and have more time and money to enjoy their lives.

Collective wellbeing should be our goal and we can use science and reason to achieve it.

Stephen Pinker wrote “We live longer, suffer less, learn more, get smarter, and enjoy more small pleasures and rich experiences. Fewer of us are killed, assaulted, enslaved, oppressed, or exploited by the others. Much suffering remains, and tremendous peril. But ideas on how to reduce them have been voiced, and an infinite number of others are yet to be conceived.”

I still believe that Agriculture has had the biggest impact on the development of society. This has had many benefits – civilisation as we know it depends upon this. However, it can be argued that it has forced humans to live in a way that we are not suited to and that we were happier before.

Klaus Schwab talks of the Fourth Industrial Revolution, where the previous three were machines, electricity and computers. The Fourth will be cyber-physical systems. Examples include genome editing, machine intelligence, blockchain and AI.

Of the technologies listed by the WEF, I think the one that will have the biggest impact in my discipline area in the next 5 years will be Computing Capabilities. This is because not only is an enabler of all other technologies but it also enables the acceleration of AI and Machine Learning which I think will have a big impact over the next few years on not only my discipline but the majority of disciplines.

Nikolai Kondratieff studied the impact of innovation on society and produced a chart referred to as Kondratieff Waves or Cycles. Certain Innovations cause spikes of prosperity in society before returning to normal.

Creativity is important as it is innovation, not replication, that will give us a competitive edge. There is no innovation without creativity. Creativity is essential for solving the social and environmental issues in the world around us which will only get worse as it increases in complexity.

We have an innate ability to be creative, but we lose that as we get older. Picasso said that all children are born artists, the problem is to remain an artist as we grow up. We need to be aware that it is wrong to believe that mistakes are the worst things you can do. The creative process is born from trying things that may at first seem wrong, but later appear right.

I do not consider myself to be a creative person under most circumstances as I struggle to think outside the box. However, I think I am actually quite good at thinking up solutions to clearly defined problems.

I had an experience where I felt encouraged to be creative. During a brain-storming session in work, the three of us were encouraged to come up with ways to solve a problem that we had, no matter how outside the box it was. As I was fairly new to my role, I was mostly a bystander but watching the other two throw up ideas without reservation for looking silly and having most of them discarded for legitimate reasons encouraged me to do likewise in future, knowing that I would not be made to feel stupid for suggesting something unworkable or unusual.

Creativity is important in my subject area as we regularly have situations where all we have is a goal and how we get there is up to us. For most tasks, there are many possible ways to go so having some creativity to explore all options for benefits and drawbacks is vital.



✓ **You have completed Section 3!**

After completing this section, you should

- Think of yourself as someone capable of being creative (if you don't already)
- Identify personal passions as sources for idea generation.

Back

Continue →

Unit 1 Topic 2

How do YOU solve a problem?

The importance of creativity

The last time I felt I was being creative was writing for this assignment. I feel at my most creative when putting thoughts into words in a way that other people can understand. It makes me feel like I'm creating order out of chaos, clarity out of obfuscation. It feels satisfying.

Being creative is good for your mental wellbeing

In a study by BBC Arts and UCL, 50,000 people participated in a creativity test. The research showed that there are three main ways we use creativity as coping mechanisms to control our emotions:

- Distraction – using creativity to avoid stress
- Contemplation – using creativity to allow us to reassess problems and make plans
- Self-development – to face challenges by building up self-esteem and confidence

Nothing beats taking part in live creative activities that involve face-to-face social interaction.

Regardless of skill level

New pursuits is more beneficial than doing one for more than ten years as activities can become less effective at regulating negative emotions

When we face hardships, creative activities are particularly beneficial for our emotions

Lower psychological wellbeing significantly impacts on student learning and experience while good wellbeing can lead to better learning and performance

Research suggests adults who make time for play have greater productivity, more creativity and greater health and happiness

This is why learning games are becoming a popular learning strategy

Creative confidence is having the ability to think freely, without fear of shame, or fear of judgement, or any fear.

According to the founders of Ideo, "creativity is something you practice, not just a talent you're born with."

There are four fears that hold most of us back from being creative:

- Fear of the messy unknown
- Fear of being judged
- Fear of the first step
- Fear of losing control

Recognising how your philosophy impacts your investigation

Research philosophy is your belief about the ways in which you will be collecting, analysing and using data in your project

Philosophy of research is split into three broad forms:

- Ontology – the study of being – researcher assumes the knowledge on what exists has already been created, but then asks what the world must be like with this assumption. Split into two branches – objective (reality is controlling the person) and subjective (person controls reality)
- Epistemology – the study of knowledge – split into three – foundationalism (all knowledge is based on a few statements that cannot be doubted), coherentism (all knowledge is true so long as it does not contradict itself) and pragmatism (if it works, it's true and will get you by for the time being)
- Axiology – the researcher is trying to find what is valuable and why – split into two – ethics (moral principles and dilemmas) and aesthetics (sensory emotional value in science, critical reflection on art, culture and nature in social science and branches of art in art)

Each approach has pros and cons

Positivism is led by science. Assumption that there is only one truth that can be gained by observing, collecting data and interpreting strictly

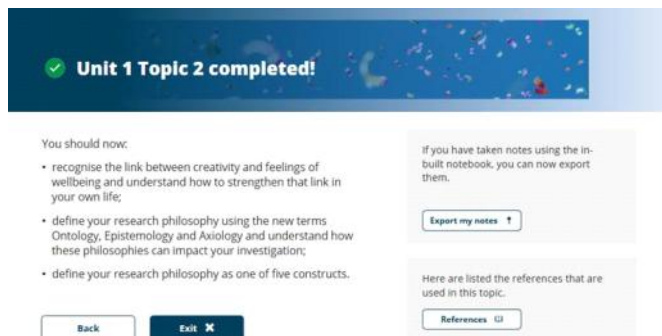
Critical realism is similar but does not require the restrictions of science to be proven.

Interpretivism assumes that access to reality is only through social constructions. You cannot use science to understand people

Postmodernism establishes established thinking by exposing ideas embedded in the way we use language to explain the reality in which we exist

Pragmatism only considers concepts relevant if they support an action

I would say that I feel most drawn to Positivism as I work with data a lot and like to think that there is only one truth that the data leads towards. With enough data, anything can be understood



The screenshot shows a digital interface with a blue header bar at the top containing a green checkmark icon and the text "Unit 1 Topic 2 completed!". Below the header, the page is divided into two main columns. The left column, titled "You should now:", contains a bulleted list of three tasks: recognizing the link between creativity and feelings of wellbeing, defining research philosophy using new terms, and defining research philosophy as one of five constructs. At the bottom of this column are "Back" and "Exit" buttons. The right column contains two sections: "If you have taken notes using the in-built notebook, you can now export them." with an "Export my notes" button, and "Here are listed the references that are used in this topic." with a "References" button.

Unit 1 Topic 3

Known Knowns

Selecting an appropriate database

Four categories, known knowns, known unknowns, unknown knowns and unknown unknowns

Unknown unknown is something we are neither aware of nor completely understand

In research, it is prudent to be aware of what is known and what is not known

This can be explored using library databases

A database is any digital resource of information

Most published work uses between 4 and 8 databases

The CRAAP test evaluates the credibility and relevance of information found:

- Currency – is it recent & has it been superseded?
- Relevance
- Authority – is it credible?
- Accuracy
- Purpose – what is it written for?

Search engines have no quality control and is disorganised

Google scholar can be used

Using Boolean search strategies for identifying research

Keywords are important words that can be used to target our search

Speech marks indicate multiple words are one search term

Asterisk indicates the search should be for all forms of the word

Booleans (And, Not, Or) can be used to refine the search

Resources found can improve future searches

Analysis paralysis can be dispelled by using a concept map

A concept map is a chart that is a visual approach to organisation, like a mind map

A concept map starts with more discipline from the top of the page and works down

It introduces structure and form

Done to expose concepts and propositions

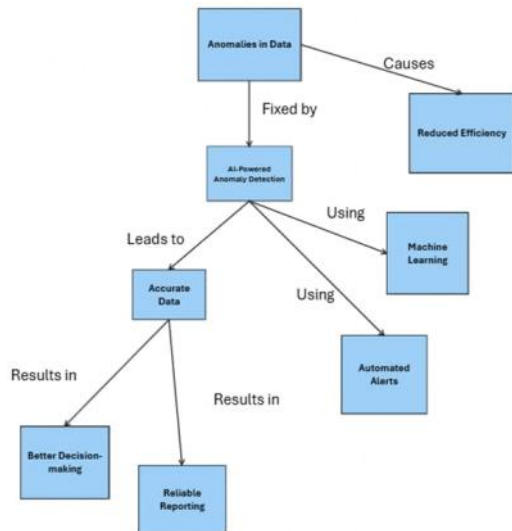
A proposition is an academic term used to describe how two concepts are joined

Can use Microsoft Whiteboard, PowerPoint, Word, Google Docs, Slides, etc

As simple a diagram as possible that organises thoughts

One page and three units max

Concepts in shapes, relationships are arrows and text on arrows explains the relationship



A focus question to outline what is going to be in or out of scope for my research will be ""How can AI-driven anomaly detection improve the accuracy and reliability of data in SharePoint Lists used for reporting?"

Unit 4 Topic 1 completed!

You should now:

- Be able to imply a numerical precision to arguments in a casual way that brings a lack of precision to our work.
- Be able to analyse data using different examples of graphs and charts and present data using a variety of methods.
- Have a deeper understanding of creating charts that communicate in the best way possible.

If you have taken notes using the in-built notebook, you can now export them.

Export my notes

Here are listed the references that are used in this topic.

References

Back

Exit

You have completed Section 3!

Upon completion of this sub-section, you should be able to create your own concept map and come up with concepts that you can use for your literature search.

Back

Continue

Unit 2 Topic 2

Surveys are a great way of collecting data. Can be Quantitative or Qualitative.

Individual Interviews can extract beliefs, values and attitudes.

Focus Groups involve around 7 people discussing a problem. This allows for a range of opinions.

Observation can be done qualitatively or quantitatively.

Secondary data can be used – from databases or websites.

The Hawthorne Effect shows that people perform better while under observation. This can bias the data collected.

For my project, the best way of collecting data would be to do observations on the SharePoint List to get Qualitative data about which errors are most likely to occur. This can then be paired with either surveys or individual interviews to assess the quality of the dataset before and after the solution is implemented. I do not believe that the Hawthorne Effect will be relevant here.

Quantitative data can be limited due to time and resources, limited results or limited replies. This needs to be accounted for.

Picking people for surveys has a number of approaches. However, as only a very limited number of people will be using the dataset, they can all be surveyed or interviewed.

Causal Relationship is an experiment – broadly manipulating a variable and measuring its effect.

For my research project, the data collection will be:

- Type – Causal Relationship
- Sample Size – All users of Dataset
- Method – Classification of errors followed by either surveying or interviewing users after solution has been in place for some time

Unit 2 Topic 2 completed!

You should now:

- Be familiar with different methods for collecting and analysing qualitative data including data sampling.
- Be able to consider how the acquisition of your data will impact your research.
- Understand your approach to collecting data for your research.

If you have taken notes using the in-built notebook, you can now export them.

[Export my notes](#)

Here are listed the references that are used in this topic.

[References](#)

[Back](#) [Exit](#)

Unit 4 Topic 1

Statistics are ways we can try to reveal patterns and trends.

When searching for “Quantitative Research of...” in Google, I get results for education, medicine, stem students and business.

With research statistics, the more accurate, the better.

There are two sorts of statistics – those used to describe our argument (Descriptive) and those used to make inferences from our argument (Inferential).

Descriptive – to turn data into a source of information to help describe the research topic. Can be numbers, charts, tables or graphs. Great way to take a snapshot of a data set and produce a summary that can be more easily consumed.

These statistics could include averages or frequencies.

Inferential – to help us conclude what the data is saying. Inference is necessary as it is usually not feasible to collect information from an entire sample.

Information gathered from these statistics can be strongest variable or proof of an opinion.

In my research, descriptive statistics could include a table showing types of error and their frequency in the dataset. An example question could be: Does error flagging reduce the error frequency and improve overall data quality?

With inferential samples, it should be random and unbiased.

Visualising data in charts is useful at helping us understand it. Different types of chart are useful for displaying different data.

To describe the correlation between academic success to student reading age results, I would use a scatter graph. This is because there are many different combinations possible, with academic success being continuous.

In my field of study, we use a lot of bar charts as it is useful at showing how many people have carried out specific training.

Aristotle said “The aim of art is to represent not the outward appearance of things, but their inward significance.”

I don’t agree with this – art does not have an aim. Artists create what they create for an infinite number of different reasons.

When creating charts, it is important to consider how the viewer will look at it.

Visual order of importance:

Size – Bigger catches the eye more and size hierarchy can show which order things should be read

Colour – Brighter colours catch the eye more

Contrast – High contrast catches the eye more

Proximity – The closer things are to each other, the more they seem related

Alignment – Important text should be more towards the left

When including charts, consider:

What are you trying to accomplish?

What story in your data do you want to share?

Are you trying to present too much?

Less is more

Unit 4 Topic 2

Qualitative research methods involve collecting and analysing data, normally non-numerical. Objective is often to understand stories, experiences and concepts of participants.

A researcher might use qualitative research methods as the topic of research is not numerical. Qualitative methods do not suit all research.

Limitations are that it is difficult to put the results into a tangible form.

Qualitative research follows four stages:

Data collection

Reading and rereading the data

Interpreting findings

Writing your report

I will organise my data in tables in the cloud.

A theme or pattern is when a group of things have similar properties or act in a predictable way.

Thematic analysis is a method used to identify patterns or themes that emerge from qualitative data. The process contains six steps:

- Familiarisation
- Generating initial codes
- Identifying themes
- Reviewing themes
- Naming themes
- Writing about themes

Excel is a useful tool for analysing qualitative data.

Identifying themes can be done using excel with column headers:

- Source #
- Note #
- Author
- Note
- Theme
- Initial Code



Unit 4 Topic 2 completed!

You should now:

- Have demonstrated knowledge and understanding of how to organise qualitative data ready for analysis.
- Have figured out some techniques to identify patterns and themes in your data.
- Have improved the understanding of the use of Excel as a qualitative analysis software.

[Back](#)

[Exit](#) ✕

If you have taken notes using the in-built notebook, you can now export them.

[Export my notes](#) ↑

Here are listed the references that are used in this topic.

[References](#) ↗