

1. Home	2
1.1 Assignment	4
1.1.1 1 Sprint 1	5
1.1.2 Sprint Log	7
1.1.2.1 1 Sprint 1	8
1.1.2.2 2 Sprint 2	9
1.1.2.3 3 Sprint 3	10
1.1.2.4 Preparation	11
1.2 Requirements	12
1.2.1 Project Description	13
1.2.2 Sprint Plan	15
1.2.3 Motivational Model	17
1.2.4 User Story	21
1.2.5 Persona	24
1.2.6 Prototype	32
1.2.7 Business case	34
1.3 Communication Archive	36
1.3.1 Client	37
1.3.2 Supervisor	38
1.4 Resources	39
1.4.1 Contacts	40

Home

COMP90082 Hacking Materials User Interface (code: HA)

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. Your project aims 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 (skikit learn, Keras), and visualising results.

Goal

Provide a user interface for standard processes in MatMiner: retrieving data from a database, users selecting features to be extracted from the database, performing simple machine learning tasks (Skikit Learn, Keras), and visualizing the results.

Project team





Dr Christian Brandl
Client



Mauro Mello Jr
Supervisor

Team Members

Useful links

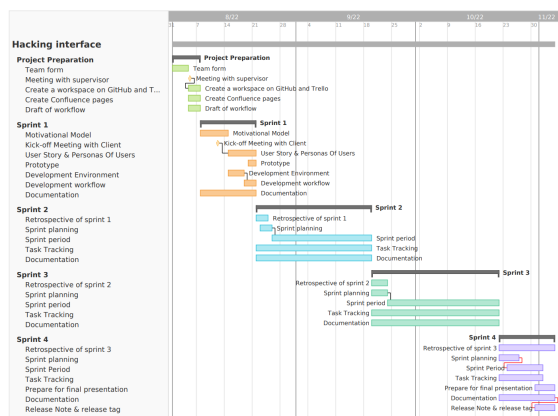
Trello	Github
	

Tasks

- ✓ **Edit this home page** - Click *Edit* in the top right of this screen to customize your Space home page
- ✓ **Create your first page** - Click the *Create* button in the header to get started
- ✓ **Brand your Space** - Click *Configure Sidebar* in the left panel to update space details and logo
- ✓ **Set permissions** - Click *Space Tools* in the left sidebar to update permissions and give others access

				
Yanan Liu (Rep)	Hongpei Lu	Jiahao Ju	Xinle Yu	Rui Zhang
Yanan Liu	Hongpei Lu	Jiahao Ju	Xinle Yu	Rui ZHANG

Project Timeline



Recent space activity

Yanan Liu



| Sprint Plan 3 •

| Persona 13 •

Xinle Yu



| Project Description 25 •

Yanan Liu



| Business case 32 •

1 | Sprint 1 1 •

| Assignment

Sprint

1 | [Sprint 1](#)

1 | Sprint 1

Sprint 1 Assessment Checklist

This checklist helps you double check your 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 (???)

- ✓ 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 Log

Preparation

1 Sprint 1

2 Sprint 2

3 Sprint 3

1 Sprint 1

- Background description, client goals and motivation

[| Project Description](#)

[| Motivational Model](#)

- User story and Persona

[| User Story](#)

[| Persona](#)

- Development environment and technical details

- Plan for Sprint 2 and 3

[| Sprint Plan](#)

- GitHub workspace

[GitHub Link](#)

- Documentation (Confluence, Trello & Meeting Notes)

[| Meetings](#)

[Trello - Sprint 1](#)

2 Sprint 2

- Task Tracking & Version Control
- Product & GitHub
- Documentation (Confluence, Trello & Meetings)

3 Sprint 3

- Task Tracking & Version Control
- Product & GitHub
- Documentation (Confluence, Trello & Meetings)

Preparation

- Create workspace on Confluence and Trello
- Meeting with Client and supervisor
- Assign tasks in Sprint 1

| Requirements

| [Motivational Model](#)

| [User Story](#)

| [Persona](#)

| [Project Description](#)

| Project Description

Background

- In computational materials engineering, the engineering design of new materials is accelerated by avoiding costly and lengthy experiments to demonstrate the properties of materials. Data mining methods discover better materials by searching computer generated databases and making simulation predictions using high throughput and high-performance computing.
- Unfortunately, for some of the first movers, such as Tesla, translating this new thinking into engineering practice is still in its infancy. One obstacle is material engineers' software tools.
- Engineers with knowledge of the material domain need to effectively access and process this data in order to make informed decisions for the final machine learning strategy.

Goal

- Provide a user interface for standard processes in MatMiner: retrieving data from a database, users selecting features to be extracted from the database, performing simple machine learning tasks (Skikit Learn, Keras), and visualizing the results.

Sponsor

- **Dr. Brandl**
- Lecturer and sponsor of UOM.
- Completed his Ph.D. in Materials Science & Engineering at the École Polytechnique Fédérale de Lausanne EPFL and the Paul Scherrer Institute (Switzerland), followed by a post-doc in the Theoretical Division at the Los Alamos National Laboratory (USA).
- His research focuses on predictive atomistic simulation approaches to enable rational materials design.

In-scope

The project is basically design and build the following:

- A web page tool can help pro user and regular user or student who has no or very little experience in machine learning analysis method to complete material analysis;

All type users requirements:

- Detailed analysis of user requirements;
- Design a beautified, user friendly interface;
- Design a login page;
- Provide hints and guidance for new users;
- Design various machine learning models in Python that allow users to select features from a database and obtain a visual result;
- Allow to add new features;
- Provide code interface for pro user;
- The workflow allows to download;
- Machine learning model accuracy reaches a sufficient height;
- Complete the improvement suggestions that may be put forward by the clients;
- At least allow 30+ users to operate concurrently;
- Report the progress and results of the project;
- Runs on unimelb cloud server;
- The final project results passed the test.

Out-of-Scope

- The team has no obligation to release back-end code to the public;
- Operating maintenance is not within the scope of the team's obligations;
- The interpretability of the machine learning model does not need to be guaranteed, which means that the team members can choose the black box model as the machine learning model;
- Due to the lack of resources, the backend server cannot guarantee the stability of multiple people entering at the same time.

Techniques

- Web Techniques
 - NodeJS is a back-end JavaScript runtime environment for building scalable network applications. In this project, it is applied to build the structure of the online tool.
- Frontend
 - React is a JavaScript library for building user interfaces. It is applied to design and achieve designed functions.
- Backend

- Django is a Python-based web framework. The backend process and management are based on it.
- Machine Learning
 - [MatMiner](#) is a Python library for data mining the properties of materials, the feature extraction from database will be achieve based on it.
 - [Scikit Learn](#) is a Python library for Machine Learning with various algorithms. It is used for predicting properties of materials.

| Sprint Plan

Assigned to @Rui Zhang, Xinle Yu, Hongpei Lu

Can download file in here: [Hacking_interface_Gantt Chart.pdf](#), see original work in [here](#).

Project Preparation (Aug 1 - Aug 7)

- Team form (Aug 1-4)
- Meeting with supervisor (Aug 5)
- Create a workspace on GitHub and Trello (Aug 5-7)
- Create Confluence pages (Aug 5-7)
- Draft of the workflow (Aug 5-7)

Sprint 1: Inception (Aug 8 - Aug 21)

- Kick-off Meeting with Client (Aug 12)
- Motivational Model (Aug 8 - Aug 14)
- User Story & Personas Of Users (Aug 15 - Aug 21)
 - Student: Box Jelly
 - Industry: Blue Ring
 - Pro-user: Red Back
- Prototype (Aug 20 - Aug 21)
- Development Environment (Aug 15 - Aug 18)
- Development workflow (Aug 19 - Aug 21)
- Documentation (Aug 8 - Aug 21)

Sprint 2: Development (Aug 22 - Sep 19)

- Retrospective of sprint 1 (Aug 22 - Aug 24)
- Sprint planning (Aug 23 - Aug 25)
- Sprint period (Aug 26 - Sep 19)
- Task Tracking (Aug 22 - Sep 19)
- Documentation (Aug 22 - Sep 19)

Sprint 3: Development (Sep 20 - Oct 21)

- Retrospective of sprint 2 (Sep 20 - Sep 23)
- Sprint planning (Sep 20 - Sep 23)
- Sprint period (Sep 24 - Oct 21)
- Task Tracking (Sep 20 - Oct 21)
- Documentation (Sep 20 - Oct 21)

Sprint 4: Product (Oct 22 - Nov 4)

- Retrospective of sprint 3 (Oct 22 - Oct 26)
- Sprint planning (Oct 22 - Oct 26)
- Sprint period (Oct 24 - Nov 1)
- Task Tracking (Oct 22 - Nov 4)
- Prepare for final presentation (Oct 31 - Nov 4)
- Documentation (Oct 22 - Nov 4)
 - Release Note & release tag (Oct 31 - Nov 4)

Gantt Chart

Hacking interface

Project Preparation

- Team form
- Meeting with supervisor
- Create a workspace on GitHub and T...
- Create Confluence pages
- Draft of workflow

Sprint 1

- Motivational Model
- Kick-off Meeting with Client
- User Story & Personas Of Users
- Prototype
- Development Environment
- Development workflow
- Documentation

Sprint 2

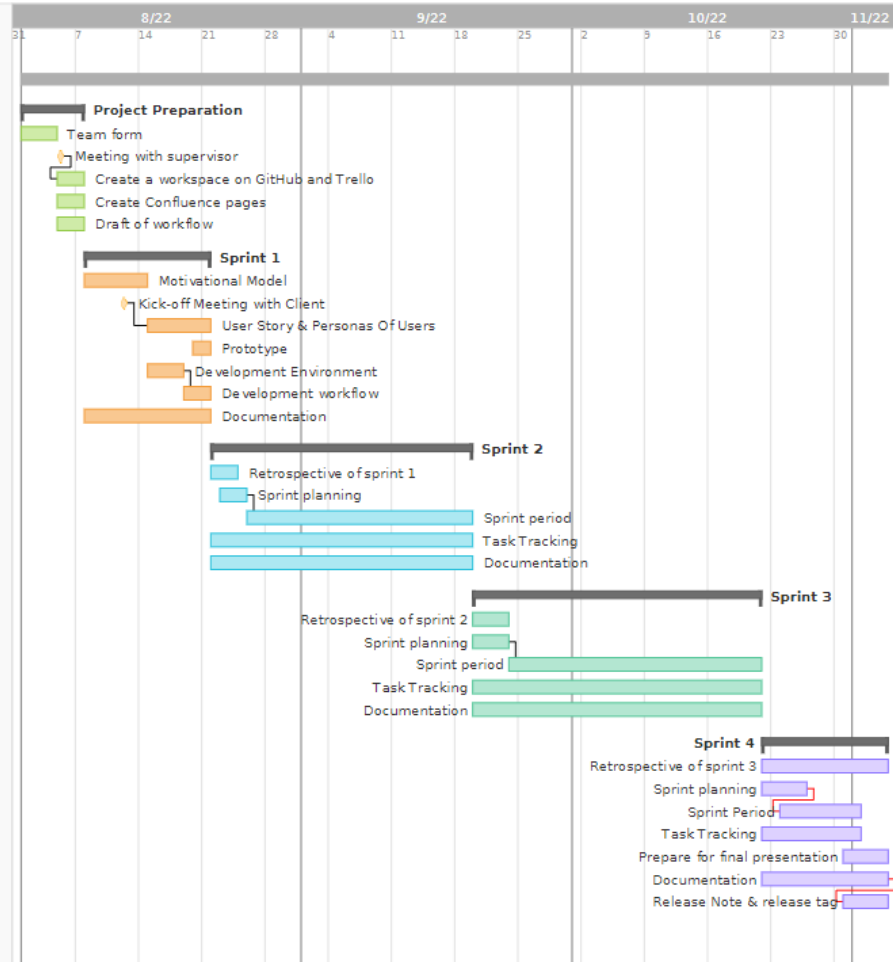
- Retrospective of sprint 1
- Sprint planning
- Sprint period
- Task Tracking
- Documentation

Sprint 3

- Retrospective of sprint 2
- Sprint planning
- Sprint period
- Task Tracking
- Documentation

Sprint 4

- Retrospective of sprint 3
- Sprint planning
- Sprint Period
- Task Tracking
- Prepare for final presentation
- Documentation
- Release Note & release tag



| Motivational Model

Assigned to [Yanan Liu](#) , [Hongpei Lu](#)

Versions

Version ID	Description	Editor	Date
1.0	First version based on initial understanding of the project and first client meeting	Yanan Liu	2022-8-17
1.1	Add a goal model based on the initial do-be-feel list	Hongpei Lu	2022-8-18
2.0	All team merged result	Yanan Liu Hongpei Lu	2022-8-19

Version 2.0

Group cooperation work, worked in [here](#)

Do-Be-Feel-Who List

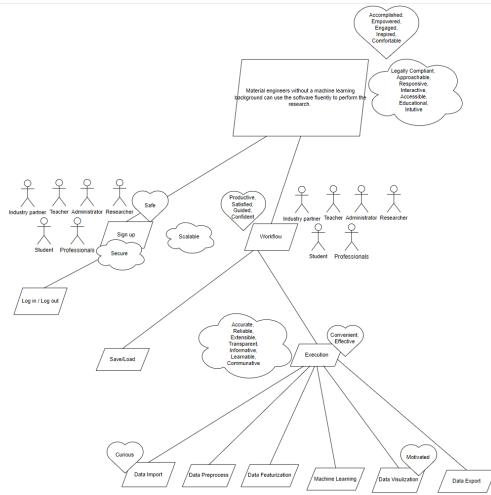
Overall goal: *Material engineers without a machine learning background can use the software fluently to perform the required data analysis*

Who (Roles)	Do(Functional Goal)	Be(Quality Goal)	Feel (Emotion Goal)
Students	Sign up	Accurate	Accomplished
Administrators	Log in/Log out	Approachable	Effective
Professionals	Import Data: Import data files (CSV, XES, Parquet) from local system	Reliable	Satisfied
Industry Partners	Import Data: Drag and drop import of files	Scalable	Productive
Teachers	Import Data: Create working spaces when importing	Extensible	Empowered
Researcher	Data Pre-processing: Overview of the current import data	Transparent (progress, error messages, notebook export...)	Engaged
Code maintainers	Data Pre-processing: Calculate descriptive statistics	Intuitive	Curious
	Data Pre-processing: Reduces noise and eliminates ambiguity	Accessible	Confident
	Data Pre-processing: Consider anonymized data	Secure	Comfortable
	Data Pre-processing: Standardizing data to bring it into the formatting range	Responsive	Guided
	Featurization data: Add multiple composition-based features	Learnable	Motivated
	Featurization data: Add multiple simple density features	Communicative	Convenient
	Machine Learning: Select the machine learning model to be used	Interactive	Safe
	Machine Learning: Define input data and output data: Splitting data into training, test, and validation sets	Legally Compliant	Inspired
	Machine Learning: Determining model features and training the model: Configure and adjust hyperparameters for optimum performance	Educational	
	Machine Learning: Evaluate model performance and establish benchmarks: Evaluate models using validation methods and validation datasets	Informative	
	Machine Learning: Evaluate model performance and establish benchmarks: Continuous measurement and monitoring of model performance		
	Machine Learning: Get model results: The most important features of the current ML model		
	Save / Load workflows		

	Export Result: Data Visualization & Comparison - Tabular data & Plotted Graph		
	Export Result: Pure data table		
	Export jupyter notebook file		
	Edit python code directly in the interface		
	Adding more database, machine learning method and plot types		
	Export Data(input)		
	Maintain software		

Goal model

New link: https://drive.google.com/file/d/1Jc3ncM2E6vEhf4OMOeaYWBSvd02prV4r/view?usp=share_link



Version 1.1

Do-Be-Feel-Who List

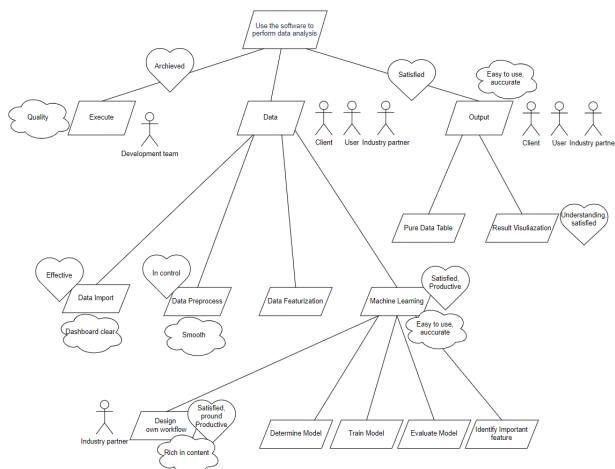
Overall goal: Material engineers without a machine learning background can use the software fluently to perform the required data analysis

Who (Roles)	Do(Functional Goal)		Be(Quality Goal)	Feel (Emotion Goal)	Concerns
Development team	<ul style="list-style-type: none"> Execute project 		<ul style="list-style-type: none"> Quality delivers 	<ul style="list-style-type: none"> Achievement 	<ul style="list-style-type: none"> time/ cost constraints Inexperience in
Client	<ul style="list-style-type: none"> Import Data 	<ul style="list-style-type: none"> Import data files (CSV, XES, Parquet) from local system Drag and drop import of files Create working spaces when importing 	<ul style="list-style-type: none"> Easy-to-use Smooth Accurate (data) Dashboard clear 	<ul style="list-style-type: none"> Effective In control Satisfied Understanding Productive 	<ul style="list-style-type: none"> Too technical
Student (User)					
Industry Partner					

Research er (User)	<ul style="list-style-type: none"> Data Pre-processing 	<ul style="list-style-type: none"> Overview of the current import data Removing unneeded columns from the data set Calculate descriptive statistics Reduces noise and eliminates ambiguity Consider anonymized data Standardizing data to bring it into the formatting range 			
	<ul style="list-style-type: none"> Featurization data 	<ul style="list-style-type: none"> Add multiple composition-based features Add multiple simple density features 			
	<ul style="list-style-type: none"> Machine Learning <ul style="list-style-type: none"> Select the machine learning model to be used 	<ul style="list-style-type: none"> Define input data and output data 	<ul style="list-style-type: none"> Splitting data into training, test, and validation sets 	<ul style="list-style-type: none"> Easy-to-use Smooth Accurate (data) Dashboard clear More ML models are available Visualization of model quality results 	<ul style="list-style-type: none"> No understanding of common assessment model methods
		<ul style="list-style-type: none"> Determining model features and training the model 	<ul style="list-style-type: none"> Configure and adjust hyperparameters for optimum performance 		<ul style="list-style-type: none"> Too technical
		<ul style="list-style-type: none"> Evaluate model performance and establish benchmarks 	<ul style="list-style-type: none"> Evaluate models using validation methods and validation datasets Continuous measurement and monitoring of model performance 		
		<ul style="list-style-type: none"> Get model results 	<ul style="list-style-type: none"> The most important features of the current ML model 		
	<ul style="list-style-type: none"> Export Result 	<ul style="list-style-type: none"> Visualization results Pure data table 			
Industry Partner	<ul style="list-style-type: none"> Better performance Complete IDE 		<ul style="list-style-type: none"> Rich in content 	<ul style="list-style-type: none"> Effective In control Satisfied Productive Proud 	

Goal model

<https://drive.google.com/file/d/1Jc3ncM2E6vEhf4OMOeaYWBSvd02prV4r/view?usp=sharing>



| User Story

Assigned to Rui ZHANG

Versions

Version ID	Description	Editor	Date
1.0	First version based on initial understanding of the project and first client meeting	Rui ZHANG	2022-8-18
1.1	Added size estimation, priority, and justifications.	Rui ZHANG	2022-8-18
1.2	Supplement the remaining User Stories	Jiahao Ju	2022-8-18
1.3	Changed priority to MoSCoW priority. Added user story 5.	Rui ZHANG	2022-8-18
1.4	Delete unrelated stories. Add some general stories.	Jiahao Ju	2022-8-19
2.0	Merged user stories from all three teams	Rui ZHANG	2022-8-20
2.1	Reordered the user stories and grouped by domain. Changed size ti estimation of days	Rui ZHANG	2022-8-21
2.2	Re group Action to new Epic and change original Epic to Domain	Yanan Liu	2022-8-21

Version 2.1

Group cooperation work, worked in [here](#)

ID		Role		Action	Epic	Domain		Goal	Size (days)	Priority
30	As a	general user	I want to	be able to view the citations for used featurizers	Input Data	Documenta tion	so that	I could be know more about the source of the featurizer (legally compliant)	1	1 - Must have
32	As a	general user	I want to	browse built-in featurizers	Input Data	ML	so that	I can discover ways of manipulating my data	1	1 - Must have
34	As a	general user	I want to	browse built-in datasets	Input Data	Server	so that	I can discover data to experiment with	1	1 - Must have
19	As a	student	I want to	quickly browse the Materials available in the database for retrieval and simulations	Input Data	UI	so that	I can quickly perform queries.	3	1 - Must have
21	As a	general user	I want to	be able to select datasets from exisiting databases	Input Data	Data Manipulation	so that	I do not have to worry about how the data is loaded	3	1 - Must have
37	As a	general user	I want to	be able to preview the input data	Input Data	UI	so that	I could explore the data	1	2 - Should have
25	As a	general user	I want to	Select specific features from a dataset	Input Data	Data Manipulation	so that	I can improve the precision of my model	3	2 - Should have
13	As a	Pro user	I want to	add new features	Input Data	ML	so that	they can be reused in the future	5	2 - Should have
28	As a	general user	I want to	be able to reference / view citation for original data sources	Input Data	Documenta tion	so that	I can retrieve data.	1	3 - Could have
18	As a	pro user	I want to	be able to apply new featurizers	Input Data	ML	so that	I can create new features	3	3 - Could have
1	As a	student	I want to	clean and tune data input	Input Data	Data Manipulation	so that	I have less noise on visualizations.	5	3 - Could have
29	As a	student	I want to	save project specific data/checkpoints	Administrat ion	Server	so that	I can pick up where I left off for specific projects	1	1 - Must have
35	As a	pro user	I want to	export model selections, parameters, and data flows	Administrat ion	Server	so that	I can save my work and share it with others	1	1 - Must have

36	As a	pro user	I want to	import exported model selections, parameters, and data flows	Administration	Server	so that	I can continue work I had previously saved	1	1 - Must have
20	As a	student	I want to	Create an account using single-sign on, restricted to the *.unimelb.edu.au domain	Administration	Server	so that	my research remains private	3	1 - Must have
23	As a	pro user	I want to	Control job execution	Administration	Server	so that	I can start, view progress of, and cancel jobs related to my project	3	1 - Must have
10	As a	pro user	I want to	be able to opt in to pro-user features	Administration	UI	so that	I can access pro user features	5	1 - Must have
38	As a	pro user	I want to	have my pro user settings persist on each visit	Administration	Server	so that	I don't have to reconfigure settings to use the features I need	1	2 - Should have
24	As a	student	I want to	receive provided hints and guidance for new users	Administration	UI	so that	I can quickly learn how to use software	3	2 - Should have
14	As a	pro user	I want to	easily find and read documentation on the pro features	Administration	Documentation	so that	I can use them with ease	5	2 - Should have
17	As a	pro user	I want to	Be kept informed about job status	Administration	Server	so that	I can avoid polling my workspace to check for results	3	3 - Could have
6	As a	pro user	I want to	have access to more processing power	Administration	Server	so that	I can run more complex operations or use more data	5	3 - Could have
31	As a	general user	I want to	able to select a Machine Learning model	Machine Learning	ML	so that	I could use it to train and run the data	1	1 - Must have
33	As a	general user	I want to	browse built-in ML models	Machine Learning	UI	so that	I can discover ways of manipulating my data	1	1 - Must have
39	As a	user	I want to	be able to select split ratio of data	Machine Learning	Data Manipulation	so that	to train and test the model	1	2 - Should have
26	As a	pro user	I want to	have the option to change the hyperparameters used in the machine learning model	Machine Learning	ML	so that	I can fine tune my test results.	3	2 - Should have
15	As a	pro user	I want to	be able use additional ML models	Machine Learning	ML	so that	I can improve accuracy	5	2 - Should have
7	As a	pro user	I want to	combine multiple models together	Machine Learning	ML	so that	I can model more complex data manipulations	5	3 - Could have
22	As a	general user	I want to	see clear annotation or explanation of data points and features	Data Visualisation	UI	so that	I can understand the results of the analysis	3	1 - Must have
8	As a	student	I want to	use different type of plotting graphs	Data Visualisation	UI	so that	I have flexibility to visualize data according to my needs.	5	1 - Must have
9	As a	general user	I want to	able to view and plot the results of the model	Data Visualisation	UI	so that	I could analysis and visualise the effects of the model	5	1 - Must have
12	As a	student	I want to	export my work to a Jupyter Notebook	Jupyter Notebook	Server	so that	I can extend my work beyond the capability of the application	5	2 - Should have
2	As a	general user	I want to	attach comments to workflow objects	Jupyter Notebook	UI	so that	I can document my work	5	3 - Could have
4	As a	Pro user	I want to	edit python code on the interface	Jupyter Notebook	UI	so that	I can have control how the ML algorithms works	5	3 - Could have
5	As a	Pro user	I want to	upload my own script (in python) if possible	Jupyter Notebook	Server	so that	I can extend the tool to support custom models and featurizers	5	3 - Could have
27	As a	pro user	I want to	be able to access new databases	External Data	Data Manipulation	so that	I can access additional data	3	2 - Should have
3	As a	Pro user	I want to	be able to add new datasets in the future	External Data	Data Manipulation	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
11	As a	student	I want to	analyze the relationship between different features		ML	so that	I can identify which features I need to select for my analysis	5	2 - Should have
16	As a	general user	I want to	add specific materials to the workflow for analysis		Data Manipulation	so that	compare the performance of the specific material my client or I choose with other material	3	3 - Could have

Version 1.4

Personas:

Alice: An design engineer from material consulting company

Bob: A project manager from material consulting company

User Story Table						
ID	As (Role)	I Want To	So That	Size Estimation	MoSCoW Priority	Justification
1	Alice / Bob	register an account	I can log in to the system with my own identity	medium	Must have	<i>Size estimation:</i> A small account management feature will work but will also involve maintenance. <i>Priority:</i> This is a prerequisite for other high priority functionalities. It is also required for security and privacy.
2	Alice / Bob	log in to the system to save or load the workflow	so that I can continue on unfinished job	small	Must have	<i>Size estimation:</i> A database or file management feature wil work. <i>Priority:</i> Save and load work can save time for workflows.
3	Alice	add specific materials to the workflow for analysis	compare the performance of the specific material my client or I choose with other material	small	Should have	<i>Size estimation:</i> We can add and merge the specific data with the main table. <i>Priority:</i> Users can also modify the main data table to accomplish this. This can be a PRO feature
4	Alice	create the machine learning workflow only by clicking on the web interface	I don't have to learn how to perform machine learning tasks in programming language	large	Must have	<i>Size estimation:</i> Requires a lot of low-level frontend, backend and machine learning jobs. <i>Priority:</i> This is the main feature of the application.
5	Alice / Bob	see hints for each workflow steps and explanation				

| Persona

Assigned to [Xinle Yu](#)

Version	Description	Editor	Date
1.0	First version based on initial understanding of the project and first client meeting	Xinle Yu	2022-8-19
1.1	Round off the sentences and adjust some description	Xinle Yu	2022-8-21
2.0	Combine three team work together	Yanan Liu	2022-8-22

Persona 1: Student (By: Boxjelly)

Elaine Kim

age: 22

residence: Melbourne

education: Material Engineering

occupation: Postgraduate student

marital status: Single

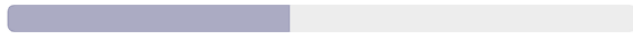


"It's SO time spending to do material researches and get decent results through just a semester."

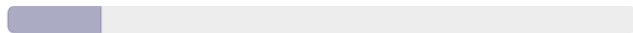
Motivation: As a material engineering graduate student, Assol gets frustrated and demotivated when she can't make sense of the data she has because she doesn't have a tool or sufficient programming/machine learning skills to process the material data. She is also frustrated that she can't use machine 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

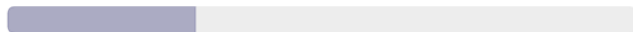
INTERNET



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

Values

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

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

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

• FLEXIBLE

- Understandable

• 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: Pro User (By: 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

Persona 3: Industry User (By: Bluering)

Gray Zhou

age: 28

residence: Ningde, Fujian, China

education: Master of Material Engineering

occupation: R & D Engineer of Polymer

marital status: Single

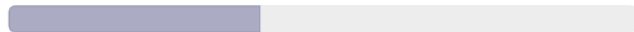


"It is fantastic to apply a multi-function online tool with ML methods in my work, but only after it shows its efficiency and reliability to me. Nobody will refuse a tool that can save his time"

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

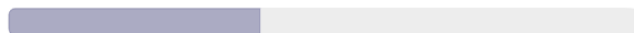
INTERNET



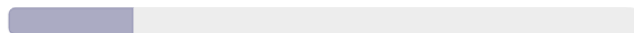
SOFTWARE



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

Values

- Easy to get started
- Efficient back-end process
- Abilities to select functions and filter results

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

Fears

- Frequently unavailability which may waste him much of time to retry
- Not enough guidance in the web or tool so him my

- Well organized visualization of interface and graphs

feel confused to find functions he wants.

- No choice of properties that he needs
- Lacking understand of what the system done, which may influence the confidence level of his report

Version 1.0



Version 1.1



Gray Zhou's persona v1.1.pdf

| Prototype

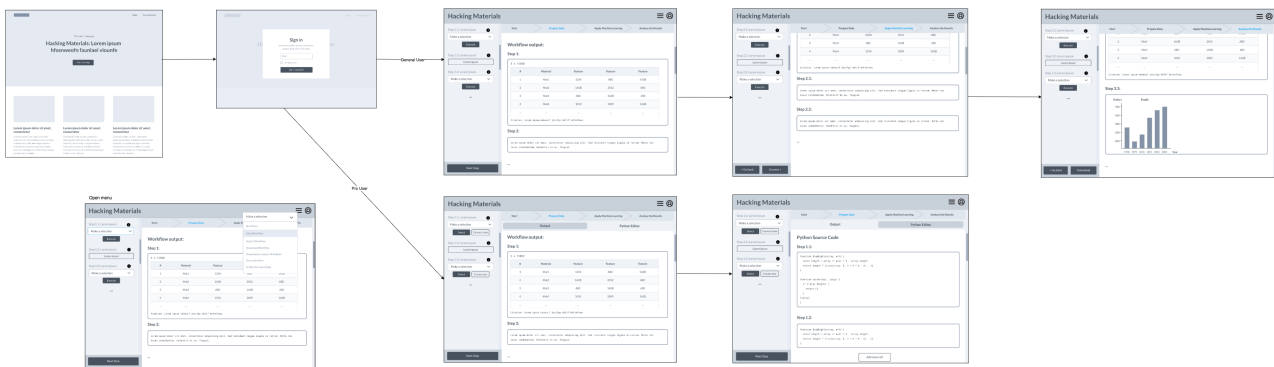
Assigned to [Yanan Liu](#)

Versions

Version ID	Description	Editor	Date
1.0	First version based on initial understanding of the project and client meeting Made in collaboration with team redback and boxjelly	Yanan Liu Ghina Yashar Felipe Leefu Huang Lin	2022-8-20

Group cooperation work, work in [marvel](#)

Version 1.0



Descriptive Notes recorded by team redback

Descriptive Notes

- 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.

| Business case

Three teams cooperation work original in [here](#), and formatted document can be downloaded here [HA Software Project - Cross-team collaboration - Business case.pdf](#)

COMP90082 Software Project SM2 2022

Project Hacking Materials Cross-Team Collaboration Model Proposal

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.

| Communication Archive

| Client

Name	Time	Tool	Detail

| Supervisor

Name	Time	Tool	Detail

| Resources

Github: <https://github.com/COMP90082-2022-SM2/HA-BlueRing>

Trello: <https://trello.com/b/Q1AZdIXb/sprint-1>

Matminer: https://nbviewer.org/github/hackingmaterials/matminer_examples/blob/main/matminer_examples/index.ipynb

LMS Project: https://canvas.lms.unimelb.edu.au/courses/126944/pages/hacking-materials-user-interface-code-ha?module_item_id=3996633

Dr. Brandl information: <https://findanexpert.unimelb.edu.au/profile/826808-christian-brandl>

Motivational Model: <https://momo-staging.eresearch.unimelb.edu.au/dashboard>

| Contacts

Teammates

Name	Email address	Student ID
Yanan Liu (Team rep)	yananl7@student.unimelb.edu.au	1289747
Hongpeil Lu	hongpeil@student.unimelb.edu.au	1275238
Jiahao Ju	jiahaoj1@student.unimelb.edu.au	1128182
Xinle Yu	xinley@student.unimelb.edu.au	1221568
Rui Zhang	zhangr.9@student.unimelb.edu.au	1294310

Client

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

Teaching Staff

Name	Email address
Mauro Mello Jr (supervisor)	mauro.mellojr@unimelb.edu.au
Eduardo Oliveira	eduardo.oliveira@unimelb.edu.au

Team RedBack

Name	Email address
Alastair Daivis (Team rep)	adaivis@student.unimelb.edu.au
Sanjeevani Avasthi	savasthi@student.unimelb.edu.au
Mamta Lopes	mlopes@student.unimelb.edu.au
Ghina Yashar	gyashar@student.unimelb.edu.au
Chunbaixue Yang	chunbaixue@student.unimelb.edu.au
Confluence Page: Home	

Team BoxJelly

Name	Email address
Zhaoqi Wang (Team rep)	zhaoqi@student.unimelb.edu.au
Dara O hEidhin	doheidhin@student.unimelb.edu.au
Yaoming Xuan	yaomingx@student.unimelb.edu.au
Radhimas Djan	djanr@student.unimelb.edu.au
Felipe Leefu Huang Li	fleefuhuangl@student.unimelb.edu.au
Confluence Page: Home	