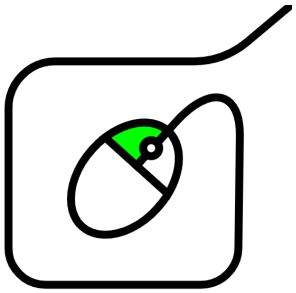

SuperGeeks

Requirements Specification

Version 1.1

November 29, 2021

Project Sponsors: Dr. Marco Gerosa, Dr. Igor Steinmacher
Team Faculty Mentor: Felicity Escarzaga
Team Members: Tim Giroux, Kyle Austria, Gustavo Valencia, Pengfei Liu

---

Accepted as baseline requirements for the project:

*For The Client:*_____ *Date:*_____

*For The Team:*_____ *Date:*_____

Table of Contents

1. Introduction

Software engineering research is critical to the success of the many industries that rely on quality software products. It provides insights that are highly valuable to users and developers alike. Empirical data about what does and does not work in software engineering is very important in the world today, especially for critical software systems in the medical and structural engineering industries. Research in this field often depends on human participants. Researchers need to coordinate with software professionals in order to track the efficacy of various engineering strategies. Without this coordination, much of this critical research would be too challenging to conduct.

While researcher-participant coordination does take place in software engineering research today, it is difficult to facilitate and limited in scope. Researchers face a number of core problematic business functions:

- *Limited scale*:

  Large-scale research projects could result in important discoveries but are often discarded because there is so much difficulty in obtaining participants.

- *Inefficient recruitment*:

  Recruitment efforts are ineffective and create a large amount of overhead for researchers.

- *Sampling bias*:

  Typically, only a subset of desired participants for a study actually respond to recruitment efforts. This introduces bias because the other desired participants will not be included in the final results of the study.

The current solutions to these problems take much more time and effort than should be necessary. Researchers might send out mass emails to software professionals based on their Github profiles, but this kind of email is often considered as spam by developers. Developers can sometimes be irritated by an email like this, or they may simply ignore it. Of those who respond, many do so only to politely decline.

Dr. Gerosa and Dr. Steinmacher, sponsors of the GeekSurvey capstone project, are researchers in the SICCS Software Engineering Research Laboratory at Northern Arizona University. They produce a large volume of software engineering research each year. This research supports the use of innovative tools, processes, and strategies to increase the quality of software products. The clients' primary means of obtaining participants for their research is mass email efforts. Email is most commonly used because research often depends on a qualified pool of subjects with very specific desired characteristics. The GeekSurvey sponsors report that this is a prominent bottleneck in their research process. They have also reported that the same problem impacts many of their colleagues. After trying existing products like Prolific

and Amazon Mechanical Turk, they found that these products lack the ability to easily filter participants based on their professional qualifications. The clients need a product that will appeal to the needs of both software professionals and researchers.

2. Solution Vision

GeekSurvey aims to be a product that could solve these problems. The plan for GeekSurvey is a public web application where users can create an account and create, manage, or participate in research studies. To facilitate cooperation between researchers and willing participants, GeekSurvey will need the following core features:

- Two account types: researchers and participants. By default, all accounts can access researcher and participant features. These two account types are an abstraction to clarify the purpose of different GeekSurvey features. Accounts can be manually given administrative access.

- Participant accounts will include information about the user such as their occupation, level of education, experience with certain technologies, and a link to their Github account.

- Researchers have the ability to create studies and specify necessary criteria for participation.

- Participants are given access to studies in which they are eligible to participate in. They can receive compensation from researchers for completing the study.

By Spring of 2022, GeekSurvey will be a high quality, extensible web application with all of the core features needed for conducting research studies. To be clear: features such as survey sampling strategies, user messaging, and support for thousands of active users, are **not** included in the requirements. Such features are considered **stretch goals**. They would be great additional features, but are not tractable for completion by Spring of 2022.

This document outlines a **minimum viable product (MVP)**, or a **working prototype**, that current developers are committed to producing by Spring of 2022. This MVP will allow the project clients and other researchers to create, manage, and actually conduct research studies using a GeekSurvey study invitation link. Upon clicking the link, participants will see a landing page for the study and will be required to make a GeekSurvey account to participate. After making an account, they will be able to discover other active studies that they are eligible for on the GeekSurvey platform.

Figure 2.1 - Client business process

This improved business process shown in Figure 2.1 allows the clients to build a network of qualified participants with GeekSurvey accounts. In this sense, the GeekSurvey MVP can be thought of as a simple, self-hosted Prolific alternative. It will encourage the growth of an engaged and personal community of researchers and participants, and all data on the platform will be private to the researchers or administrators running the GeekSurvey instance.

3. Project Requirements

In order to achieve the envisioned solution, developers will have to work towards a set of requirements for the software product. The more specific requirements for GeekSurvey will be informed by higher level key features envisioned from a client perspective. The key features are as follows:

- User accounts.
- Profiles with enrollment criteria.
- Study creation with necessary criteria for participation.
- Study enrollment and completion for eligible accounts.
- A payment system to create incentives for participants to use the platform.

The following subsections will outline the specific requirements that GeekSurvey will need to include by the end of initial development in Spring of 2022. These requirements will be split into three categories: functional, non-functional, and environmental.

3.1 Functional Requirements

GeekSurvey must be capable of the following set of functionalities. These functionalities are verifiable because they are directly related to the user experience of the application. GeekSurvey cannot function as a working prototype without these core features and their underlying requirements.

- **Studies**
  Studies are the primary means of cooperation between researchers and participants on the platform. A study in GeekSurvey is a bundle of data containing everything necessary to conduct a minimum viable research study.

    - *Enrollment Criteria*
      Studies will need to have enrollment criteria to filter for desired participants. This criteria can include general demographic information and software engineering-related qualifications. This will also include a maximum number of participants for the study.

    - *Survey Link*
      Studies will contain a link to a survey created on an external survey platform. This link will be added by the researcher who created the study.

    - *Landing Page*
      GeekSurvey will generate a landing page for a study that will demonstrate the important features of the study. This landing page will include the description and title of the study, as well as the profile information of the researcher who created the study. It will have an enroll button or link. Users who are not logged in to GeekSurvey will be prompted to sign in or create an account.

    - *Description and Title*
      Studies will have a custom description and title that will be displayed to potential participants before enrollment.

    - *Invite Link*
      GeekSurvey will provide a unique link for each study that can be used as an invitation link. This link will lead users to the study's landing page. From there, users can enroll in the study.

    - *Participation Tracking*
      GeekSurvey will need to keep track of enrolled participants to a study. It will also be necessary to record which of these participants completed the survey associated with a study. There will need to be some data-structure associated with each study where this information can be stored and updated.

- *End Date*
  Upon creation, studies can be given a custom end date. GeekSurvey will not allow for participation in the study after the end date has passed.

- *Management*
  After creation, study information can be managed by the study owner through some menu. This will include the ability to manually remove enrolled participants in order to protect from bad results. Participant removal will be reflected in the maximum allowed participants.

- **User Accounts**
  GeekSurvey will have generic user accounts. This includes shared functionality between the two account types: Researchers and Participants.

  - *Create and Update Profile*
    Accounts will be managed with a simple username and password. Users will be encouraged to create a strong password. User accounts and authentication will be managed by Django's core authentication framework. Accounts will be associated with a customizable profile. GeekSurvey must include a user-friendly interface for creating and customizing these accounts.

- **Researcher Accounts**
  The primary purpose of a researcher account is to create and manage studies on the GeekSurvey platform. A researcher account will contain important information about the researcher, such as their name and a customizable account description.

  - *Create Studies*
    GeekSurvey must provide a user interface for study creation. This should allow researchers to customize the description and title. They should also be able to add a survey link.

  - *Manage Studies*
    Researchers must be able to manage their studies. At any given time, a researcher may have multiple active and inactive studies. Researchers should be able to see enrolled participants and their completion status for a given study. They should also be able to edit information and remove the study.

- **Participant Accounts**
  Participant accounts have demographic information to determine their eligibility for studies. If they are eligible for an active study, they can enroll and participate.

- *Discover Studies*
  GeekSurvey must provide a means for participants to discover studies for which they are eligible. This can be accomplished by a curated home page for participants. The home-page can show a list of active studies that the participant is eligible to participate in. Participants can also optionally receive email notifications with suggested studies (based on eligibility).

- *Enroll in Studies*
  Participants must be able to use a user interface to enroll in a study. Once enrolled, they will be added to a list of enrolled participants to the study. A user will be able to see a list of studies that they have enrolled in on their profile page.

- *Take a Study Survey*
  After enrolling, a user will be able to access the study survey link. They can click this link to be redirected to the external survey platform. Upon completing the survey, they will be redirected back to GeekSurvey and their participation will be recorded in the GeekSurvey backend.

- **Payment System**
  GeekSurvey will include a payment system to create incentives for participants to use the platform. This payment system will be developed in a sandbox environment using fake external payment accounts. GeekSurvey will not be required to manage any real funds.

  - *Internal Payments*
    GeekSurvey accounts will have an associated balance representing an amount of funds in the account. This information will be represented in the database. GeekSurvey account balances can be transferred between GeekSurvey accounts.

    - *Compensation for a Study*
      Upon creating a study, researchers can specify a payment amount to be given to each participant who completes the study. Upon completing a study, participants will receive this specified amount to their GeekSurvey account. The amount will be taken from the researcher GeekSurvey account balance. To prevent bad participants from receiving funds, this can only happen with approval from the researcher.

  - *External Payments*
    GeekSurvey must be able to create “real” (sandboxed, but using real payment modes) transactions that occur external to the platform. The account balances in GeekSurvey represent these “real” or external transactions. GeekSurvey will

need to manage external transactions to and from a GeekSurvey external account. This GeekSurvey external account is where “real” user funds will be held.

- *Fund User Accounts*
  When a user funds their account, an external transaction will take place moving funds from the user’s external account to the GeekSurvey external account. Funds will be added to the GeekSurvey external account, and an associated amount will be reflected on the user’s GeekSurvey account balance.

- *Cash-out User Accounts*
  When a user cashes out their account balance, an amount will be transferred from the GeekSurvey external account to the user’s external account. This amount will be based on the user's GeekSurvey account balance, and will be subtracted from this balance to reflect the transaction.

- **Administrative Access**
  User accounts can be given administrative access. Accounts with administrative access will be able to access and manage critical information on GeekSurvey.

  - *Manage Studies*
    Administrators will have full access to all studies on the platform. They will be able to edit or remove studies and view all information related to any given study. This will be possible through a user-interface in the web browser.

  - *Manage User Accounts*
    Administrators will be able to view all user accounts. They will be able to view and edit account information.

3.2 Non-Functional Requirements

The GeekSurvey deliverable in Spring of 2022 will need to meet the following non-functional or performance requirements. These requirements clearly define benchmark qualities that the final product must have.

- **Privacy**
  Researchers must not have access to personal information that is not required for a study. Privacy is a principle for protecting user information that must be respected across all GeekSurvey functionalities.

- **Accessibility**
  GeekSurvey is a web application that must be online for users to access. For the application to function properly, it must be accessible to users without difficulties or disruptions.

  - *Handle simultaneous user sessions*
    GeekSurvey will be managing user sessions whenever a user is logged into the platform. It will serve dynamic, personalized web files to each user. As a benchmark, GeekSurvey must be able to provide this functionality to multiple (at least two) users at the same time.

  - *Limited backend delay*
    Backend operations must be performed efficiently enough so as not to produce more than 5 seconds of delay before sending an HTTP response back to the user.

  - *Works with modern browsers*
    GeekSurvey must work on modern versions of Chrome, Firefox, and Edge web browsers.

- **Usability**
  GeekSurvey must be simple for researchers and participants to use. At no point should users be confused by the web interface.

  - *Mobile friendly frontend*
    The GeekSurvey web interface will include a number of different dashboards for various purposes, such as account creation, study creation, and study discovery. Anything that is possible to do on GeekSurvey using a desktop web browser should also be possible using a smartphone (iPhone 8) web browser. Web interfaces should ideally be aesthetically pleasing to mobile users, but the requirement is that they are at least usable.

3.3 Environmental Requirements

There are a few environmental requirements that are imposed on the GeekSurvey prototype. GeekSurvey has a few external dependencies that it must provide support for. The web application must also comply with modern web standards.

- **Dependencies**
  GeekSurvey will be a modern web application using the latest stable version of Django as of November 2021. It will likely need to use various other python packages. Therefore, GeekSurvey is dependent on certain versions of these technologies. The necessary versions of these technologies will be packaged with GeekSurvey as a single container image. GeekSurvey will also have a notable external dependency that cannot be containerized: a survey platform. Support for Google Forms must be built into the GeekSurvey application code.

  - *Containerization*
    GeekSurvey will be developed for deployment as a Podman container. This means that the project code-base must include a configuration file to define the creation and boot processes for the container.

  - *Survey Platform*
    GeekSurvey will support Google Forms surveys to be included in studies. The application must be able to recognize when a participant user has completed one of these Google Forms surveys.

- **Standards**
  GeekSurvey must comply with modern web standards in order to be a proper web application.

  - *HTTPS*
    GeekSurvey must be able to serve web files over an HTTPS connection and must comply with HTTPS standards. During the development process, GeekSurvey will not have an official domain name, so there is no reason to acquire an SSL certificate. However, it is worth noting that **geeksurvey.net** and **geeksurvey.org** are available domain names as of November 2021. During development, GeekSurvey will be hosted over a secure connection provided by an NAU web server. The web application must return the appropriate HTTP response codes for unauthorized access and undefined URLs.

- **Documentation**
  GeekSurvey must be maintainable and extensible. The necessary information for installation, basic configuration, and maintenance must be made available as part of the

MVP in Spring of 2022. The web application codebase will be well documented to facilitate future development.

4. Potential Risks

In order to prepare for any difficulties that may arise in the process of developing GeekSurvey, it is necessary to outline some potential risks.

- *Redirect failure*
  When a participant user is completing a survey on an external platform, they will be redirected to GeekSurvey with some information to record their completion. It is possible and somewhat likely that the redirect could be unsuccessful due to a bad internet connection. Redirect failure could cause incorrect data in GeekSurvey: the external survey may record the results, but GeekSurvey will still think the user has not completed the survey. This risk can be mitigated by making sure GeekSurvey is constantly online while active studies are taking place.

- *Data loss*
  GeekSurvey will store its critical information in a database. It would be harmful to the platform if this data were lost due to misplacement, hardware problems, or some kind of file corruption. Thus, data loss is somewhat likely to happen. In order to mitigate this risk, GeekSurvey developers can avoid using critical data on the platform until some database backup mechanism is implemented.

- *Spam*
  Spam is a problem for any public web application, so it is very likely to happen. This can ruin the platform, as real human users may be outnumbered by bots. Spam can have especially bad consequences for GeekSurvey because the utility of the platform depends on quality, human participants. It can be mitigated by requiring a CAPTCHA for account creation. It may also be possible to offload the problem by managing account creation through OAuth2 from another platform such as Google or GitHub.

- *Cyber attack*
  Account profiles on GeekSurvey will contain some private information about the users. It is important that this information remains secure and that authorized administrators retain control over the platform. It's very likely that a cyber attack against GeekSurvey will take place, since it's a public web application. Therefore, it is critical for GeekSurvey to take preventative measures against cyber attacks. Cyber attacks against any public web form can be almost guaranteed, since many are performed automatically. This can be mitigated by programmatically checking for potentially malicious user inputs and never putting user input directly into a database query.

5. Project Plan

A plan for development will be necessary to ensure the completion of core requirements by Spring of 2022. Developers will work on core features grouped into roughly two-week sprints between December of 2021 and May of 2022. Features will be grouped into the following sprints and will be developed in the order in which they appear:

- Home Page
- User Account Creation
- Study Creation
- Study Discovery
- Study Completion
- Admin Menu
- External Payments

According to this development plan, all necessary work towards the final GeekSurvey deliverable will be completed sometime in March of 2022. This leaves additional time as a precautionary measure in case development takes longer than expected. Additional time may also be used for improving the core product or implementing stretch goal features. Figure 5.1 visually demonstrates how the development effort will be organized to accomplish the envisioned solution.

GeekSurvey																	
SuperGeeks	Project Start:	12/13/2021	**12/13**	**12/20**	**12/27**	**1/3**	**1/10**	**1/17**	**1/24**	**1/31**	**2/7**	**2/14**	**2/21**	**2/28**	**3/7**	**3/14**	**3/21**
TASK	**START**	**END**															
1. Home Page	12/13/2021	12/24															
2. User Account Creation	1/3/2022	1/18/2022															
3. Study Creation	1/18/2022	2/3/2022															
4. Study Discovery	2/1/2022	2/16/2022															
5. Study Completion	1/18/2022	2/3/2022															
6. Admin Menu	2/3/2022	2/14/2022															
7. External Payments	2/21/2022	3/7/2022															

Figure 5.1 - SuperGeeks Gantt Chart

6. Conclusion

GeekSurvey will be a web application to solve the research coordination problem identified by Dr. Gerosa and Dr. Steinmacher. It will be a single containerized virtual machine which runs both the web framework and the database. During the process of development, the container will be deployed to an NAUITS server. At the end of initial development in Spring of 2022, the container will be delivered to the project sponsors along with documentation for maintenance, further development, and deployment to other platforms. The project sponsors will also have the option to inherit the NAUITS Podman account used during development. This would allow the sponsors to use the online web application without manually deploying to another hosting service and paying the associated fees.

In order to ensure the completion of the core features, developers will adhere to the plan for development shown in Figure 5.1. The envisioned working prototype is to be completed in Spring of 2022. As of December 2021, some significant progress has already been made. A base foundation for GeekSurvey is already deployed to a development server through NAIUTS. By the end of initial development, GeekSurvey will be freely and publicly available, open source, easily extensible, and simple to deploy to any containerized hosting platform. It will have the potential to provide a real service to many researchers around the world.

7. Glossary

Prolific: An existing web application that serves as a platform for conducting research studies.

Amazon Mechanical Turk: An existing web application that can be used to conduct research studies**.**

Stretch Goal: A goal for the end product that developers cannot promise but would like to achieve if possible.

Minimum Viable Product (MVP): This is the goal for Spring 2022. A working product that includes the minimum core features.

Working Prototype: This is used interchangeably with **MVP**.

Container: A packaged software product with all of its dependencies.

Docker: A software product for managing containers.

Podman: A software product for managing containers; reverse compatible with Docker.

HTTPS: A secured version of HTTP, the standard protocol for internet communications.

SSL Certificate: A certificate used for authentication in HTTPS.

Spam: Junk information on the internet, often distributed by bots.

CAPTCHA: Completely automated public turing test to tell computers and humans apart.

OAuth2: An industry standard protocol for authentication.