

Hacking Materials User Interface

Team Boxjelly - Project Structure

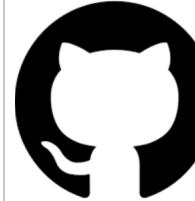


| | |
|---|----|
| 1. - Home | 2 |
| 1.1 Assessments | 4 |
| 1.1.1 Milestones | 5 |
| 1.1.2 Sprint 1 Assessment Checklist | 6 |
| 1.1.3 Sprint 1 - Reflection | 8 |
| 1.2 Specifications | 9 |
| 1.2.1 Project Description | 10 |
| 1.2.2 Motivational Model | 11 |
| 1.2.3 Personas | 14 |
| 1.2.4 User Stories | 22 |
| 1.2.5 Prototype | 25 |
| 1.2.6 Business Case | 27 |
| 1.2.7 Plan | 29 |
| 1.3 Sprint Artefacts | 30 |
| 1.3.1 Sprint Review | 31 |
| 1.3.1.1 Trello Board link | 32 |
| 1.3.2 Retrospectives | 33 |
| 1.3.3 Product planning | 34 |

- Home

Hacking Materials User Interface

COMP90082-HA-2022

| Slack | Trello | GitHub |
|---|---|--|
|  |  |  |

About

The aim of this project is to provide a user interface for engineering for a standard process in [Matminer](#). Retrieving data from databases, user selection of features to be extracted within the databases, performing simple machine-learning tasks (scikit learn, Keras), and visualising results.

In computational materials engineering, engineering new materials is accelerated by avoiding expensive and lengthy experiments to demonstrate materials' performance. Data mining methods discover better materials by searching computer-generated databases with simulations predicting using high-throughput and high-performance computing. Unfortunately, the translation of this new thinking into engineering practice is still in its infancy with some frontrunners (e.g., Tesla). A barrier is materials engineers' software tools. An engineer with materials domain knowledge needs to access and process these data efficiently to make an informed decision for eventual machine learning strategies.

The [Matminer](#) python library provides a framework to simplify the process of data retrieval, feature extraction, machine learning and visualisation. But even a python tool is a high barrier in engineering practice.

Contact Details

Client

| Name | Email |
|---------------------|--|
| Dr Christian Brandl | christian.brandl@unimelb.edu.au |

Staff

| Name | Email | Role |
|------------------|--|---------------------|
| Mauro Mello Jr | mauro.mellojr@unimelb.edu.au | Supervisor |
| Eduardo Oliveira | eduardo.oliveira@unimelb.edu.au | Subject Coordinator |

Team contacts: **BoxJelly** 

| Name | Email | Role | Student ID |
|----------------|--|----------|------------|
| Zhaoqi Wang | zhaoqiw@student.unimelb.edu.au | Team Rep | 1018092 |
| Dara O hEidhin | doheidhin@student.unimelb.edu.au | | 1006587 |

| | | | |
|-----------------------|--|--|---------|
| Felipe Leefu Huang Li | fleefuhuangl@student.unimelb.edu.au | | 1202652 |
| Radhimas Djan | djanr@student.unimelb.edu.au | | 1146240 |
| Yaoming Xuan | yaomingx@student.unimelb.edu.au | | 1167302 |

Team contacts: Redback

| Name | Email | Role | Student ID |
|----------------------------|--|----------|------------|
| Alastair Daivis (Team rep) | adaivis@student.unimelb.edu.au | Team Rep | 359 101 |
| Chunbaixue Yang | chunbaixue@student.unimelb.edu.au | | 1208333 |
| Ghina Yashar | gyashar@student.unimelb.edu.au | | 1274878 |
| Mamta Lopes | mlopes@student.unimelb.edu.au | | 1157314 |
| Sanjeevani Avasthi | savasthi@student.unimelb.edu.au | | 1101265 |

Team contacts: Bluering

| Name | Email | Role | Student ID |
|----------------------|--|----------|------------|
| Yanan Liu (Team rep) | yananl7@student.unimelb.edu.au | Team Rep | 1289747 |
| Hongpeil Lu | hongpeil@student.unimelb.edu.au | | 1275238 |
| Jiahao Ju | jiahaoj1@student.unimelb.edu.au | | 1128182 |
| Rui Zhang | rzzhan2@student.unimelb.edu.au | | 1221568 |
| Xinle Yu | xinley@student.unimelb.edu.au | | 1294310 |

Recent space activity



Dara O hEidhin

| Sprint 1 Assessment Checklist updated a minute ago • [view change](#)

| Milestones created 9 minutes ago



Felipe Leefu Huang Lin

| Plan updated 11 minutes ago • [view change](#)

| Specifications updated 12 minutes ago • [view change](#)



Dara O hEidhin

| Assessments created 12 minutes ago

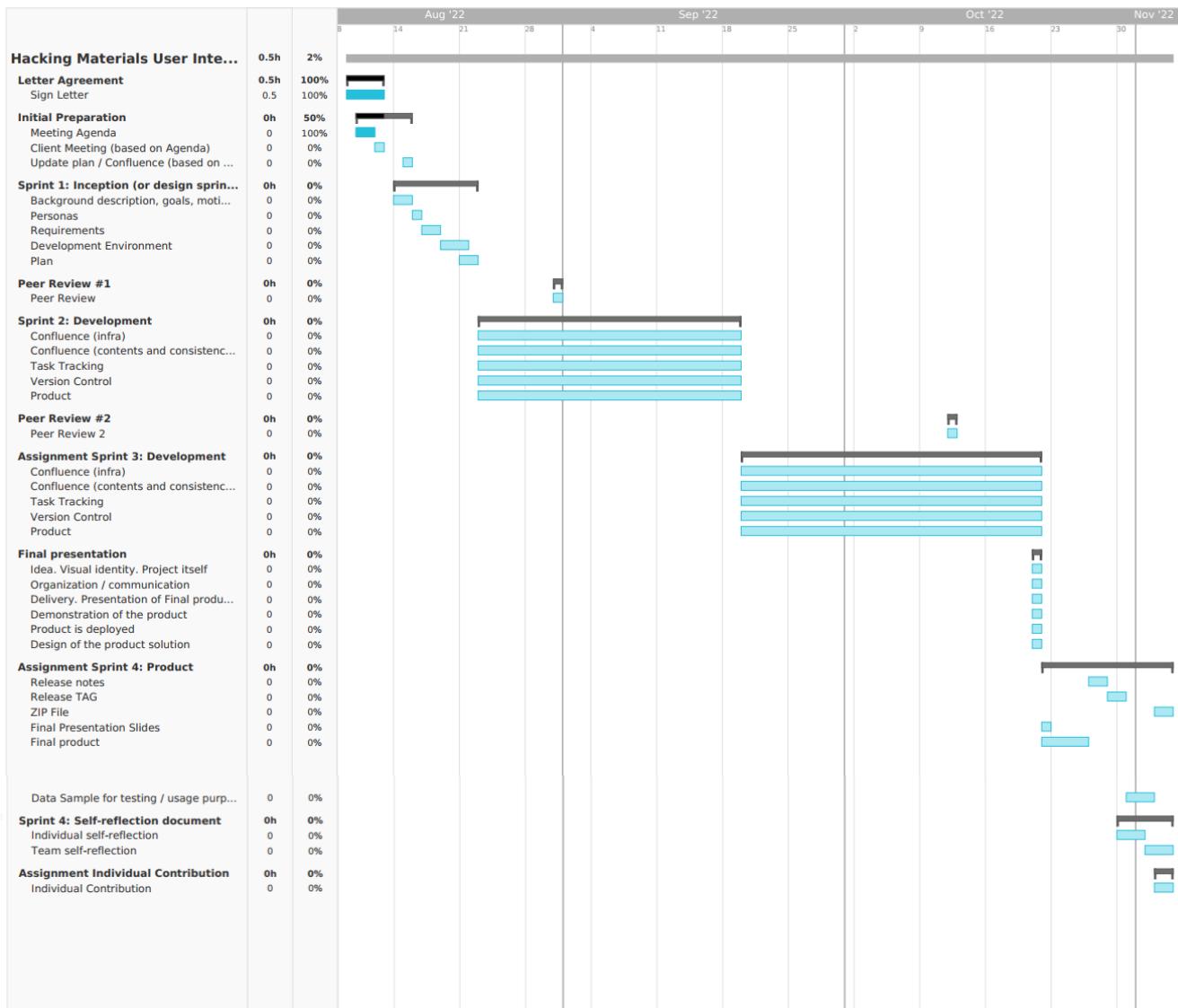
Space contributors

- Dara O hEidhin (a minute ago)
- Felipe Leefu Huang Lin (11 minutes ago)
- Radhimas Djan (50 minutes ago)
- Yaoming Xuan (an hour ago)
- Zhaoqi WANG (2 hours ago)
- ...

| Assessments

- Milestones
- Plan
- - [Sprint 1 - Reflection](#)

Milestones



Sprint 1 Assessment Checklist

Sprint 1 Assessment Checklist

This checklist helps you double check our work for Sprint 1.

Background description, client goals, motivation

- Project overview, background and goals were created.
- DO-BE-FEEL list and GOAL MODEL were created.
- The goal model is consistent with the client understanding of the problem and with DO-BE-FEEL list.

Personas

Make sure that your Personas satisfies the following criteria:

- 2-3 personas were developed to help with requirements validation.
- Personas are based on the research done by students and the discussion with industry partners.
- Personas are inclusive and diverse.

Analysis of requirements (User Stories or Use Cases)

- The analysis of requirements was performed on most of the existing requirements.
- The [new set of] requirements is consistent to the scope of the project, completely cover the new capabilities required by the client and are well documented/structured/organized on Confluence.
- The requirements can be documented in the form of user stories or use cases, supplementary specification of design/implementation/deployment requirements, prototypes, and others. It may also be necessary to be explicit about what is not in scope to define the scope boundary more clearly.

Development environment

- Confluence is organized (cover page, project details, requirements, technical details about the project, meeting minutes and so on).
- Trello (or Github projects or JIRA) is created, structured and organized.
- README file is updated and provide details about the project, workflow (branches/naming conventions and so on).

Plan

- Approved 3-team collaboration plan approved by Eduardo
- A plan (or discussion on what to do next) was provided (requirements to develop, technologies to use, infrastructure to deploy the project) for Sprint 2 and Sprint 3.
- Requirements were estimated and prioritized.
- Backlog items can be found in Trello (or Github project or JIRA).

Meetings

- Meetings are recorded in Confluence and only. They were NOT exported to Github as they're part of internal process.

GitHub

- Folders are structured (On Canvas, visit Assignment -> "Sprint 1: Confluence Space, project background and elicitation documents" page: you can find requirements for folders' structure.)
- Sprint 1 documents were exported from Confluence and added to the repository (and are updated)
- README file is updated and explain the team's repository
- A baseline tag was generated for this Sprint (On Canvas, visit Assignment -> "Sprint 1: Confluence Space, project background and elicitation documents" page: you can find requirements for the baseline tag)

Additional Information

do you have any other additional information you'd like to share with us? Please add it here.

Sprint 1 - Reflection

Sprint retrospective (Reflection):

What went well:

During this sprint, what went well was how the team could adapt and act quickly on deciding how to collaborate within three groups. Throughout sprint 1, we decided to make a weekly meeting. Before the meeting, each group will do their respective task, and then during the meeting, they will work together on the task. By doing this, the team is able to work together and finishes the user stories, motivational model, persona, business case, and QA plan. At the end of the sprint, the group planned to distribute the work for the next sprint by spreading tasks using epic. Each group will be assigned to an epic to make it easier to split up the works. Based on our experience in this sprint, we are confident that in the next sprint, the three groups will be able to work as a team on our tasks respectively and finish them on time.

What did not go well:

- The main thing that did not go well in this sprint was the administrative task on the universities side. Our supervisor suggests us the three groups work together as a team. However, the universities are pretty slow to give access to the tools the team needs, such as GitHub and confluence access to each group. We are not able to work as efficiently due to this. Fortunately, for sprint 1, there hasn't been a development that requires a GitHub, so it isn't affecting the work on the team.
- Finding the common availability between 15 people for the meeting is difficult. We mitigate this by having team representatives on each team and communicating it between the team. These approaches work from what we experience, but it is quite a high effort for the team representative and creates an increased risk of information getting lost in transmission.

What to improve:

The main thing we need to improve on in the next sprint is the meeting with other groups because we found that scheduling meeting between 3 teams is sometimes tricky. Improving it is crucial because, for the next sprint, we planned to distribute two epic task for each group. However, there will be user stories that correlate with each other. Due to this, we will need to ensure that the member working on the user stories is in the meeting if it's associated with other group user stories.

What we think of the client interactions:

We feel that in sprint 1, Dr. Christian Bradl was very cooperative throughout the process. He always responds to our email, explains things clearly on the specification of the task, and explains what features are good to have or could have. He is also cooperative in setting up a meeting while considering the three teams' times.

| Specifications

| Links |
|---------------------|
| Project Description |
| Motivational Model |
| Personas |
| User Stories |
| Prototype |
| Business Case |
| Plan |

| Project Description

Goals

- Help material engineers and people with interest in the field to predict the properties about a new material with simulations to avoid redundant experiments.
- The idea is reasonable but still in fancy in practice. We need to make it happen.

Sponsor

- Dr. Brandl
 - Lecturer and sponsor in UOM studying computational material engineering.
 - Has some knowledge about python and jupyter notebook, but knows very little about programming and algorithm.
 - Very cooperative but doesn't want to be bothered after work hours.
 - Regular meeting: 1:00 PM on Friday.

Users

- Regular user
 - General user with knowledge and has interest in material engineering.
 - Has no or very little experience in python, programming or algorithms.
 - Would benefit from an easy-to-use application to assist in retrieve data and perform applied machine learning functionalities.
- Pro user
 - Regular user with access to additional features:
 - In control of the codes of the ML methods and the datasets.
 - Able to modify the python codes in the interface.
 - Able to introduce new datasets from new sources.
 - Able to add new features and delete existing features (the latter is optional).
 - Account management (not mentioned in the meeting but should have this function).
 - For now, the pro user is Dr. Brandl himself.
- Students
 - Graduate engineering students studying materials.
 - Has no experience in python, programming or algorithms (i.e. knows nothing about ML).
 - Able to search and understand wiki pages.

Structure

- Web page
 - Able to demonstrate diagrams.
 - Able to show more details about data entries in the diagram on click.
 - Provide hints and guidance for new users (low priority).
 - Provide interface for pro users to code directly.
 - Login page.
- Backend server
 - at least allow 30+ users to operate concurrently, since there are 30 students in need of this software this semester.
 - Runs on unimelb cloud server (for now).
- Machine learning module
 - Feature engineering
 - Able to generate a report about the importance of every feature in a specific task.
 - Analyze the relationship between different features (optional).
 - Allow pro users to add new features.
 - Allow all users to define which features to use.
 - Algorithms
 - Provide various ML algorithm options (and their parameters perhaps) for users to choose from.
 - The logic can be exported into a jupyter notebook file (the nature of the exported file may change in the future).
 - Pro users should be able to modify the ML algorithms, even upload their own script.

Others

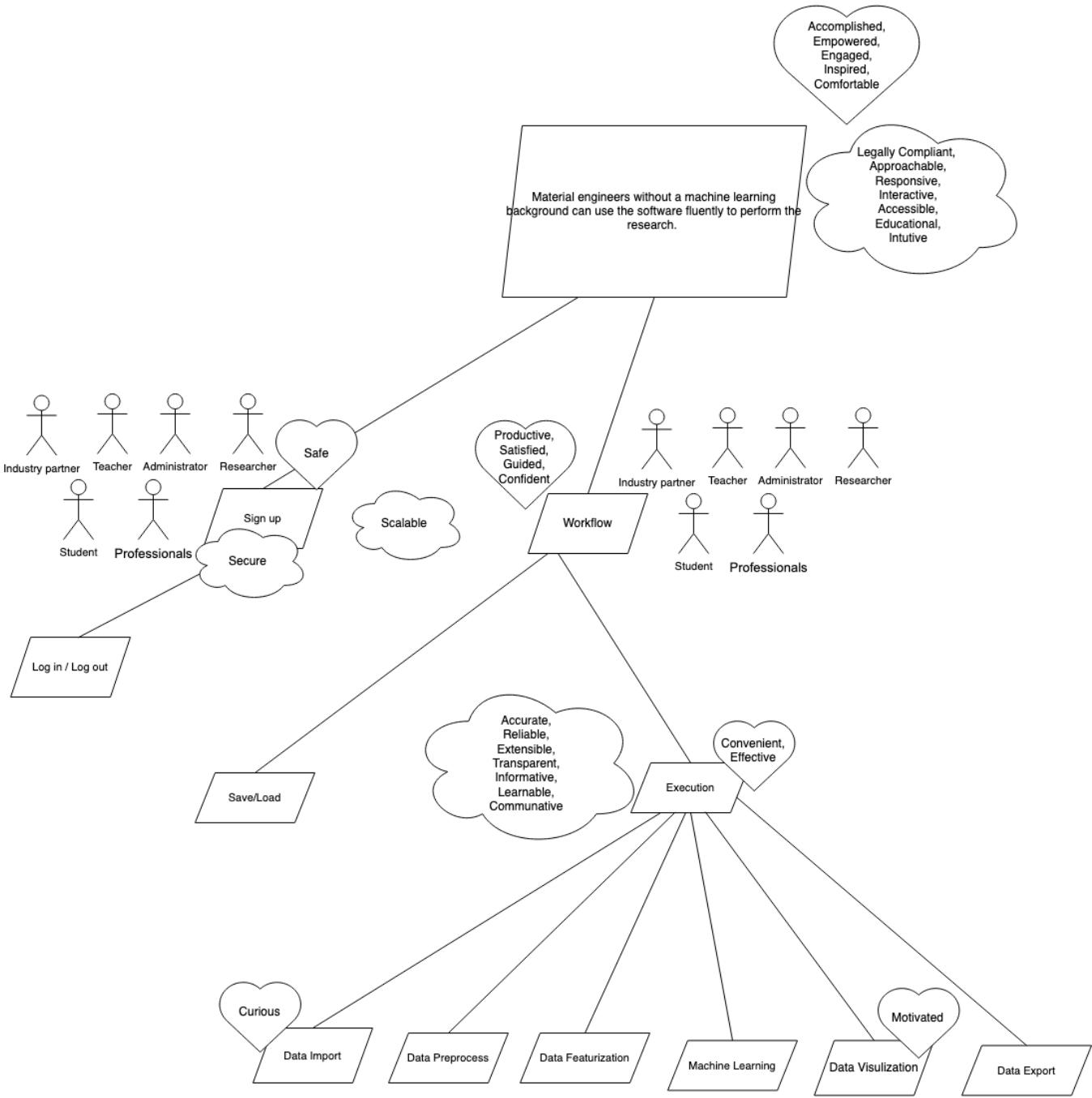
- Dataset is clean and reliable. It shouldn't require much preprocessing.
- The project should be based on [matminer](#).

| Motivational Model

Do-Be-Feel List - Collaboration Version 1 (Sprint 1 - Final)

| Who (users) | Do (functional goals) | Be (qualitative goals) | Feel (emotional goals) |
|-------------------|---|--|------------------------|
| Students | Add more database, machine learning method and plot types | Accessible | Accomplished |
| Administrators | Compare data using tables & plots | Accurate | Comfortable |
| Professionals | Data Pre-processing: Calculate descriptive statistics | Approachable | Confident |
| Industry Partners | Data Pre-processing: Consider anonymized data | Communicative | Convenient |
| Teachers | Data Pre-processing: Overview of the current import data | Educational | Curious |
| Researcher | Data Pre-processing: Reduces noise and eliminates ambiguity | Extensible | Effective |
| Code maintainers | Data Pre-processing: Standardizing data to bring it into the formatting range | Informative | Empowered |
| | Data Visualization: Data processing: Tabular data & Plotted Graph | Interactive | Engaged |
| | Edit python code directly in the interface | Intuitive | Guided |
| | Export input data | Learnable | Inspired |
| | Export jupyter notebook file | Legally Compliant | Motivated |
| | Export output data tables and figures | Reliable | Productive |
| | Featurization data: Add multiple composition-based features | Responsive | Safe |
| | Featurization data: Add multiple simple density features | Scalable | Satisfied |
| | Import Data: Create working spaces when importing | Secure | |
| | Import Data: Drag and drop import of files | Transparent (progress, error messages, notebook export...) | |
| | Import Data: Import data files (CSV, XES, Parquet) from local system | | |
| | Log in/Log out | | |
| | Machine Learning: Define input data and output data: Splitting data into training, test, and validation sets | | |
| | Machine Learning: Determining model features and training the model: Configure and adjust hyper parameters for optimum performance | | |
| | Machine Learning: Evaluate model performance and establish benchmarks: Continuous measurement and monitoring of model performance | | |
| | Machine Learning: Evaluate model performance and establish benchmarks: Evaluate models using validation methods and validation datasets | | |
| | Machine Learning: Get model results: The most important features of the current ML model | | |
| | Machine Learning: Select the machine learning model to be used | | |
| | Maintain software | | |
| | save/load workflows | | |
| | Sign up | | |

Goal Model - Collaboration Version 1 (Sprint 1 - Final)

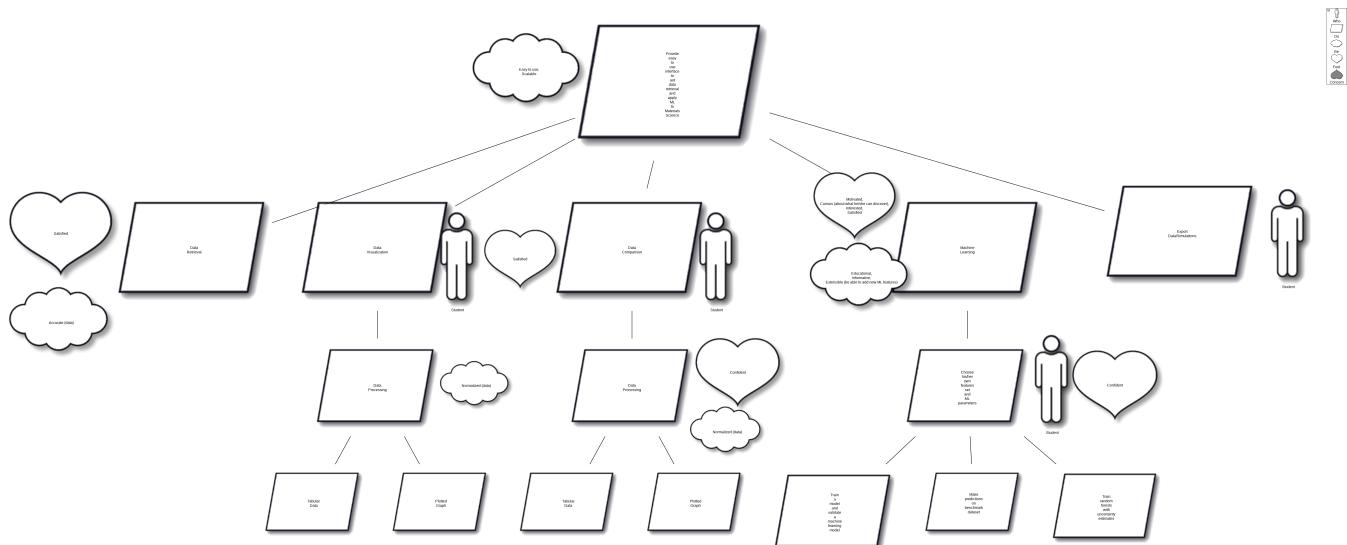


Do-Be-Feel List - BoxJelly Version 1

| Who | Do (Functional Goal) | Be (Quality Goal) | Feel (Emotional Goal) |
|----------|--|-------------------|-----------------------|
| Pro user | Create new features based on existing features | Easy to use | Motivated |
| Pro user | Edit python code directly in the interface | Accurate (data) | Satisfied |
| Pro user | Adding more database, machine learning method and plot types | Reliable | Satisfied |
| Students | Data Retrieval | Normalized (data) | Interested |
| Students | Data Visualization | Educational | Confident |
| | Data Processing | Tabular data | |

| Students | Data Comparison | Plotted Graph | Informative | Curious (about what he/she can discover) |
|------------------------------|---|---|---|--|
| Students Machine Learning | Choose his/her own features set and ML parameters | Train a model and validate a machine learning model | Extensible (be able to add new ML features) | Inspired / Satisfied |
| | | Make predictions on benchmark dataset | Scalable | |
| | | Train random forests with uncertainty estimates | Reliable | Confident |
| Students | Export Data/Simulations | | Easy to use | Convenient |
| Students | Running environment | | Stable | Satisfied |
| Students | Log in | | Intuitive | Safe |

Goal Model - BoxJelly Version 1



Change log

| Version date | Editor | Comment |
|--------------|------------------------|--|
| 08 Aug 2022 | Dara O hEidhin | Initial template |
| 15 Aug 2022 | Yaoming Xuan | Fill in the information about the students |
| 17 Aug 2022 | Dara O hEidhin | Merged do-be-feel and goal model into one page |
| 17 Aug 2022 | Felipe Leefu Huang Lin | Restructure Do-Be-Feel list |
| 17 Aug 2022 | Felipe Leefu Huang Lin | Upload first draft of goal model |
| 18 Aug 2022 | Zhaoqi WANG | Filled in some blanks |
| 18 Aug 2022 | Zhaoqi WANG | Added row on Do-Be-Feel list |
| 18 Aug 2022 | Dara O hEidhin | Added row on Do-Be-Feel list |
| 18 Aug 2022 | Radhimas Djan | Added row on Do-Be-Feel list |
| 21 Aug 2022 | Zhaoqi WANG | Updated with merged results (Final) with other teams |
| 22 Aug 2022 | Felipe Leefu Huang Lin | Added 3 teams collaboration Do-Be-Feel list and Goal Model version |

| Personas

Persona 1 - Prepared by team BoxJelly

Elaine Kim

age: 22
residence: Melbourne
education: Material Engineering
occupation: Postgraduate student
marital status: Single



"It's SO time spending to do material researches and get decent results through just a semester."

Motivation: As a material engineering graduate student, Assol gets frustrated and demotivated when she can't make sense of the data she has because she doesn't have a tool or sufficient programming/machine learning skills to process the material data. She is also frustrated that she can't use machine learning algorithms to help her engineer new materials even though she is told by her supervisor that this idea works in theory.

Comfort With Technology

INTERNET



SOFTWARE



MOBILE APPS



SOCIAL NETWORK



Criteria For Success:

Design a usable online tool which Assol can easily perform Materials data requests and retrievals and on a simple click of a button can predict the property of a material with features and ML algorithms chosen by her

Needs

- Easy-to-use Material science data processing and retrieval interface application
- A tool to predict property of a material with ML methods
- A powerful tool or technology supports material researching
- Get a HD on research assignment in each semester

Wants

- A data mining application that helps her research
- A better understanding on how Machine Learning can help her to learn more about a material
- Ability to use ML algorithms as a black box
- Freedom to select features on her own terms
- A tool to accelerate research progress

Values

- Convenience
- Quickness
- Safety
- High information density
- Flexible
- Understandable

Fears

- Spends hours without getting anything done because she neither has a adequate tool to do data mining, nor the programming skill to analyze the data herself.
- Looks at the generated diagram without knowing what it means

- Conducts countless experiments just to figure out the properties of the materials
- School assignments are strongly limited by time.
- Hard to choose suitable ML algorithms.

Persona 2 - Prepared by team BlueRing

Gray Zhou

age: 28
residence: Ningde, Fujian, China
education: Master of Material Engineering
occupation: R & D Engineer of Polymer
marital status: Single



"It is fantastic to apply a multi-function online tool with ML methods in my work, but only after it shows its efficiency and reliability to me. Nobody will refuse a tool that can save his time"

Motivation: Gray Zhou is a R & D Engineer of polymer in a battery factory. His work is searching for better materials for battery production. Gray spends lots of time testing different materials every day, but some of tests are waste of time because of the poor performance observed. He needs a system that can predict some useful properties of materials so that he can remove samples with low predicted performance and boost the research. Although his company provides some solutions, but they are awkward and only have limited functions.

Comfort With Technology

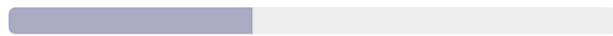
INTERNET



SOFTWARE



MOBILE APPS



SOCIAL NETWORK



Criteria For Success:

Provide a website or online-tool with quick, visual interface which can help him in daily development of new materials. A successful product should help him save noticeable time and energy on data processing and provide reliable prediction of properties.

Needs

- Retrieve and extract required data, process the data with ML methods to get some properties
- Provide graphs about properties he interested in which can be modified with interface
- Help him find the material with best predict properties

Wants

- Ability to interact with the graph to further compare serval materials in detail
- Explain what ML method the system applied and how it helps the prediction

Values

- Easy to get started
- Efficient back-end process
- Abilities to select functions and filter results
- Well organized visualization of interface and

Fears

- Frequently unavailability which may waste him much of time to retry
- Not enough guidance in the web or tool so him my feel confused to find functions he wants.
- No choice of properties that he needs

- ~~the organized visualization of interface and graphs~~

- ~~the choice of properties that he needs~~

- Lacking understand of what the system done, which may influence the confidence level of his report

Persona 3 - Prepared by team RedBack

Alex

age: 45

residence: Melbourne

education: Masters Degree in Physics

occupation: Materials Engineer

marital status: Divorced without kids



"There has got to be a better way to do this."

Alex has a Masters degree in Material Science. He has been working as a Materials Engineer for 15 years, he has been promoted to a senior role over time. Because physical experiments can take years, his job requires him to narrow down candidate materials for experimentation using Machine Learning and simulations

Comfort With Technology

PROGRAMMING WITH PYTHON



MACHINE LEARNING



CLOUD BASED STORAGE



MATERIAL SCIENCE



Criteria For Success:

Efficiency in finding the right materials

Accuracy of the results

Client satisfaction

Needs

- Products to accelerate his workflow
- Access to wide variety of related tools and resources

Wants

- Suitable models and featurizers for different use cases
- Demonstrate reproducible results to his clients
- Share resources with others
- Refining generated workflow to reuse

Values

- Extensible
- Accuracy
- Reliability
- Responsive
- Scalable
- Transparent

Fears

- Tool is too inflexible
- Losing access to progress on his work
- Not being able to verify his results
- Not having support with the tool

Change log

| Date | Version | Author | Comment | File |
|-------------|---------|------------------------|---|--|
| 15 Aug 2022 | 1.1 | Felipe Leefu Huang Lin | First draft |  persona.pdf |
| 16 Aug 2022 | 1.2 | Yaoming Xuan | version 2 of the persona |  persona.pdf |
| 16 Aug 2022 | 1.3 | Felipe Leefu Huang Lin | Used name generator to create a random name |  student.pdf |
| 16 Aug 2022 | 1.4 | Yaoming Xuan | Converted to docx file |  persona....960.docx |
| 16 Aug 2022 | 1.5 | @Yaoming Xuan | small modifications |  persona ...r15.docx |

| | | | | |
|-------------|-----|------------------------|---|--|
| 17 Aug 2022 | 1.6 | Zhaoqi WANG | <p>Suggestions on Quote, Needs, Wants, Values, Fears and Logo (Brand).</p> <p>Issues:</p> <ol style="list-style-type: none"> 1. can't upload more picture 2. extended to 2 pages 3. Name seems to be automatically changed from time to time |  Assol Ana...rator.pdf |
| 20 Aug 2022 | 1.7 | Yaoming Xuan | Generate a pdf file with latest content |  persona ver17.pdf |
| 22 Aug 2022 | 1.8 | Zhaoqi WANG | small modifications |  persona_1.8.pdf |
| 22 Aug 2022 | | Felipe Leefu Huang Lin | Added personas created by teams BlueRing and RedBack | |
| 22 Aug 2022 | | Dara O hEidhin | Embedded a pngs of our personas for site PDF export compatibility | |

Link to make additional changes: <https://personagenerator.com/7300af72-1c96-11ed-8d77-d742c7e0ee51/7300af73-1c96-11ed-8d77-5f5915d1141f>

| User Stories

Prioritization Technique

We used the MoSCoW prioritization classification.

Must have - must be included in the scope of the project, we defined this all the must have user stories can create a MVP

Should Have - should be included in the scope of the project

Could Have - could be included in the scope of the project

Won't Have - will not be included in the scope of the project

Team BoxJelly - Version 1

| Initiative | Epics | Users | ID | User Story | Prioritization (MoSCoW) |
|---------------------|---------|---------|----|--|-------------------------|
| Project Development | Epic #1 | Student | 1 | As a student I would like to be able to use different type of plotting graphs so that I have flexibility to visualize data according to my needs. | Must Have |
| | | | 2 | As a student I would like to be able to quickly browse the Materials available in the database for retrieval and simulations so that I can quickly perform queries. | Must Have |
| | | | 3 | As a student I would like to have access to content about basics of Machine Learning applied to Materials Science so that I become more motivated and engaged in applying ML to materials science. | Must Have |
| | | | 4 | As a student I would like to be able to use Matminer features without the need to write Python code so that I can spend more time doing data analysis. | Must Have |
| | | | 5 | As a student I would like to create a password protected user account so that my research remains private | Must Have |
| | | | 6 | As a student I would like to save project specific data/checkpoints so I can pick up where I left off for specific projects | Could Have |
| | | | 7 | As a student I would like to analyze the relationship between different features so that I can filter out unnecessary features that could skew the results | Must Have |
| | | | 8 | As a student I would like the application to provide hints and guidance for new users so that I can quickly learn how to use software | Should Have |
| | | | 9 | As a student I would like to export my work to a Jupyter Notebook so that I can extend my work beyond the capability of the application | Could Have |
| | | | 10 | As a student I would like the application to clean and tune data input so that I have less noise on visualizations. | Could Have |
| | | | 11 | As a student I would like the application to provide correct Machine Learning algorithms so that I can be confident of analyzing results. | Must Have |
| | | | 12 | As a student I would like to be able to make a machine learning model without fully understanding how to program them from start so that I am able to work with my task as soon as possible. | Must Have |
| | | | 13 | As a student I would like the application to give a hint during my work on what should I do next so that I can be confident with my workflow. | Could Have |
| | | | 14 | As a student I would like to be able to create an account so that the system can remember me and my work. | Must have |

| Initiative | Epics | Users | ID | User Story | Prioritization (MoSCoW) |
|---------------------|---------|----------|----|---|-------------------------|
| Project Development | Epic #1 | Pro User | 1 | As a pro user, I need to be able to add new datasets in the future so that if there's a new dataset that can be used on a new project, it can be added instantly. | Must Have |
| | | | 2 | As a pro user, I would like to edit python code on the interface so that I can have control how the ML algorithms works | Must Have |
| | | | 3 | As a pro user, I would like to upload my own script (in python) if possible. | Could have |
| | | | 4 | As a pro user, I would like to add new features so that in the future if they are useful it can be added. | Must Have |
| | | | 5 | As a pro user, I need to be able to manage all the accounts so that in case one of the students have technical issue I can help them. | Must Have |
| | | | 6 | As a pro user, I must also be able to act as normal users, so that I'm also able to use it in my research and figure out what issues other users may encounter. | Must have |
| | | | 7 | As a pro user, I would like to be able to cite all the machine learning algorithm I prepared for the user so that in the future, if user want to know more about how the algorithm work, they can refer to the original paper | Could have |
| | | | 8 | As a pro user, I would like to make sure that the application will be scalable to a lot of users so that when its peak hour, the application will not crash | Must have |
| | | | 9 | As a pro user, I must write documentation on how the code works so that when there's a developer that wants to extend the program, they can quickly understand it | Should have |

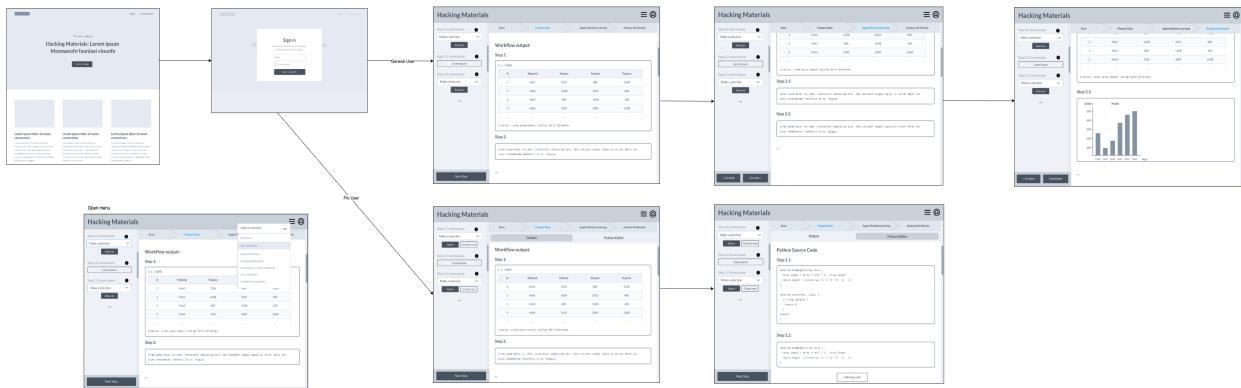
Teams Collaboration - Version 1.2 (Sprint 1 - Final)

| ID | | Role | | Action | Epic | | Goal | Size (days) | Priority | Assigned Team |
|----|------|--------------|-----------|--|------------------|---------|---|-------------|-----------------|---------------|
| 30 | As a | general user | I want to | be able to view the citations for used featurizers | Input Data | so that | I could be know more about the source of the featurizer (legally compliant) | 1 | 1 - Must have | RedBack |
| 32 | As a | general user | I want to | browse built-in featurizers | Input Data | so that | I can discover ways of manipulating my data | 1 | 1 - Must have | RedBack |
| 34 | As a | general user | I want to | browse built-in datasets | Input Data | so that | I can discover data to experiment with | 1 | 1 - Must have | RedBack |
| 19 | As a | student | I want to | quickly browse the Materials available in the database for retrieval and simulations | Input Data | so that | I can quickly perform queries. | 3 | 1 - Must have | RedBack |
| 21 | As a | general user | I want to | be able to select datasets from existing databases | Input Data | so that | I do not have to worry about how the data is loaded | 3 | 1 - Must have | RedBack |
| 37 | As a | general user | I want to | be able to preview the input data | Input Data | so that | I could explore the data | 1 | 2 - Should have | RedBack |
| 25 | As a | general user | I want to | Select specific features from a dataset | Input Data | so that | I can improve the precision of my model | 3 | 2 - Should have | RedBack |
| 13 | As a | Pro user | I want to | add new features | Input Data | so that | they can be reused in the future | 5 | 2 - Should have | RedBack |
| 28 | As a | general user | I want to | be able to reference / view citation for original data sources | Input Data | so that | I can retrieve data. | 1 | 3 - Could have | RedBack |
| 18 | As a | pro user | I want to | be able to apply new featurizers | Input Data | so that | I can create new features | 3 | 3 - Could have | RedBack |
| 1 | As a | student | I want to | clean and tune data input | Input Data | so that | I have less noise on visualizations. | 5 | 3 - Could have | RedBack |
| 29 | As a | student | I want to | save project specific data/checkpoints | Administration | so that | I can pick up where I left off for specific projects | 1 | 1 - Must have | BoxJelly |
| 35 | As a | pro user | I want to | export model selections, parameters, and data flows | Administration | so that | I can save my work and share it with others | 1 | 1 - Must have | BoxJelly |
| 36 | As a | pro user | I want to | import exported model selections, parameters, and data flows | Administration | so that | I can continue work I had previously saved | 1 | 1 - Must have | BoxJelly |
| 20 | As a | student | I want to | Create an account using single-sign on, restricted to the *.unimelb.edu.au domain | Administration | so that | my research remains private | 3 | 1 - Must have | BoxJelly |
| 23 | As a | pro user | I want to | Control job execution | Administration | so that | I can start, view progress of, and cancel jobs related to my project | 3 | 1 - Must have | BoxJelly |
| 10 | As a | pro user | I want to | be able to opt in to pro-user features | Administration | so that | I can access pro user features | 5 | 1 - Must have | BoxJelly |
| 38 | As a | pro user | I want to | have my pro user settings persist on each visit | Administration | so that | I don't have to reconfigure settings to use the features I need | 1 | 2 - Should have | BoxJelly |
| 24 | As a | student | I want to | receive provided hints and guidance for new users | Administration | so that | I can quickly learn how to use software | 3 | 2 - Should have | BoxJelly |
| 14 | As a | pro user | I want to | easily find and read documentation on the pro features | Administration | so that | I can use them with ease | 5 | 2 - Should have | BoxJelly |
| 17 | As a | pro user | I want to | Be kept informed about job status | Administration | so that | I can avoid polling my workspace to check for results | 3 | 3 - Could have | BoxJelly |
| 6 | As a | pro user | I want to | have access to more processing power | Administration | so that | I can run more complex operations or use more data | 5 | 3 - Could have | BoxJelly |
| 31 | As a | general user | I want to | able to select a Machine Learning model | Machine Learning | so that | I could use it to train and run the data | 1 | 1 - Must have | BlueRing |
| 33 | As a | general user | I want to | browse built-in ML models | Machine Learning | so that | I can discover ways of manipulating my data | 1 | 1 - Must have | BlueRing |
| 39 | As a | user | I want to | be able to select split ratio of data | Machine Learning | so that | to train and test the model | 1 | 2 - Should have | BlueRing |

| | | | | | | | | | | |
|----|------|--------------|-----------|--|--------------------|---------|--|---|-----------------|----------|
| 26 | As a | pro user | I want to | have the option to change the hyperparameters used in the machine learning model | Machine Learning | so that | I can fine tune my test results. | 3 | 2 - Should have | BlueRing |
| 15 | As a | pro user | I want to | be able use additional ML models | Machine Learning | so that | I can improve accuracy | 5 | 2 - Should have | BlueRing |
| 7 | As a | pro user | I want to | combine multiple models together | Machine Learning | so that | I can model more complex data manipulations | 5 | 3 - Could have | BlueRing |
| 22 | As a | general user | I want to | see clear annotation or explanation of data points and features | Data Visualisation | so that | I can understand the results of the analysis | 3 | 1 - Must have | TBD |
| 8 | As a | student | I want to | use different type of plotting graphs | Data Visualisation | so that | I have flexibility to visualize data according to my needs. | 5 | 1 - Must have | TBD |
| 9 | As a | general user | I want to | able to view and plot the results of the model | Data Visualisation | so that | I could analysis and visualise the effects of the model | 5 | 1 - Must have | TBD |
| 12 | As a | student | I want to | export my work to a Jupyter Notebook | Jupyter Notebook | so that | I can extend my work beyond the capability of the application | 5 | 2 - Should have | TBD |
| 2 | As a | general user | I want to | attach comments to workflow objects | Jupyter Notebook | so that | I can document my work | 5 | 3 - Could have | TBD |
| 4 | As a | Pro user | I want to | edit python code on the interface | Jupyter Notebook | so that | I can have control how the ML algorithms works | 5 | 3 - Could have | TBD |
| 5 | As a | Pro user | I want to | upload my own script (in python) if possible | Jupyter Notebook | so that | I can extend the tool to support custom models and featurizers | 5 | 3 - Could have | TBD |
| 27 | As a | pro user | I want to | be able to access new databases | External Data | so that | I can access addional data | 3 | 2 - Should have | TBD |
| 3 | As a | Pro user | I want to | be able to add new datasets in the future | External Data | so that | if there's a new dataset that can be used on a new project, it can be added instantly | 5 | 3 - Could have | TBD |
| 11 | As a | student | I want to | analyze the relationship between different features | | so that | I can identify which features I need to select for my analysis | 5 | 2 - Should have | TBD |
| 16 | As a | general user | I want to | add specific materials to the workflow for analysis | | so that | compare the performance of the specific material my client or I choose with other material | 3 | 3 - Could have | TBD |

| Prototype

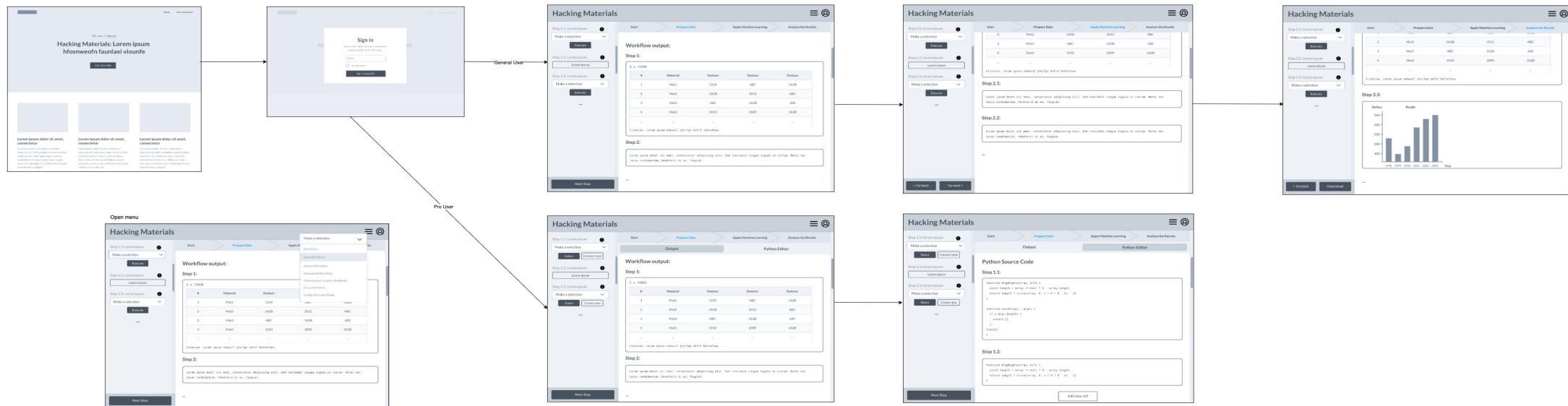
Teams Collaboration - version 1 (Sprint 1 - Final)



Descriptive Notes (recorded by team redback)

- Landing page:
 - Static page with information about the app and project
 - Link to access the app
 - On click, it opens a login modal
 - Once user is logged in, they're redirected to the app
- Single page app:
 - Top bar:
 - User profile button at the top opens a menu to give the user the option to log out
 - Menu button at the top has options to import or save a workflow, download it in different formats, start over, a link to the documentation and a toggle to enable pro view.
 - General user:
 - The workflow is divided into major and minor steps. Each major step would have its own page. User can go back and forth between the major steps as needed.
 - Left panel:
 - All the minor steps are numbers and named to guide the user
 - Inputs can be of different types
 - Each step has a tooltip button that would open a modal with guidance information about the step
 - The steps and options in the left panel should always be the same no matter what selections the user made in previous steps. Any step that requires customised inputs would open in a modal.
 - Example 1: Step 3.1 might be "Selecting a plot type". As there is a known, limited list of different plot types, this step may be a drop-down menu that is displayed directly in the left panel.
 - Example 2: Step 3.2 might be customising the selected plot's configuration options. As different plot types may need different configuration options, these options will not be displayed in the panel directly. Instead, the panel will include only a button that says "Configure plot", which would open a modal with the specific options applicable to the selected plot type.
 - Pinned buttons at the bottom of the panel: navigate between the different major steps. Last step page may also have a button to download the full workflow.
 - Viewing window:
 - At the top of the viewing window, the user can see the progression of major steps with the current step highlighted.
 - The output of each minor step is labelled with the step number and contained inside a box. The output inside the box is the same output produced by running the python code, simply copied over for transparency.
 - The outputs from the previous pages are also always displayed, so it's not just the outputs of the current page.
 - Where a resource with citations is used, the citations will be automatically printed after the output of the step where the resource was selected.
 - Pro user:
 - Left panel: has all the same options as a general user, plus additional buttons to configure their own settings as needed
 - Viewing window: the window has 2 tabs:
 - Output: same the as the viewing window of the general user
 - Python source code:
 - An editable view of all the code generated by their selections, looks similar to a Jupyter notebook.
 - User can add new cells as desired
 - Brings up the following question: what happens if the user edits the code generated by one of the steps? This may lead to inconsistencies between what is shown in the step's input field and what the code now actually does. This is an implementation decision so is not a major concern right now, but one option that we decided to show in the prototype is that the step's input in the left panel would change to say "Custom" or something similar, indicating that the configuration was changed.

Teams Collaboration - version 1 (Sprint 1 - Final)



Change log

| Version date | Editor | Comment |
|--------------|--|--|
| 20 Aug 2022 | Felipe Leefu Huang Lin | First prototype draft made in collaboration with team redback and bluering |
| | | |

| Business Case

Project Hacking Materials Cross-Team Collaboration Model Proposal

COMP90082 Software Project SM2 2022

The current situation

We currently have three teams working with the same client, Dr Christian Brandl, on the project Hacking Materials ("HA"). The client expressed that he would not be interested in three different versions of the same product, and would prefer we work together to be able to build one more complete final product.

Additionally, it is clear from our discussions with the client that any resulting single product will be difficult to decompose into totally independent components. This means that the deliverables produced by each team will be dependent on those produced by other teams, requiring collaboration on design, development, and project management tasks.

This issue is compounded by the fact that each team has a separate workspace set up by the university, in particular the separate GitHub repositories. This makes it extremely difficult to share resources including source code and development resources, and to determine how much progress is being made by each team, and on which tasks.

Our biggest concern is that our teams will be assessed separately. Working on the same product makes us inherently reliant on each other, which could affect the assessment of each team. We understand that this complicates the situation, and have come up with a working agreement that we believe would allow us to deliver the product that the client wants while minimizing the risks to the assessable components of the project.

Obstacles

We have identified several issues with the way the project is currently being approached, including the following:

- Project scope is larger than what could realistically be accomplished by one student team.
- If the teams were to work completely separately on different parts of the product, it will be difficult to organize ourselves in a way that ensures the needs of all teams are met.
- Finding common availability between 15 people for meetings is difficult. As such, cross-team communication so far has been done mainly through team representatives. This approach is slow, high effort for team representatives (as they need to be aware of everything their teams are doing at all times) and creates an increased risk of information getting lost in transmission.

Proposed work structure / way of working

The structure we are proposing is as follow:

- The teams would work on the same GitHub repository.
- Each team works as a full-stack team focusing on a particular epic of stories,
- Someone from each team is involved in each technical area of the product (Backend/ Frontend/ Machine Learning).
- The members of the different teams who are within the same technical area cooperate to ensure they follow the same standards and processes. This will help spread some of the cross-team communication load away from the team representatives.
- Each team will need to review the other teams' work to ensure that it doesn't interfere or affect their own work.
- Ownership of epics will be allocated to teams, but the teams will work together to ensure all high priority user stories are delivered first. This creates fewer dependencies between teams and allows each team to deliver fully functional components without relying on the others.
- Each team will be responsible for their own Confluence space, but the teams will maintain a similar structure to make navigation easy. Major structural changes will be reported and shared with the other teams during cross-team meetings.
- The teams will share certain deliverables within their Confluence spaces so that the client would not have three different versions of the same document.
- Sprint documents exported from each team's Confluence space will be included within their own folder in the shared GitHub repository.
- The teams will take turns organizing meetings with the client, supervisor and each other.

Requested actions

- **Decision from teaching staff:** University administration hasn't made a decision on assessment criteria, or whether the teams can collaborate on this project. This prevents us from adopting a shared working model because we are not sure whether our project submissions will be accepted by the university, or whether we will be otherwise penalized for taking this approach. This was discussed with our supervisor, Mauro, and it was agreed that a response would need to be available by Tuesday 3pm.
- **Shared GitHub repository:** If our proposal is accepted, we would need a shared GitHub repository so that Dr. Brandl will not have three copies of each deliverable.



HA Software Proj...usiness case.pdf

| Plan

Scenario 1 - If the projected is divided between the three teams then we expect to split each epic equally between the three teams.

Scenario 2 - If we work as separate teams the user stories will be addressed in order of priority and entered into Trello. Please refer to Business Case and User Stories

Collaborative changes, on GitHub, will pass a reviewing process with each sprint being released with a tag.

| Sprint Artifacts

| Links |
|------------------------------------|
| - Sprint Review |
| - Retrospectives |
| - Product planning |

| Sprint Review



Trello Board link

<https://trello.com/b/PzZuNQMk/boxjelly-sprint>

| Retrospectives

Sprint retrospective (Reflection):

| Sprint | Link |
|----------|--|
| Sprint 1 | Sprint 1 - Retrospective |