Plan Overview

A Data Management Plan created using DMPonline

Title: Leveraging Generative Artificial Intelligence for Personalized Customer Acquisition: A Study on Marketing Strategies in the AI Age

Creator: Mike Brachten

Affiliation: Utrecht University

Template: Utrecht University DMP

Project abstract:

The proposed study aims to evaluate the effectiveness of Generative Artificial Intelligence (GenAI) in personalizing marketing emails to enhance customer acquisition for tech startups. Specifically, it investigates how AI-personalized emails impact user engagement and conversion rates compared to generic emails. Utilizing large language models (LLMs) such as GPT-4o, GPT-3.5, Gemini 1.5 Pro, and Llama 3, the study will generate and send personalized emails to a diverse group of participants primarily recruited through LinkedIn and a partner organization's CRM system.

The study employs a quantitative, between-subjects, randomized design, with a control group receiving non-personalized emails and multiple treatment groups receiving AI-personalized emails. Data on email opens, clicks, and replies will be collected and analyzed to determine the effectiveness of the personalized emails. Measures to ensure participant privacy and data protection include pseudonymization, the use of a Data Processing Agreement (DPA) with OpenAI, and the implementation of open-source LLMs.

This research addresses a significant gap in empirical studies on the real-world impact of AI in marketing, providing valuable insights into the practical applications and potential benefits of AI-driven personalization. The findings are expected to enhance theoretical understanding, inform marketing strategies, and contribute to the development of ethical guidelines for AI use in marketing. The study's minimal burden on participants and its substantial potential contributions to both scientific knowledge and practical marketing practices underscore its importance and relevance.

ID: 154496

Start date: 01-05-2024

End date: 30-04-2025

Last modified: 21-06-2024

Leveraging Generative Artificial Intelligence for Personalized Customer Acquisition: A Study on Marketing Strategies in the AI Age

Data Collection

1.1 Will you re-use existing data ?

If yes: explain which existing data you will re-use and under which terms of use.

- No, I will be collecting/generating new data

1.2 Describe your data.
Fill the table below with a brief description of the data, including the type, format and volume.

#	Data Description	Data Type	Format	Total Volume
1	Identifying information (Name, email, URLs, ID*)	Tabular	.csv	2MB
2	User profiles (ID*, Gender, Age, Location, Job title, Organization, Personal interests/descriptions)	Tabular	.csv	5MB
3	Study outcomes (ID, Treatment (LLM used), Email opens, clicks, replies, conversions)	Tabular	.csv	1MB
4	Email content	Tabular	.csv	10MB
5	Analysis dataset (ID, Treatment, Email opens, clicks, replies, conversions, gender, age, location)	Tabular	.spss	5MB

Data Documentation

2.1 Describe the documentation and metadata that you will use to to make your data reproducible and interoperable. Describe which files you will provide, along with a brief description of the information they will contain, to make your data reproducible and interoperable. Describe the information that you will provide to make the data items in questions 2.1 reusable and interoperable. If using a specific metadata standard, please mention this below.

If I have to, I plan on publishing only the Analysis dataset, which contains the pseudonymization ID, Treatment, Email opens, clicks, replies, conversions, gender, age, location.
I presume the metadata that has to be published is the coding of the treatment variable; e.g. "0" corresponds to control, "1" is GPT-4o, "2" is GPT-3.5, etc.

2.2 Describe the folder structure you will provide to make your data reproducible and interoperable.
Describe the folder structure, naming conventions and/or version control you will use for this project.

>Project Folder
  >>Thesis information
    >>>Final_Thesis.pdf
  >>Dataset
    >>>README.txt
    >>>Data_V#.spss

Data Storage

3.1. Select the storage solution where you will store and back-up your data.
Select the locations where your data will be stored. You may select more than one. Please describe the storage solution and the backup strategy of your storage solution if it does not appear in the list below.

- Other (please specify below)

The storage location will depend on the type of the data:

1. Identifying information: This data will be stored on my personal device*.
2. User profiles: This data will be stored on my personal device*.
3. Study outcomes: This data will be stored on my personal device* and on the server running the email tracking software**.
4. Email content: This data will be stored on my personal device* and it is ineveitable that these are also stored on the SMTP server of the sending server.
5. Analysis dataset: This data will be stored on my personal device*.

* My personal device is a MacBook Air M1 2020 with a FireVault-encrypted drive with a strong password. All data on my personal device is backed up to my personal STACK storage with 2FA (TransIP, Privacy information)
** The server is either a self-hosted VPS or webhost provided webspace running PHP scripts that track email opens and replies for an identifier.

Data Privacy and Security

4.1 Will you be collecting or using personal data ?

Personal data is any data which, alone or in combination with other information, can identify a living person. Such data must abide by the GDPR and requires additional safeguards and documentation to be processed lawfully.

- Yes, I will collect and/or use personal data

Please refer to the GDPR compliance statement and DPIA in my research proposal for more detailed information. I'm processing personal data (specified in Data Documentation section of this document) that is publically available.

4.2 What is the legal basis by which you are collecting and/or processing this data ?
If you are uncertain as to which legal basis applies to your type of research, please do not hesitate to contact us at info.rdm@uu.nl or by using the "Request feedback" button and leaving a comment alongside this question.

- Public interest

As this research is the first of its kind (empirical experiment testing the effectiveness of GenAI personalization) it provides essential insight that guides marketeers, researchers, the scientific community, and ethical codes of conduct. Outcomes can inform our knowledge of the potential of AI to be persuasive, which is in the public interest to inform consecutive decisions about how we manage this new technology safely.

4.3 Select the privacy and security measures you will employ to protect the privacy of your data subjects. Check all that apply.

- Access control
- Encryption
- Minimization
- Pseudonymization
- Secure storage

4.4 Who is the controller of the personal data ?
The controller of the personal data is the entity which determines what is done with the data. In most cases the controller is Utrecht University.

Utrecht university is the controller of the collected personal data. Nevertheless, the principal investigator of the research project will ensure that the data is handled and processed in accordance with the GDPR.

4.5 How will ownership and intellectual property rights of the data be managed?
Describe who controls access to the data and who determines what is done to the data.

The principal investigator (MSc student, Mike Brachten) will determine who has access to the data within the research group. All intellectual property rights belong to Utrecht University.

During the project, the MSc student has access to the data. Upon request, the supervisor can access the data for verification purposes of the thesis grading.
The research group only consists of the principal investigator, first supervisor, daily examiner, and possibly a partner organization providing the data.

Data Selection, Preservation & Sharing

5.1 Describe the data you will be preserving and the storage solution where it will be preserved?
Describe which data will be preserved under long-term storage. You may refer back to the data described in question 1.2 to specify which data will be preserved. Explain where you will preserve your data, and how procedures are applied to ensure the survival of the data for the long term.

For knowledge sharing purposes, only the pseudonymized dataset can be archived. The other data is deleted after the thesis is completed, with exception of directly identifiable information, which is deleted after four weeks after sending the follow-up email (so respondents can opt-out).

5.2 Describe the data you will be sharing and the repository where it will be shared?
Describe which data you will be sharing. Select where you will make your data findable and available to others. If selecting "Other" please specify below which repository and provide a URL.

Please also write below if you will apply any conditions to the re-use of your data. (i.e. Creative commons license or Data Transfer Agreement).

- Other

I don't have a respository yet, I will inform this decision with UU-recommended systems.

5.3 Are specialized, uncommon or expensive software, tools or facilities required to use the data?

Please list any specialized, uncommon or expensive software, tools or facilities that are absolutely required to obtain, use or handle your data, if any.

SPSS is likely required, but the data can also be exported to .csv for portability.

Data Management Costs and Resources

6.1 What are the foreseeable research data management costs and how do you expect to cover them ?
Please specify the known and expected costs involved in managing, storing and sharing your data. Also explain how

you plan to cover these costs.

None, only for storage and archiving if desired by the University, but I assume UU will pay for this.

6.2 Who will be responsible for data management?
Please specify who is responsible for updating the DMP and ensuring it is being followed accordingly.

The MSc student is responsible for data management.

6.3 State if you contacted an RDM consultant from Utrecht University to help you fill out your DMP.

Please list their name and date of contact.
This is mandatory forNWO grants.

No.