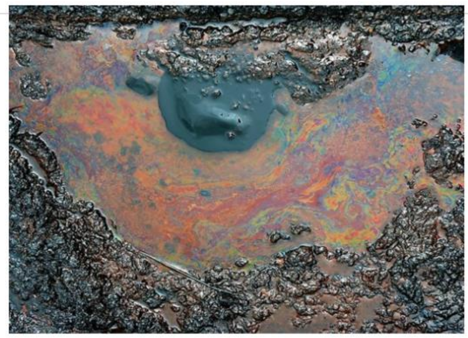
Phytoremediation Data Management Final Report



Group 9 Data Management - 21306
Etse Lossou
Isa Molewijk
Yanuar Bomantara
Tim Rangé
Jurrian Doornbos

Table of Contents

1 Introduction

1.1 Client Description

Founded in 2020, the Scape foundation, designer of urban spaces in Amsterdam and Rotterdam, is dedicated to enhancing the biodiversity and liveability of urban communities for all living things. To achieve this aim, the foundation connects researchers, designers, and developers to create sustainable and future-proof green ecosystems through open-source projects, research, data, and tools. In line with its mandate, the foundation in collaboration with researchers from the Wageningen University and Research embarked on a phytoremediation project.

The idea of phytoremediation is borne out of the fact that there are many polluted areas (soil and water) in Dutch cities (Figure 1) that are polluted with contaminants mostly because of previous industrial activities. Often, conventional remediation methods such as excavation, landfill, thermal treatment, leaching, and soil washing are very costly (Gerhardt et al., 2017; Petrová, Rezek, Soudek, & Vaněk, 2017), hence the need for an alternative method that considers the use of plants to clean up the contaminants from the polluted soil.

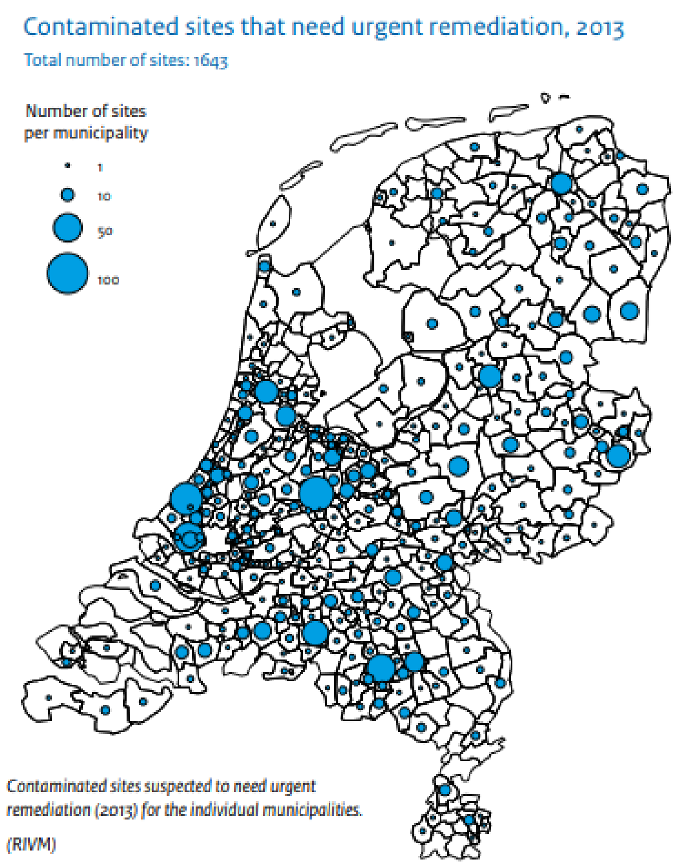


Figure 1. Map of polluted soils in the Netherlands

This method is called phytoremediation. As shown in Figure 2, there are six different phytoremediation techniques namely Phytostabilisation, phytostimulation, phytodegradation, phytovolatisation, phytoextraction and rhizofiltration. Each of these techniques is based on the type of contaminant involved (heavy metal, crude oil, chlorinated hydrocarbons, etc). Jeroen Schutt and Levi Biersteker, the two leadings Biologist researchers at the WUR are custodian of an extensive work and data on different plants that can be used to remediate polluted soils in the Netherlands. However, this data is presented in an excel format, which limit the potential of the data to be queried effectively.

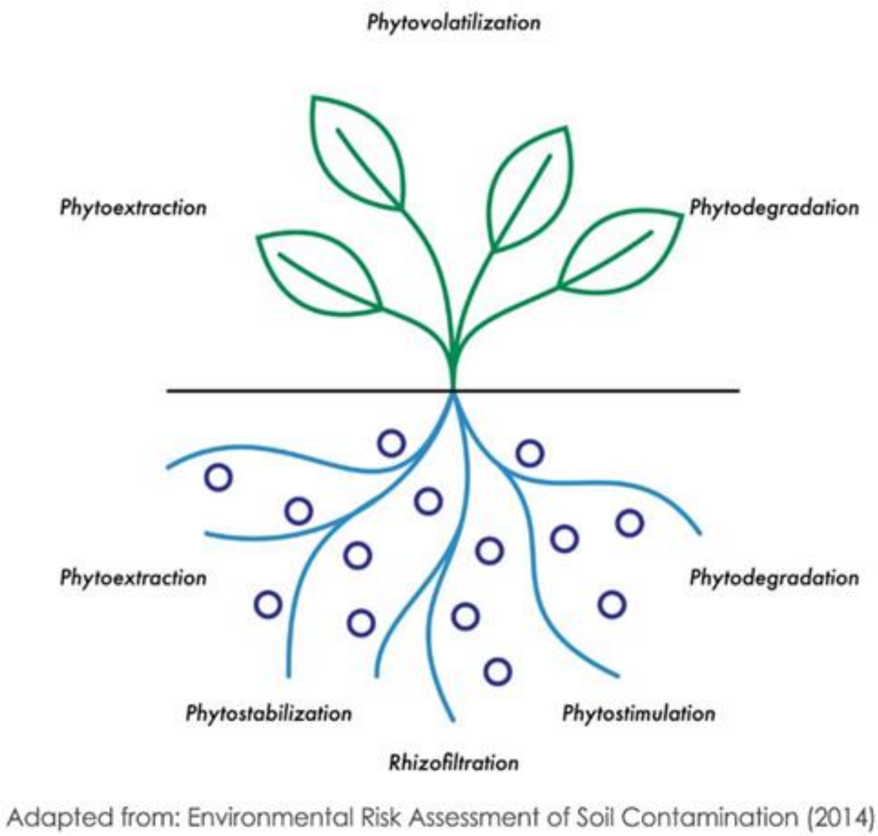


Figure 2. Types of phytoremediation

1.2 Client requests

Scape Foundation and WUR (the client) want to give urban planners, municipality officials, designers and, architects ideas to get rid of polluted soils in Dutch cities in a sustainable way. The client has available data that is currently in an excel format and they want a comprehensive phytoremediation database developed, that will show the relation between plants and between plant associations.

1.3 Objective Statement

The database must be reproducible and able to perform the following:

1. Provide detail information about plants species. The current data has close to 80 plants whereas about 66000 plants species have been identified. The database should therefore be created in such a way that it accepts addition plant species in future.
2. Determine which plant performs what type of phytoremediation based on a specific type of contaminant. This is a vital information for the client because plants can be selected based on

specific characteristics of the plants set by the user. Some of these characteristics are acidity, moisture content, salinity, etc.

3. Keep record of users of the database. The database should be able to retrieve information of users, team they belong to, projects they are working on and the location of the project.
4. The possibility of the database to be incorporated in a web interface where users can easily interact with the data. Considering the variety of users with different level of understanding of phytoremediation, this web interface will help urban planners and designers find which plant to use in a certain area, based on the chemical properties of the soil and the properties of the plant like colour, size and form. (**The client has their platform and intend to incorporate the database that will be created in this project).**

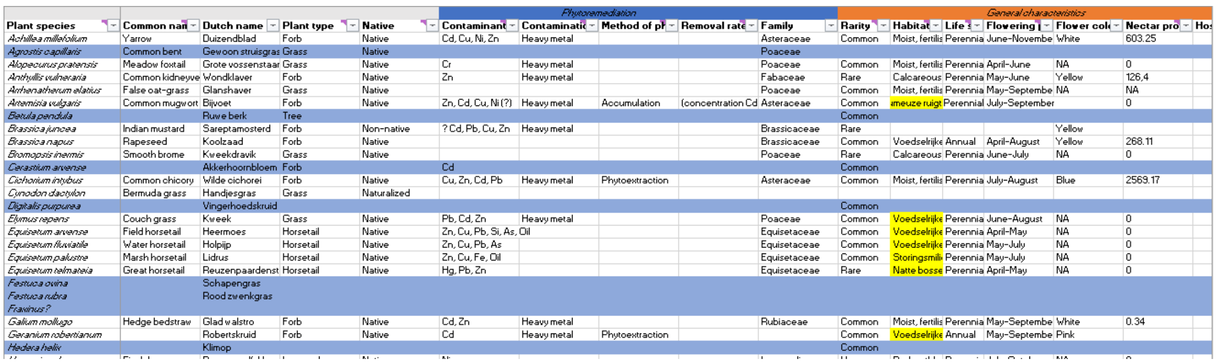
From the objectives listed above, and based on the information that will be made available in the database, the following queries can be performed:

1. What is the total number of plant species per contaminant type?
2. Which plants perform what type of phytoremediation based on a specific contaminant type?
3. What type of soil is fixable through a specific phytoremediation technique?
4. Which users are working on which projects? Where are the project sites located? Are the projects on-going?
5. What are the best plants specific based on a specific field report?

2 Materials and methods

2.1 Data description

We received an excel sheet produced by a Wageningen student team (Academic Consultancy Training project) consisting of three sheets of phytoremediation data (figure 3). More specifically, it was a list of plants by Latin name, common name, type of phytoremediation linked to the plant, plant characteristics and environmental conditions. The database we were asked to create should link contaminants in the soil to plants which can be used as a remedy. The file listed a total of 67 plants, which would be expanded in the future. There are a lot of null values in the file which would be filled in later.



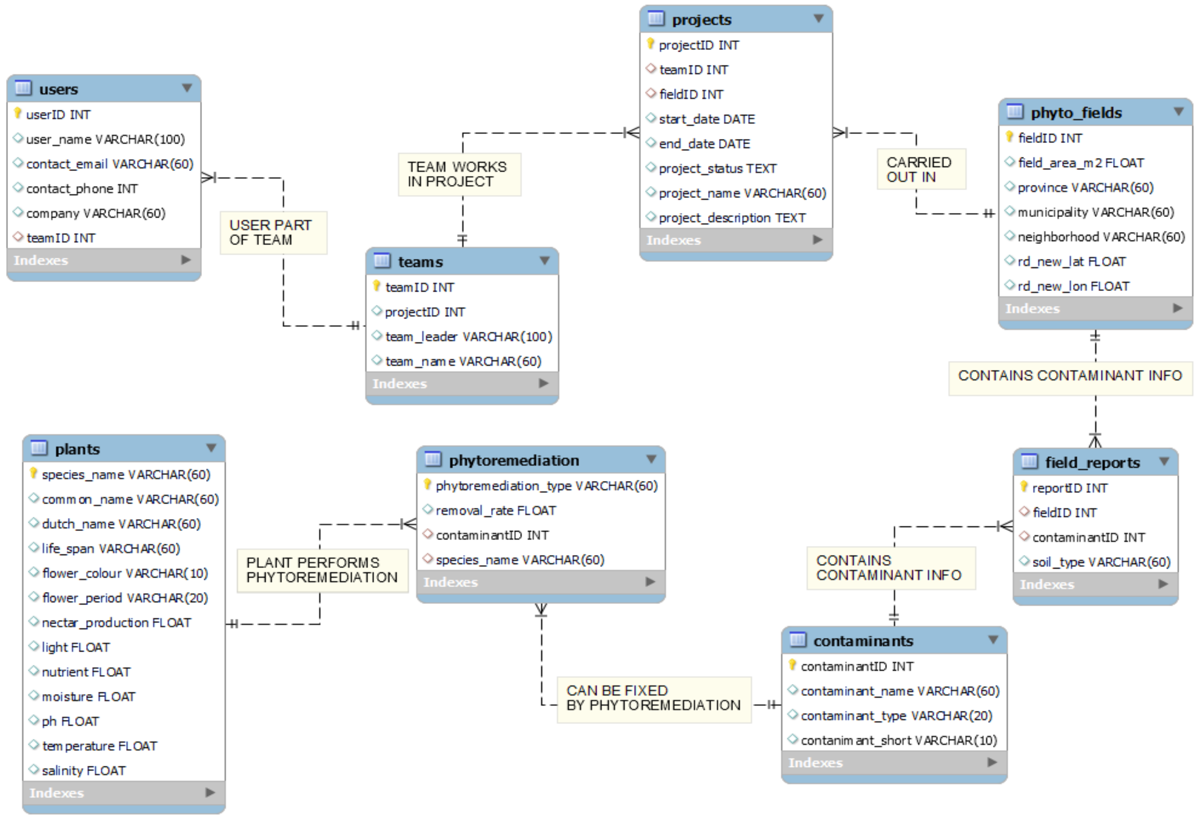
					Phytoremediation					General characteristics						
Plant species	Common nan	Dutch name	Plant type	Native	Contaminant	Contaminatic	Method of pł	Removal rate	Family	Rarity	Habitat	Life s	Flowering	Flower col	Nectar pro	Ho
Achillea millefolium	Yarrow	Duizendblad	Forb	Native	Cd, Cu, Ni, Zn	Heavy metal			Asteraceae	Common	Moist, fertilis	Perennia	June-Novembe	White	603.25	
Agrostis capillaris	Common bent	Gewoon struisgras	Grass	Native					Poaceae							
Alopecurus pratensis	Meadow foxtail	Grote vossenstaar	Grass	Native	Cr	Heavy metal			Poaceae	Common	Moist, fertilis	Perennia	April-June	NA	0	
Anthyllis vulneraria	Common kidneyve	Wondklaver	Forb	Native	Zn	Heavy metal			Fabaceae	Rare	Calcareous	Perennia	May-June	Yellow	126,4	
Arrhenatherum elatius	False oat-grass	Glanshaver	Grass	Native					Poaceae	Common	Moist, fertilis	Perennia	May-Septembe	NA	NA	
Artemisia vulgaris	Common mugwort	Bijvoet	Forb	Native	Zn, Cd, Cu, Ni (?)	Heavy metal	Accumulation	(concentration Cd	Asteraceae	Common	meuze ruigt	Perennial	July-September		0	
Betula pendula		Ruwe berk	Tree							Common						
Brassica juncea	Indian mustard	Sareptamosterd	Forb	Non-native	? Cd, Pb, Cu, Zn	Heavy metal			Brassicaceae	Rare				Yellow		
Brassica napus	Rapeseed	Koolzaad	Forb	Native					Brassicaceae	Common	Voedselrijke	Annual	April-August	Yellow	268.11	
Bromopsis inermis	Smooth brome	Kweekdravik	Grass	Native					Poaceae	Rare	Calcareous	Perennia	June-July	NA	0	
Cerastium arvense		Akkerhoornbloem	Forb		Cd					Common						
Cichorium intybus	Common chicory	Wilde cichorei	Forb	Native	Cu, Zn, Cd, Pb	Heavy metal	Phytoextraction		Asteraceae	Common	Moist, fertilis	Perennia	July-August	Blue	2569.17	
Cynodon dactylon	Bermuda grass	Handjesgras	Grass	Naturalized												
Digitalis purpurea		Vingerhoedskruid								Common						
Elymus repens	Couch grass	Kweek	Grass	Native	Pb, Cd, Zn	Heavy metal			Poaceae	Common	Voedselrijke	Perennia	June-August	NA	0	
Equisetum arvense	Field horsetail	Heermoes	Horsetail	Native	Zn, Cu, Pb, Si, As, Oil				Equisetaceae	Common	Voedselrijke	Perennia	April-May	NA	0	
Equisetum fluviatile	Water horsetail	Holpijp	Horsetail	Native	Zn, Cu, Pb, As				Equisetaceae	Common	Voedselrijke	Perennia	May-July	NA	0	
Equisetum palustre	Marsh horsetail	Lidrus	Horsetail	Native	Zn, Cu, Fe, Oil				Equisetaceae	Common	Storingsmili	Perennia	May-July	NA	0	
Equisetum telmateia	Great horsetail	Reuzenpaardenst	Horsetail	Native	Hg, Pb, Zn				Equisetaceae	Rare	Natte bosse	Perennia	April-May	NA	0	
Festuca ovina		Schapengras														
Festuca rubra		Rood zwenkgras														
Fraxinus?																
Galium mollugo	Hedge bedstraw	Glad walstro	Forb	Native	Cd, Zn	Heavy metal			Rubiaceae	Common	Moist, fertilis	Perennia	May-Septembe	White	0.34	
Geranium robertianum		Robertskruid	Forb	Native	Cd	Heavy metal	Phytoextraction			Common	Voedselrijke	Annual	May-Septembe	Pink		
Hedera helix		Klimop								Common						

Figure 3. Screenshot of phytoremediation data available in Excel

2.2 Preprocessing data

As the client asked us to look at the usability of the database with urban developers as an end-user in mind, we created tables linked to "user", "project", and "team". For this we used dummy data, to show how the database would interact if actual users would be linked to the project. The user would file a field report and through the contaminants in the soil would be linked to plants. As many plants could be linked to many contaminants, we created an in-between table ("phytoremediation") for the many-to-many relationship. This table was created manually with a row per relationship. As the original excel file had a lot of null values, we also created dummy data in the plant table. We did a lot of research on the specific plants to make it as accurate as possible, learning a lot about phytoremediation in the process. To create the relationships we created identifiers ID in tables, to serve as primary key and foreign key of other tables.

2.3 Entity Relationship model



2.4 Entities Analysis

In the end the entities are linked as a sort of train, which we would change if we had more time, for this project it made the most sense. The main entity is "plants" which has a many-to-many relationship with "contaminants" linked with the "phytoremediation" table. With more time the plant table could be split up in more different tables and other factors could be added like location or rarity. The top part of the ER-model starts with "users" with a many-to-one relationship with "teams" as many users could be in the same team (wish can also be shown through queries). Teams could work on many different projects. "Phyto_fields" contains locational data of the project and is linked to many "field_reports" which contains a contaminant ID. Therefore, a user submitting a field report can get plants recommended for their contaminated soil.

3 Results

The goal of the project is to develop a database which is able to suggest plants, based on a few parameters of the soil, such as contaminants, or plant requirements, such as required light and salinity of the soil for example. This has been achieved, the design of this database is presented in chapter 2.

To increase usability of this database for a wider audience, apart from software developers or database managers, this chapter will present three activities with the database. The first are some of the SQL queries, as created by the project group. The second result is the Jupyter Notebook, where the database is accessed, and a few example interactions between SQL and Python are explained. The third is a longer-term plan for the database and applications.

The client also noted that the database and queries were very good, and very useful to them. A version of the database will be hosted by them as soon as possible, for direct use.

3.1 Queries

To show the usability of the database, a variety of queries were created, these are presented in Appendix II. The main goal of these queries is to show that the database functions according to the ER-Diagram, and the important queries, as required by the client, are very much possible. This is the same as the SQL-report delivered in Week 6 of the project.

An important query, for the client, was a query to have contaminant information, or other location features, and it would return plants that could be used. Versions of this query are in the query appendix (1,2,6,7).

3.2 Jupyter notebook

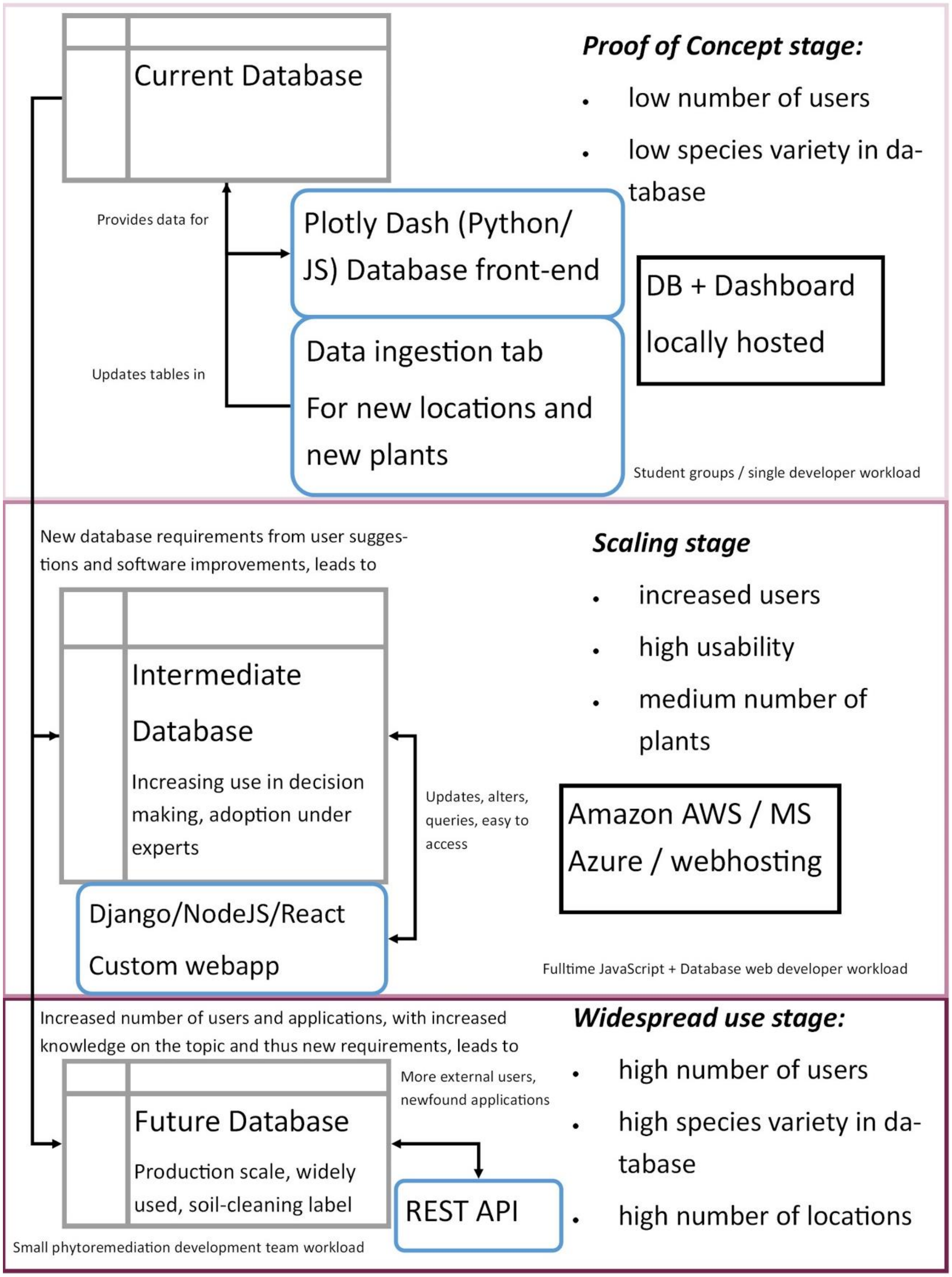
An important goal of the project was to make the database accessible to a variety of actors, even thought there was little time for us to develop a front-end for this project, we have spent time helping the next users of the database, by creating a Jupyter Notebook.

The Jupyter Notebook was created to help our direct commissioner Levi Biersteker with understanding SQL in a programming environment. The used SQL python-package is MySQL.

The notebook can be found under the /PR3_Notebook/ folder in the project folder. This folder contains a database creation script (.sql file), an SQLite version of the database (.db file), and the notebook itself.

3.3 Suggested longer-term plan

The main commissioner, Lars van Vianen suggested that we create a longer-term implementation and pathway to a larger, and easier to use version of the database, our perspective on such a pathway is presented in the diagram below.



The diagram is read from top to bottom, and is separated into three stages, the Proof of Concept Stage, the Scaling Stage and Widespread Use stage. Each stage has a new version of a database, and ways to access and interact with the database.

The Proof of Concept stage is characterized by a low number of users, and not that many plants in the database. In this PoC stage, the current database is used as a backend for a Plotly Dash dashboard. This dashboard shows the data, by selecting contaminants that are present in the soil, a design of this is already shown in the Introduction. Furthermore, a data-ingestion tab, for new plant species and field locations is present for urban planners and designers to interact with the database. The database is locally hosted.

The Scaling stage is characterized by an increase in number of users, due to an understanding of the need of such a database. Lessons from the previous front-end and back-end database are considered when designing this new version. More features are added, in the form remediation-labels (whether it is a good choice to plant or not), a lot more plants are added, with more detailed information, as there is more understanding of how phytoremediation works. The front-end is now being redeveloped using a JavaScript web framework, resulting in a more intuitive design, with more features to interact, query, and keep track of your own contaminated fields over time. Interaction between users is also increasing, and exploratory searches in the database are possible, to see the best-case scenarios and lessons from other phytoremediation projects. This database and webapp is now professionally hosted at a webhosting service, such as Amazon AWS.

The Widespread use stage is characterized by new applications and non-professional users accessing the database, for internal use by plant-stores, home-users and other plant-tracking apps or even large DIY-stores, to label their plants according to phytoremediation categories. This requires a REST-API for the database, so that a high number of users can access the database for their own use and applications. The webapp created in the previous stage is still accessible but holds its own niche in the phytoremediation field. The database itself is an accurate representation of all the plants and interactions with the soil these plants have, users have methods and ways to store their data of large and small projects using phytoremediation.

4 Conclusions

This project created a newly designed and functional phytoremediation database for Scape Foundation and WUR. With this database, the client can help urban planners and other users select the best plants from the options available to them during their phytoremediation projects.

This database is also capable of managing and storing information of different sites and projects where phytoremediation needs to occur. From the database created, queries were generated to answer specific research questions that the client wanted to be solved.

In consultation with the client, specific entities were created, and relationships established. The database is created in such a way that there is room for scaling-up since the project mentioned there are over 66000 plants that would potentially be added to the database. The scaling-up can only be carried out by authorized access to ensure data integrity.

From the phytoremediation database, plants species can be queried with ease based on their specific characteristics. The core objective of this database is to identify plants that can identify a given contaminant in a polluted soil. This objective was achieved and was demonstrated in the queries that were generated in this project (Queries 1, 2, 6 and 7).

Also, from the database, users can be identified, projects they are working on and where the projects are taking place can also be queried. This type of query was developed by the team working on this project because it did not form part of the original data.

Furthermore, the client expressed an interest in getting advice from the group on future development of this database and methods of interaction with this database. This advice is presented in the form of a diagram, this diagram shows a possible timeline that shows the future of a front-end and the database itself. Within this diagram, the database presented in this work is the starting-point and proof of concept.

Finally, to make the database as useful as possible to the client, a Jupyter notebook was developed for the client with programming experience, but little SQL knowledge. This Jupiter notebook serves to bridge the gap between the current database and potential future implementations, as shown in the diagram of 3.3.

5 Recommendations

Next to the results and conclusions, some recommendations for the database are presented below.

- Add where the toxins go in the plant as a new table, joining phytoremediation type with this new toxin_location table, this requires some more detailed information on the relationship between phytoremediation and the toxin location.
- Since phytoremediation is location-based, geospatial calculation on the database could be included in the field report. As such, pollution level (score card?) could be visualized in the interface. This will require the use of Postgres because it supports many GIS functionalities like finding nearest neighbour, distance calculation from one point to another, etc.
- In future, the plant table can be split into multiple tables (area characteristics, plant requirements)

References

Gerhardt, K. E., Gerwing, P. D., & Greenberg, B. M. (2017). Opinion: Taking phytoremediation from proven technology to accepted practice. Plant science, 256, 170-185.

Petrová, Š., Rezek, J., Soudek, P., & Vaněk, T. (2017). Preliminary study of phytoremediation of brownfield soil contaminated by PAHs. Science of the Total Environment, 599, 572-580.

Rijkswaterstaat. (2014). Into Dutch Soils. A swift overview of how the Dutch manage their soils: a source of inspiration for your own practice.

Appendix

I: Data Dictionary

User			
<u>UserID</u>	Primary Key	INT	Unique identification number for in the database
User_name		VARCHAR(60)	Name of the user
Contact_email		VARCHAR(60)	Email address of the user
Contact_phone		INT	Phone number of the user
Company		VARCHAR(60)	Company where the user works at
Role		*VARCHAR(60)*	Job title of user, Urban designer, or manager, or developer

Team			
<u>TeamID</u>	Primary Key	INT	Unique identification number for in the database
ProjectID	Foreign Key	INT	Foreign key of the project
UserID	Foreign Key	INT	Foreign key of the user
Team_name		VARCHAR(60)	The name that the team decided to use
Team_leader		VARCHAR(60)	Leader of the team, could be a user

Project			
<u>ProjectID</u>	Primary Key	INT	Unique identification number for in the database
FieldID	Foreign Key	VARCHAR(60)	Foreign key of the field used in the project
TeamID	Foreign Key	VARCHAR(60)	Foreign key of the team being part of the project
Start_date		DATE	Date that the project started
End_date		DATE	Date that the project ended
Status		TEXT	Status of the project, a small description of how the project is going
Project_name		VARCHAR(60)	Name of the project, such as the name of the field that is being phytoremediated

Phyto_fields			
<u>FieldID</u>	Primary Key	INT	Unique identification number for in the database
Field_area		FLOAT	Size of the area that is being remediated
Province		VARCHAR(60)	Province where the field is
Municipality		VARCHAR(60)	Municipality where the field is
Neighborhood		VARCHAR(60)	Neighbourhood where the field is

RD_new_lat	FLOAT	Point location Y in the RD_new coordinate system
RD_new_lon	FLOAT	Point location X in the RD_new coordinate system

Reports			
ReportID	Primary Key	INT	Unique identification number for in the database
FieldID	Foreign Key	INT	Foreign key of the field
ContaminantID	Foreign Key	INT	Foreign key of the contaminant

Contaminant			
ContaminantID	Primary Key	INT	Unique identification number for in the database
Contaminant_name		VARCHAR(60)	Long name of the contaminant
Contaminant_type		VARCHAR(60)	How the contaminant is classified, such as heavy metal or crude oil
Contaminant_short		VARCHAR(10)	The chemical notation of the contaminant

Phytoremediation			
Phytoremediation_type	Primary Key	VARCHAR(60)	Unique identification name of the phytoremediation type
Removal_rate		FLOAT	[Content removed/year] that this type of phytoremediation will achieve
ContaminantID	Foreign Key	INT	Foreign key of the contaminant that can be removed with this phytoremediation type
Species_name	Foreign Key	VARCHAR(60)	Foreign key of the plant that uses this phytoremediation type

Plants			
Species_name	Primary Key	VARCHAR(60)	Unique Latin Name of the plant
Common_name		VARCHAR(60)	Common name of the plant, in English
Dutch_name		VARCHAR(60)	The name of the plant, in Dutch
Life_span		VARCHAR(30)	Months that the plant grows
Flower_colour		VARCHAR(10)	Colour of the flower on the plant
Nectar_pro		FLOAT	Nectar production of the plant, in
Light		INT	Light requires for the plant, Ellenberg Indicator (range: 1-9)
Nutrient		FLOAT	Nutrient intake of the plant, Ellenberg Indicator (range: 1-9)
Moisture		FLOAT	Moisture content of the soil required, Ellenberg Indicator (range: 1-9)
pH		FLOAT	pH of the soil required in pH, Ellenberg Indicator (range: 1-9)

Temp	FLOAT	Temperature for the plant, Ellenberg Indicator (range: 1-9)
Salinity	FLOAT	Salinity of the soil, Ellenberg Indicator (range: 1-9)

II: SQL queries

Query 1: Selecting the best plants, given a field report

This query returns the report of a field. This contains the full information about the plant, the type of soil, the contaminant and the phytoremediation.

```
SELECT field_reports.soil_type, contaminants.contaminant_name,
contaminants.contaminant_type, plants.species_name, plants.dutch_name,
plants.life_span, plants.flower_colour, plants.flower_period,

phytoremediation.removal_rate, phytoremediation.phytoremediation_type

FROM field_reports

LEFT JOIN contaminants ON contaminants.contaminantID =
field_reports.contaminantID

LEFT JOIN phytoremediation ON phytoremediation.contaminantID =
field_reports.contaminantID

LEFT JOIN plants ON phytoremediation.species_name = plants.species_name

WHERE field_reports.reportID = 2;
```

reportID	fieldID	contaminantID	soil_type	contaminantID	contaminant_name	contaminant_type	contaminant_short	phytoremediation_ID	phytoremediation_type	removal_rate	contaminantID	species_na
2	2	15	sand	15	Selenium	Heavy metal	Se	62	Accumulation	61	15	Schedonoru
2	2	15	sand	15	Selenium	Heavy metal	Se	66	Rhizodegradation (DDE)	65	15	Lotus cornic

Query 2: Selecting plants, and some features, given soil pollutants and plant requirements

As shown in the report below, the query returns all plants that are associated with specific contaminants and the type of phytoremediation they perform.

```
SELECT plants.species_name, plants.dutch_name, plants.life_span,
plants.flower_colour, plants.flower_period, contaminants.contaminant_type,
contaminants.contaminant_short AS chemical, phytoremediation.removal_rate,
phytoremediation.phytoremediation_type

FROM plants

JOIN phytoremediation ON plants.species_name = phytoremediation.species_name

JOIN contaminants ON phytoremediation.contaminantID =
phytoremediation.contaminantID

WHERE contaminant_name = "Cobalt" AND plants.light > 5 AND plants.ph > 6 #
change these for different plants and types

GROUP BY plants.species_name, phytoremediation.phytoremediation_type;
```

species_name	dutch_name	life_span	flower_colour	flower_period	contaminant_type	contaminant_short	removal_rate	phytoremediation_type
Cerastium arvense	Akkerhoornbloem	Annual	NA	June-July	Heavy metal	Co	1	Phytoextraction
Salix viminalis	Katwilg	Perennial	NA	May-June	Heavy metal	Co	12	Phytostabilization
Taraxacum officinale	Gewone paardenbloem	Perennial	Yellow	April-June	Heavy metal	Co	13	Phytostabilization
Brassica juncea	Sareptamosterd	Annual	Yellow	July-October	Heavy metal	Co	14	Phytostabilization
Brassica juncea	Sareptamosterd	Annual	Yellow	July-October	Heavy metal	Co	16	Phytoextraction
Plantago major	Grote weegbree	Perennial	NA	May-November	Heavy metal	Co	22	Phytoextraction
Plantago major	Grote weegbree	Perennial	NA	May-November	Heavy metal	Co	25	phytostabilization
Silene vulgaris	Blaassilene	Perennial	White	July-September	Heavy metal	Co	35	phytostabilization
Salsola kali	Stekend loogkruid	Annual	Pink	July-September	Heavy metal	Co	4	Phytoextraction
Cichorium intybus	Wilde cichorei	Perennial	Blue	July-August	Heavy metal	Co	45	phytostabilization
Equisetum telmateia	Reuzenpaardenstaart	Perennial	NA	April-May	Heavy metal	Co	49	phytostabilization
Elymus repens	Kweek	Perennial	NA	June-August	Heavy metal	Co	55	Rhizodegradation (DDE)
Elymus repens	Kweek	Perennial	NA	June-August	Heavy metal	Co	56	Accumulation
Alopecurus pratensis	Grote vossenstaart	Perennial	NA	April-June	Heavy metal	Co	7	Phytostabilization
Medicago sativa	Luzerne	Perennial	Purple	June-September	Heavy metal	Co	71	Rhizodegradation (DDE)
Artemisia vulgaris	Bijvoet	Perennial	NA	July-September	Heavy metal	Co	73	Accumulation
Pteridium aquilinum	Adelaarsvarn	Perennial	NA	July-August	Heavy metal	Co	79	Accumulation
Pteridium aquilinum	Adelaarsvarn	Perennial	NA	July-August	Heavy metal	Co	83	Rhizodegradation (DDE)
Equisetum arvense	Heermoes	Perennial	NA	April-May	Heavy metal	Co	92	Rhizodegradation (DDE)
Equisetum arvense	Heermoes	Perennial	NA	April-May	Heavy metal	Co	94	Phytoextraction

Query 3: Number of plants per contaminant type

This query provides a summary of all the plants that are available in the database per contaminant type.

```
SELECT COUNT(species_name) AS amount_of_plants, contaminant_name FROM
phytoremediation

LEFT JOIN contaminants

ON phytoremediation.contaminantID = contaminants.contaminantID

GROUP BY contaminant_name

ORDER BY amount_of_plants DESC;
```

amount_of_plants	contaminant_name
21	Zinc
18	Cadmium
13	Copper
13	Lead
6	Nickel
4	Arsenic
2	Chromium
2	Oil
2	pesticide
2	Petroleum
2	Selenium
2	Cobalt
2	Alamenium
2	Iron
1	Strontium
1	Mercury
1	Manganese
1	hydrocarbons
1	Mineral Oil
1	Silicon

Query 4: how who is working in the North Netherlands (latitude higher than 500000)?

In case there is the need to return all the users working in a specific area based on just their knowledge of latitude or longitude, this query is the ideal example.

```
SELECT user_name, team_name, project_name, province
FROM users
LEFT JOIN teams
ON users.teamID = teams.teamID
LEFT JOIN projects
ON teams.teamID = projects.teamID
LEFT JOIN phyto_fields
ON projects.fieldID = phyto_fields.fieldID
WHERE rd_new_lat > 500000;
```

user_name	team_name	project_name	province
Cornelis	Gamma	Redwood	Groningen
Pieter	Gamma	Redwood	Groningen
Zacharias	Gamma	Redwood	Groningen
Gerard	Etha	Bay Garden	Friesland
Teunis	Etha	Bay Garden	Friesland

Query 5: Select all users researching a field in Gelderland

This code returns all the users that are doing research in Gelderland. To be better understood, this code can include more columns.

```
SELECT user_name FROM users WHERE teamID in (
	SELECT teamID FROM teams WHERE projectID in (
		SELECT projectID from projects WHERE fieldID in (
			SELECT fieldID FROM phyto_fields WHERE province = 'Gelderland')));
```

	user_name
▶	Anton
	Dirk
	Hendrik
	Victor

Query 6: type of polluted soil fixable through for phytoremediation

This query retrieves the type of polluted soil from which the plant removes the pollutant through phytoremediation. Since the contaminant dummy data in the field_report does not have all the contaminant types, the query returns null values for some of the plant species. Thus, it is important to return only values where soil_type is not null.

```
SELECT species_name, contaminant_name AS Contaminant, contaminants.contaminant_short
AS Chemical_symbol, phytoremediation.phytoremediation_type AS Phytoremediation,
soil_type AS Soil from phytoremediation

        LEFT JOIN contaminants

        ON phytoremediation.contaminantID = contaminants.contaminantID

        LEFT JOIN field_reports

        ON contaminants.contaminantID = field_reports.contaminantID

        WHERE soil_type IS NOT NUL;
```

species_name	Contaminant	Chemical_symbol	Phytoremediation	Soil
Plantago major	Alamenium	Al	phytostabilization	sand
Lycopodium clavatum	Alamenium	Al	phytostabilization	sand
Schedonorus arundinaceus	Selenium	Se	Accumulation	sand
Lotus corniculatus	Selenium	Se	Rhizodegradation (DDE)	sand
Lolium perenne	pesticide	p,p'-DDE	Rhizodegradation (DDE)	mineral
Medicago sativa	pesticide	p,p'-DDE	Rhizodegradation (DDE)	mineral
Salix viminalis	Mineral Oil	Mineral oi	Phytostabilization	sand
Rumex acetosa	Copper	Cu	Phytostabilization	mineral
Brassica juncea	Copper	Cu	Phytoextraction	mineral
Achillea millefolium	Copper	Cu	Phytoextraction	mineral
Plantago major	Copper	Cu	phytostabilization	mineral
Lycopodium clavatum	Copper	Cu	phytostabilization	mineral
Cichorium intybus	Copper	Cu	phytostabilization	mineral
Mentha arvensis	Copper	Cu	Rhizodegradation (DDE)	mineral
Artemisia vulgaris	Copper	Cu	Accumulation	mineral
Osmunda regalis	Copper	Cu	Accumulation	mineral
Pteridium aquilinum	Copper	Cu	Accumulation	mineral
Equisetum palustre	Copper	Cu	Rhizodegradation (DDE)	mineral

Query 7: Identify which plants perform what type of phytoremediation based on the contaminant

This query identifies all the plants that perform phytoremediation based on a specific type of contaminant. Changing the contaminant type in the code will produce a diverse set of plants.

```
SELECT plants.species_name AS Plant, plants.dutch_name AS Plant_in_Dutch,
plants.life_span, contaminants.contaminant_type, contaminant_name,
phytoremediation.phytoremediation_type AS Phytoremediation  FROM
phytoremediation

LEFT JOIN plants

ON phytoremediation.species_name = plants.species_name

LEFT JOIN contaminants

ON contaminants.contaminantID = phytoremediation.contaminantID

WHERE contaminant_type = "crude oil"; # Depending on the contaminant type
```

Plant	Plant_in_Dutch	life_span	contaminant_type	contaminant_name	Phytoremediation
Equisetum palustre	Lidrus	Perennial	Crude oil	Oil	phytostabilization
Equisetum arvense	Heermoes	Perennial	Crude oil	Oil	Phytoextraction
Lolium perenne	Engels raaigrass	Perennial	Crude oil	Petroleum	Accumulation
Medicago sativa	Luzerne	Perennial	Crude oil	Petroleum	Rhizodegradation (DDE)
Lotus corniculatus	Gewone rolklaver	Perennial	Crude oil	hydrocarbons	Accumulation

Query 8: Which users work on which project?

This snippet of code checks which users are working on which projects. It also returns the status of the project. It is observed from the dummy data that While some users belong to more than one projects, other users do not have a project.

```
SELECT user_name as Client, contact_email as email, company as Company,
province, municipality, team_name AS Team, team_leader as Leader, project_name
as Project, project_status as Project_status from users, teams, projects,
phyto_fields

WHERE users.teamID = teams.teamID

AND teams.teamID = projects.teamID

AND projects.projectID = phyto_fields.fieldID

ORDER by client ASC;

## OR ##

SELECT user_name AS Client, contact_email as email, company AS Company,
province, municipality, team_name AS Team, team_leader as Leader, project_name
AS Project, project_status AS Project_status

FROM users

LEFT JOIN teams

ON users.teamID = teams.teamID

LEFT JOIN projects

on teams.teamID = projects.teamID

LEFT JOIN phyto_fields

ON projects.fieldID = phyto_fields.fieldID

WHERE Project_status = "Finished" OR Project_status = "Progress"

ORDER BY client ASC;
```

Client	email	Company	province	municipality	Team	Leader	Project	Project_status
Anton	anton@mail.nl	Andromeda	Noord-Holland	Amsterdam	Alfa	Anton	Blue Sky	Finished
Anton	anton@mail.nl	Andromeda	Gelderland	Wageningen	Alfa	Anton	Hemisphere	Finished
Cornelis	cornelis@mail.nl	Backward	Groningen	Groningen	Gamma	Cornelis	Redwood	Finished
Dirk	dirk@mail.nl	Bodes	Gelderland	Wageningen	Alfa	Anton	Hemisphere	Finished
Dirk	dirk@mail.nl	Bodes	Noord-Holland	Amsterdam	Alfa	Anton	Blue Sky	Finished
Ferdinand	ferdinand@mail.nl	Circinus	Flevoland	Almere	Zeta	Ferdinand	Sky Garden	Finished
Gerard	gerard@mail.nl	Coma Pinwheel	Friesland	Sneek	Etha	Teunis	Bay Garden	Finished
Hendrik	hendrik@mail.nl	Cosmos Redshift	Gelderland	Wageningen	Alfa	Anton	Hemisphere	Finished
Hendrik	hendrik@mail.nl	Cosmos Redshift	Noord-Holland	Amsterdam	Alfa	Anton	Blue Sky	Finished
Isaak	isaak@mail.nl	Eye of Sauron	Noord-Holland	Amsterdam	Beta	Isaak	Utara	Finished
Isaak	isaak@mail.nl	Eye of Sauron	Zuid-Holland	Den-Haag	Beta	Isaak	Samara	Finished
Lodewijk	lodewijk@mail.nl	Malin	Zuid-Holland	Den-Haag	Beta	Isaak	Samara	Finished
Lodewijk	lodewijk@mail.nl	Malin	Noord-Holland	Amsterdam	Beta	Isaak	Utara	Finished
Maria	maria@mail.nl	Medusa	Zuid-Holland	Den-Haag	Beta	Isaak	Samara	Finished
Maria	maria@mail.nl	Medusa	Noord-Holland	Amsterdam	Beta	Isaak	Utara	Finished
Nico	nico@mail.nl	Sculptur	Zuid-Holland	Rotterdam	Epsilon	Rudolf	Grey Bay	Progress
Otto	otto@mail.nl	Magellanic	Flevoland	Almere	Zeta	Ferdinand	Sky Garden	Finished
Pieter	pieter@mail.nl	Mayalls	Groningen	Groningen	Gamma	Cornelis	Redwood	Finished