Plan Overview

A Data Management Plan created using DMPonline

Title: Inspiring Customers in Retail Malls Using Generative Artificial Intelligence: A Theory-driven Design Science Approach

Creator: Xiangru Liu

Principal Investigator: Xiangrui Liu

Data Manager: Xiangrui Liu

Affiliation: The University of Sheffield

Template: Postgraduate Research DMP (The University of Sheffield)

Project abstract:

Generative AI’s potential in retail remains underexplored, particularly in physical retail environments where its role in sparking customer inspiration is poorly understood. This study addresses these gaps using a theory-driven Design Science Research approach. Drawing on the information systems (IS) delegation framework and customer inspiration theory, we design and implement a GenAI artifact in a large retail mall. This artifact features reflexive (real-time query-responsive) and anticipatory (content-generating) agent roles. We test it through field experiments via the mall’s digital platform, examining whether the GenAI chatbot’s reflexive guidance and AI-driven personalized content elevate customers’ “inspired-by” and “inspired-to” states relative to control conditions. Theoretically, the study extends customer inspiration theory by introducing GenAI as a novel inspiration stimulus, and refines the IS delegation framework by empirically demonstrating how AI agents assume reflexive and anticipatory roles in service delivery. Practically, it provides evidence-based guidelines for deploying GenAI in retail malls, illustrating how such innovations enhance shopper engagement and sales performance.

ID: 173960

Start date: 02-09-2024

End date: 30-09-2027

Last modified: 17-06-2025

Inspiring Customers in Retail Malls Using Generative Artificial Intelligence: A Theory-driven Design Science Approach

Defining your data

- What digital data (and physical data if applicable) will you collect or create during the project?
- How will the data be collected or created, and over what time period?
- What formats will your digital data be in? (E.g. .docx, .txt, .jpeg)
- Approximately how much digital data (in GB, MB, etc) will be generated during the project?
- Are you using pre-existing datasets? Give details if possible, including conditions of use.

We will collect digital data comprising customer interaction logs (queries, clicks, and browsing history), survey responses measuring customer inspiration, and transaction records detailing purchase incidence and amounts. Physical data collection is not applicable to this project. Data will be collected through Shin Kong Place Chongqing's WeChat mini-program during two structured field experiments. Experiment 1 (reflexive GenAI chatbot interactions) and Experiment 2 (anticipatory GenAI-generated personalized shopping guides) will take place over approximately 3 months. Data will be primarily collected in structured formats (.csv, .xlsx) for customer logs, survey responses, and purchase transactions. Additional textual data (chatbot interactions) will be stored in plain text (.txt). We anticipate generating approximately 0.5-1 GB of digital data, comprising mostly textual data from interactions and structured numerical survey and transactional records. Pre-existing CRM datasets provided by Shin Kong Place Chongqing, including historical customer demographics and purchase behaviors, will be used. Access to these datasets is governed by a confidentiality agreement ensuring compliance with data protection regulations and usage strictly limited to this research project.

Looking after data during your research

- Where will you store digital data during the project to ensure it is secure and backed up regularly? (University research storage)
- How will you name and organise your data files? (An example filename can help to illustrate this)
- If you collect or create physical data, where will you store these securely?
- How will you make data easier to understand and use? (E.g. include file structure and methodology in a README file)
- Will you use extra security precautions for any of your digital or physical data? (E.g. for sensitive and/or personal data)

All digital data collected during the project will be securely stored and regularly backed up on the University's dedicated research storage system, which ensures secure access control and regular data backup protocols. Data files will follow a structured naming convention including the project title, data type, and date of creation. For example: GenAI_CustomerSurvey_YYYYMMDD.csv, GenAI_ChatbotLogs_YYYYMMDD.txt, GenAI_TransactionRecords_YYYYMMDD.xlsx. A comprehensive file will be created, clearly documenting the data structure, file naming conventions, methodology used for data collection, experiment descriptions, and data analysis procedures. This file will serve as an accessible guide to facilitate data understanding and future reuse. Given the sensitivity of personal customer data, additional security measures will be employed. These will include restricted access

control, encryption for storage and transmission of sensitive data, and adherence to strict university and industry standards for data privacy and confidentiality.

Storing data after your research

- Which parts of your data will be stored on a long-term basis after the end of the project?
- Where will the data be stored after the project? (E.g. University of Sheffield repository ORDA, or a subject-specific repository)
- How long will the data be stored for? (E.g. standard TUoS retention period of minimum 10 years after the project)
- Who will place the data in a repository or other long-term storage? (E.g. you, or your supervisor)
- If you plan to use long-term data storage other than a repository, who will be responsible for the data?

Anonymized and aggregated datasets, including summarized customer interaction data, anonymized survey results on customer inspiration, and summarized transaction data (e.g., aggregated purchase trends), will be stored long-term. No personal or sensitive data will be stored beyond the necessary retention period. Data will be stored securely in the University of Sheffield's repository (ORDA), which provides secure, sustainable, and long-term access compliant with research standards and institutional policies. Data will be stored in a minimum of 10 years after project completion. The primary researcher will be responsible for placing data in the University repository, supervised and reviewed by my academic supervisor to ensure adherence to the university's data storage guidelines. Moreover, we do not plan to use alternative long-term storage solutions outside the University repository. All long-term data management responsibilities will remain with the researcher under the oversight of the academic supervisor.

Sharing data after your research

- How will you make data available outside of the research group after the project? (E.g. openly available through a repository, or on request through your department)
- Will you make all of your data available, or are there reasons you can't do this? (E.g. personal data, commercial or legal restrictions, very large datasets)
- If there are reasons you can't share all of your data, how might you make as much of it available as possible? (E.g. anonymisation, participant consent, sharing analysed data only)
- How will you make your data as widely accessible as possible? (E.g. include a data availability statement in publications, ensure published data has a DOI)
- What licence will you apply to your data to say how it can be reused and shared? (E.g. one of the Creative Commons licences)

Not all data can be openly shared due to confidentiality agreements, commercial sensitivities, and data protection regulations involving personal customer data. Only anonymized and aggregated datasets without identifiable personal information will be made publicly accessible. To maximize data availability, data will undergo rigorous anonymization processes to remove personally identifiable information. Only aggregated summaries and statistical analyses, free from individual-level sensitive

information, will be shared. We will include a clear data availability statement in all related publications and reports, providing details on how to access the datasets. All data published in the ORDA repository will be assigned a Digital Object Identifier (DOI) to enhance discoverability and accessibility. The data will be made available under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, allowing users to reuse, adapt, and redistribute the data, provided appropriate credit is given to the original source.

Putting your plan into practice

- Who is responsible for making sure your data management plan is followed? (E.g. you with the support of your supervisor)
- How often will your data management plan be reviewed and updated? (E.g. yearly and if the project changes)
- Are there any actions you need to take in order to put your data management plan into practice? (E.g. requesting University research storage via your supervisor.)

The primary researcher will be primarily responsible for ensuring adherence to the data management plan, with regular support, review, and guidance from my academic supervisor. The data management plan will be reviewed annually, or more frequently if there are significant changes to the project methodology, data collection procedures, or institutional requirements. It will also be updated whenever necessary to reflect any adjustments in project scope or data handling requirements. Key initial actions include requesting access to the University research storage facilities through my supervisor, setting up appropriate data backup protocols, and securing approval for data anonymization methods. Additionally, regular meetings with my supervisor will be scheduled to ensure continuous oversight of the data management activities.