Carnegie Mellon University
Heinz College

chatham UNIVERSITY

Data Management System Implementation Project

Final Presentation

Team Introduction

Jeeranat Sithivaraporn - *Project Manager*

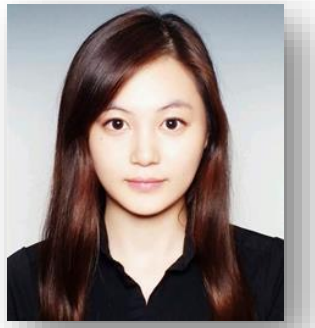
MISM

Eleanor Kenny - *Financial Manager*

MISM

Daniel Ohara

MISM



Min-Ji Huh

MISM – BIDA

Anshu Agrawal

MISM – BIDA

Agenda

Carnegie Mellon University
Heinz College

1. Overview and Scope
2. Front-end
3. Back-End and Demo
4. Documentation
5. Challenges and Future Work

chatham UNIVERSITY

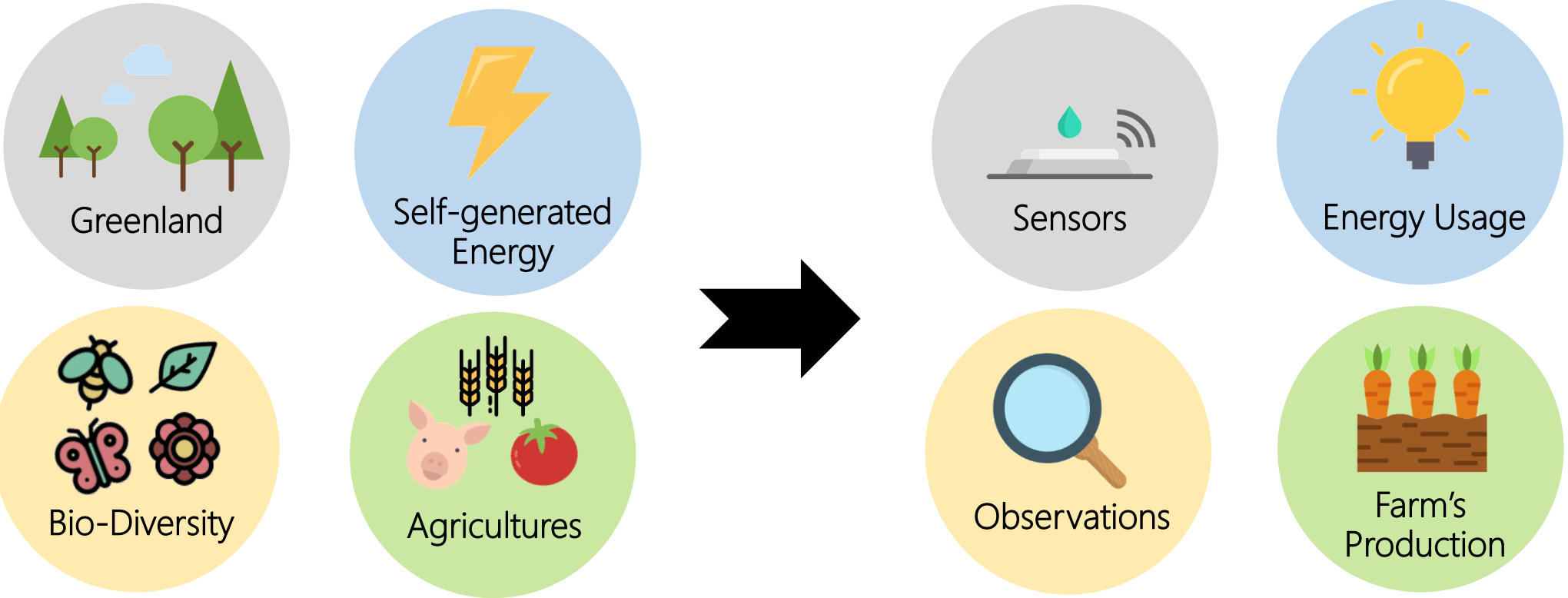
Carnegie Mellon University
Heinz College

1

OVERVIEW AND SCOPE

Eden Hall Campus
Falk School of Sustainability

Carnegie Mellon University
Heinz College



"Eden Hall Campus is seeking solutions to manage and centralize the growing amount of data, and provide data access to researchers and students."

chatham UNIVERSITY

Complications

Carnegie Mellon University
Heinz College

No data management system

- Data is scattered in non-uniform platforms
- No centralized data access point

Variability of the data

- Different sources
- Different formats
- Different collection processes/intervals

Limited technical resources

- Development and maintenance is difficult
- No user-friendly interface

Objective and Deliverables

Carnegie Mellon University
Heinz College

Objective

The Heinz College will work with Chatham University Eden Hall Campus to design and implement a centralized and scalable information management system for faculties and students to store and analyze the data.

Deliverables

Front-end
(Web Portal)

Back-end
(Data Management System)

Documentation

Functionality Scope

Carnegie Mellon University Heinz College

Functionalities	Use cases	Faculty	Granted Student	Others
Authorization and Authentication	Log-in	✓	✓	-
	Create Access	✓	-	-
	Manage Access	✓	-	-
Create New Tables	Add new data stream	✓	-	-
	Add Validation Rules	✓	-	-
	Manage Validation Rules	✓	-	-
Upload	Batch Files	✓	✓	-
	Observation Records	✓	✓	-
Download/ Query	Query Observation Data	✓	✓	✓
	Query By Type	✓	✓	✓
	Query By Site	✓	✓	✓
	Preview Data	✓	✓	✓
	Download as CSV	✓	✓	✓

chatham UNIVERSITY

2

Carnegie Mellon University
Heinz College

FRONT-END

Site Map

Main

Jeeranat

localhost:8080/InterestingPicture1/index2.jsp

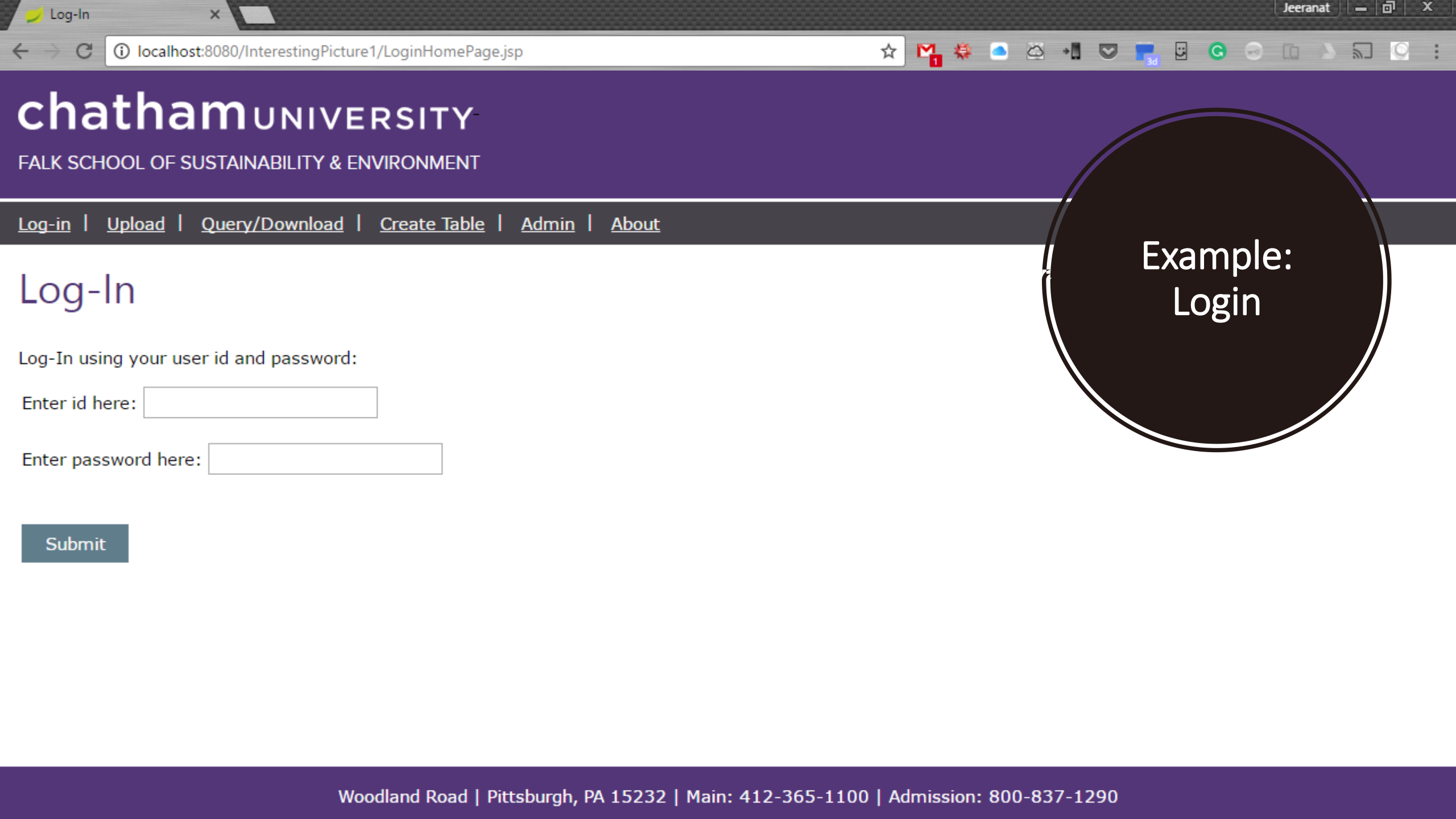
chatham UNIVERSITY

FALK SCHOOL OF SUSTAINABILITY & ENVIRONMENT

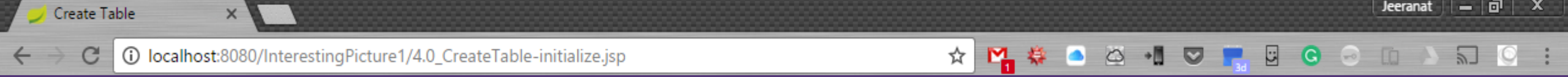
Log-in | Upload | Query/Download | Create Table | Admin | About

Example:
Main Page

Woodland Road | Pittsburgh, PA 15232 | Main: 412-365-1100 | Admission: 800-837-1290



Example: Login



chatham UNIVERSITY

FALK SCHOOL OF SUSTAINABILITY & ENVIRONMENT

Log-in | Upload | Query/Download | Create Table | Admin | About

Example: Create Table

Create Table

Welcome to create table page. Please fill in the fields below to create your table.

Number of Columns:

Type of Delimeter:

Comment Line Signifier:

Number of header lines to skip:

Data generation frequency:

Type of Sensor:

Site Prefix:

Sensor Model:

Sensor Manufacturer:

Location:

Next

Upload

localhost:8080/Capstone/2.5.1_Upload-Epidoptera.jsp

chatham UNIVERSITY

FALK SCHOOL OF SUSTAINABILITY & ENVIRONMENT

Log-in | Upload | Query/Download | Create Table | Admin | About

Example: Upload

Epidoptera(Moth) Upload Observation.

Fill in appropriate observation form.

Genus:

Species:

Common Name:

Family: Arctiidae

Sex: Unknown

Time of Sampling: Morning

Date of Reference Image: mm/dd/yyyy

Identified by:

Native Status: Unknown

Collected and Pinned: Yes

Submit

Query/Download

localhost:8080/Capstone/3.1_Query-ChooseType.jsp

chatham UNIVERSITY

FALK SCHOOL OF SUSTAINABILITY & ENVIRONMENT

Log-in | Upload | Query/Download | Create Table | Admin | About

Example: Query/ Download

Query Type.

Choose a type of query you wish to make

Please choose one of the following options.

- (•) By Site
- ( ) By Sensor
- ( ) Observations

Next Before

chatham UNIVERSITY

3

Carnegie Mellon University
Heinz College

BACK-END

Solutions

Carnegie Mellon University
Heinz College

No data management system

Variability of the data

1. Data Storage

- ✓ Table Mapping
- ✓ Column Mapping
- ✓ Rules Mapping

2. Data Engineering

- ✓ Validation Rules
- ✓ ETL Model

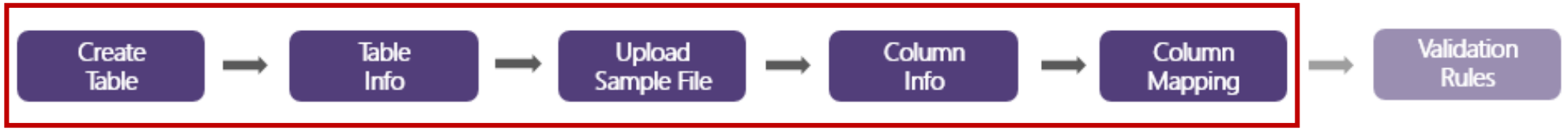
3. Data Aggregation

- ✓ Horizontal Join
- ✓ Vertical Join

1. Data Storage

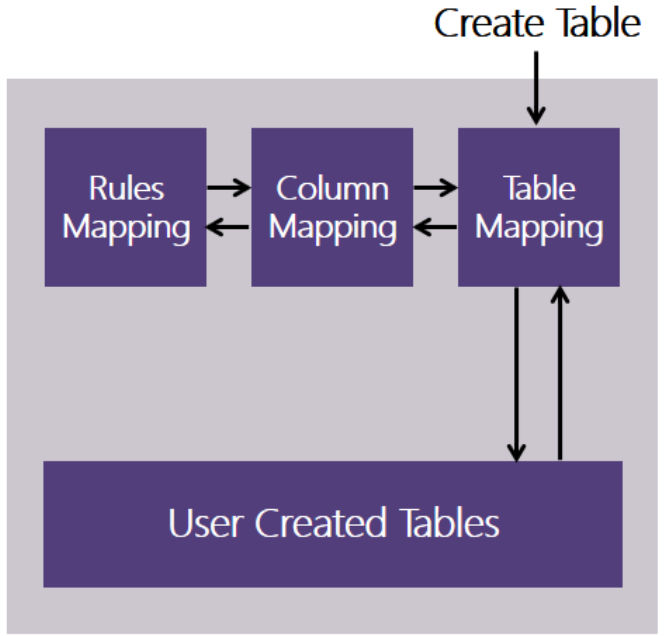
Carnegie Mellon University
Heinz College

FRONT



BACK

Data Storage Solution manages variety and scalability of the data using **Table Mapping**, and **Column Mapping** tables. The metadata of all tables are stored in these 2 tables that provide the mapping between tables. Therefore, tables of any size and format can be joined together in the future.

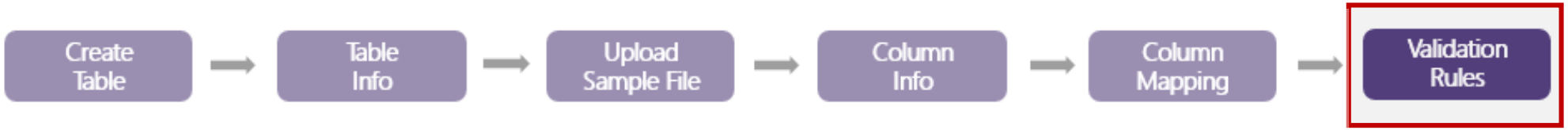


chatham UNIVERSITY

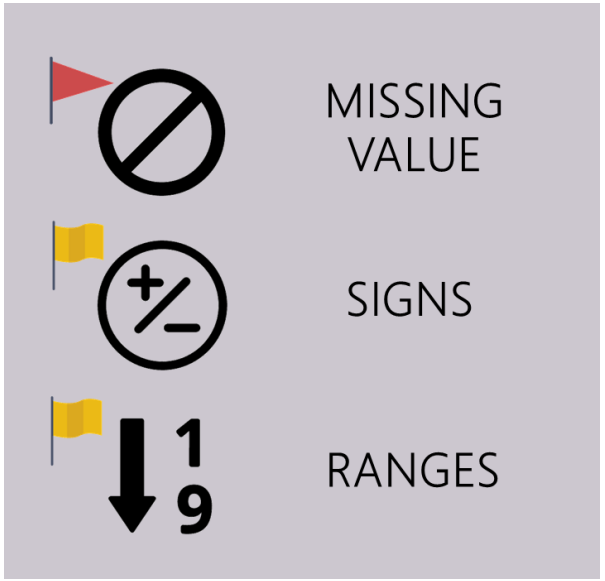
2. Data Engineering – Validation Rules

Carnegie Mellon University Heinz College

FRONT



BACK



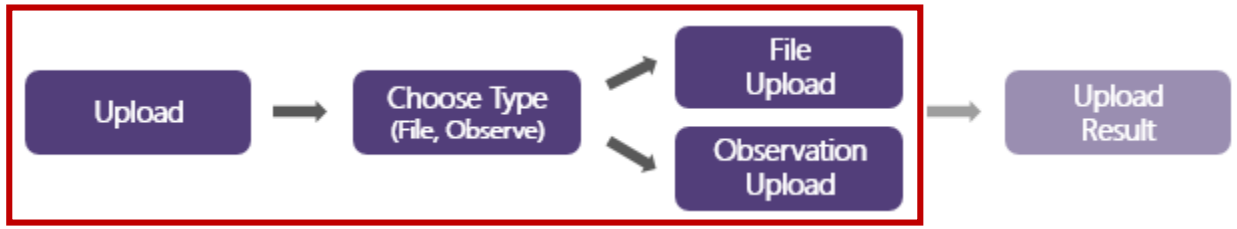
Validation Rules allows users select criteria to assign "RED" or "YELLOW" flags to the data, to eliminate or mark as "Flagged".

- **Missing Values**: Include/exclude null values from the data set
- **Signs**: Include/Exclude positive and negative values from the data set
- **Ranges**: Include/Exclude specific ranges of values from each column

2. Data Engineering – ETL Model

Carnegie Mellon University Heinz College

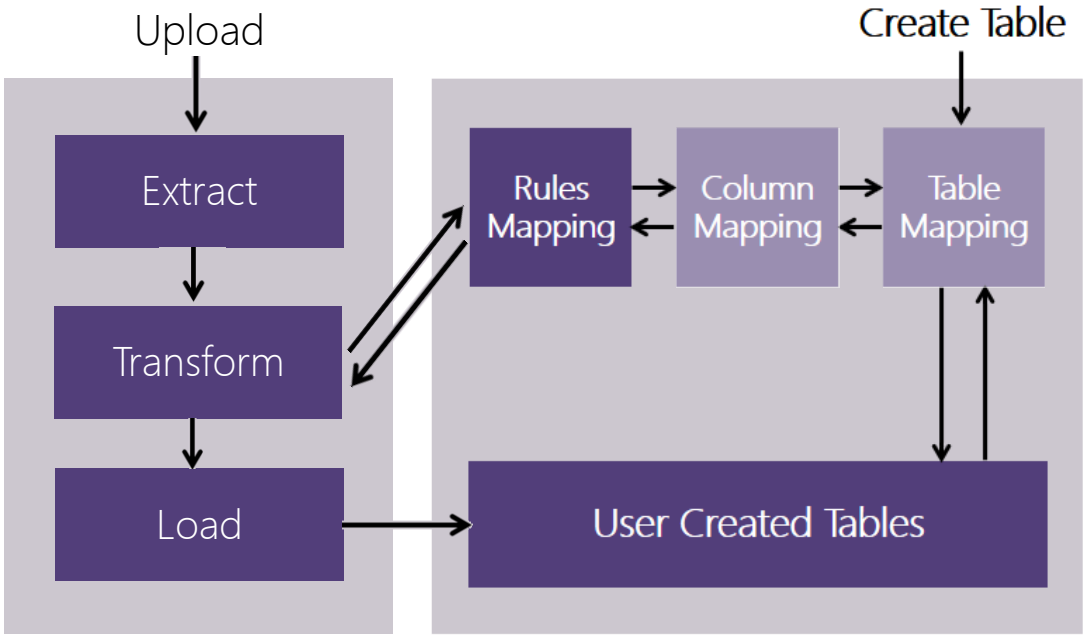
FRONT



BACK

Extract-Transform-Load (ETL)

provides logic to filter the source file, by fetching the information from "Table Mapping" table, and insert the right data to the selected table.

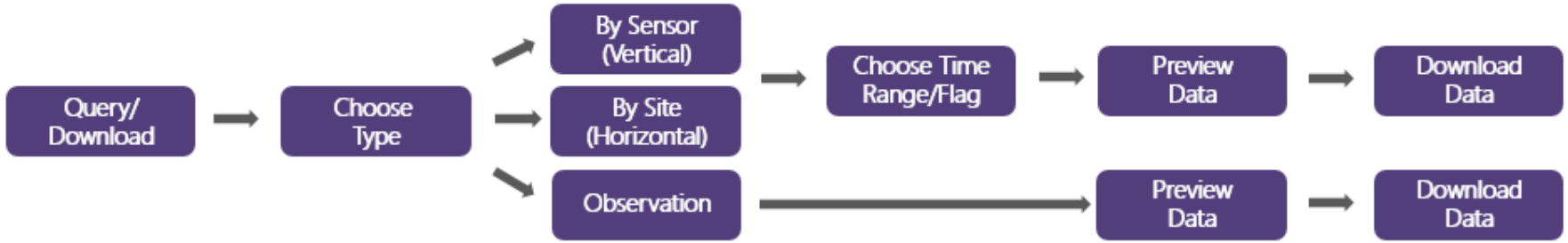


chatham UNIVERSITY

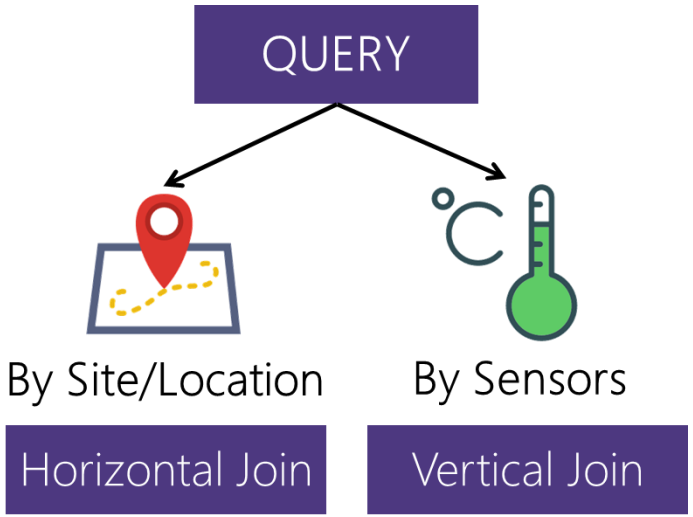
3. Data Aggregation

Carnegie Mellon University Heinz College

FRONT



BACK



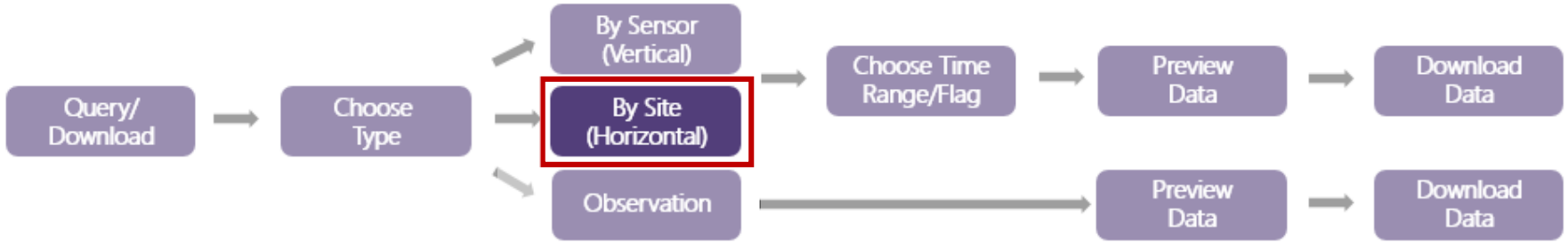
Data Aggregation provides scalable joining logics which can handle the variability in the data.

- **Horizontal Join**: resolves differences in collection intervals for multiple sensors at the same site
- **Vertical Join:** provides a logic for joining differing data schemas from different sites or locations.

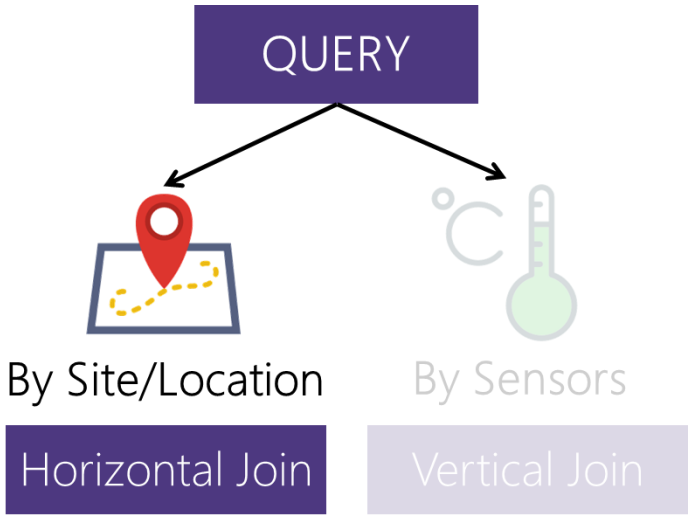
3. Data Aggregation – Horizontal Join

Carnegie Mellon University Heinz College

FRONT



BACK



Example of join result from DO sensor and Temp sensor

Base Table (8 mins)

Time	DO
11/21/2015 0:08	430.82
11/21/2015 0:16	431.19
11/21/2015 0:24	431.19

Join Table (15 mins)

Time	Temp
11/21/2015 0:11	6.727
11/21/2015 0:26	8.747
11/21/2015 0:41	6.767

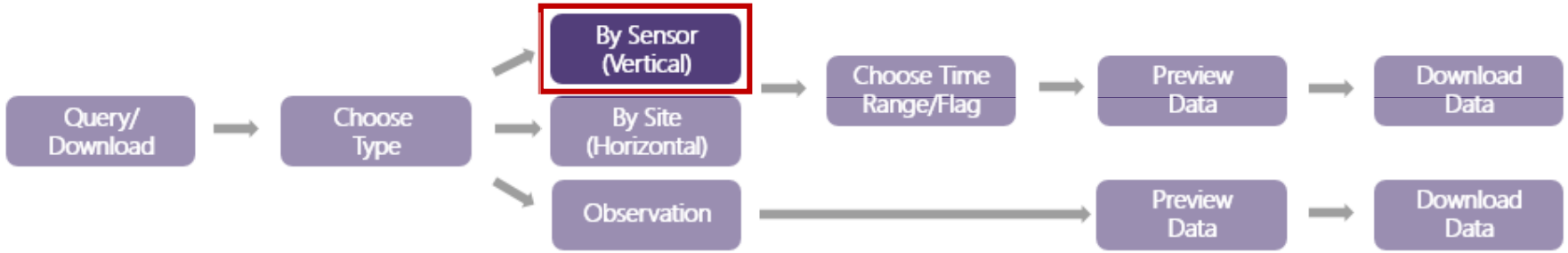
Time	DO	Temp
11/21/2015 0:08	430.82	6.727
11/21/2015 0:16	431.19	-
11/21/2015 0:24	431.19	8.747

chatham UNIVERSITY

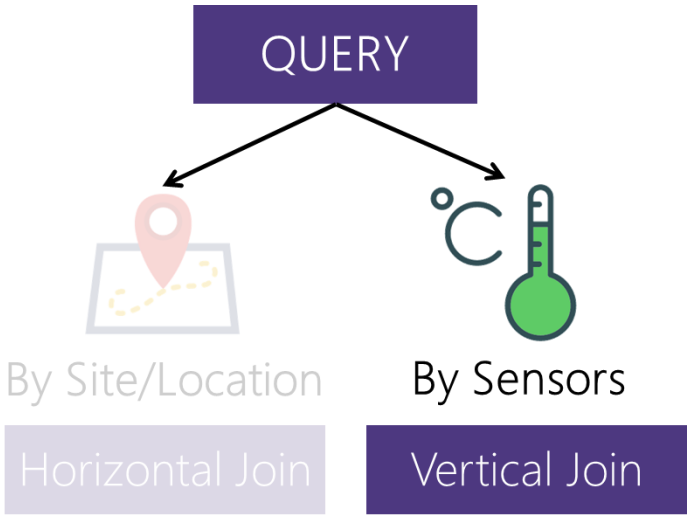
3. Data Aggregation – Vertical

Carnegie Mellon University Heinz College

FRONT



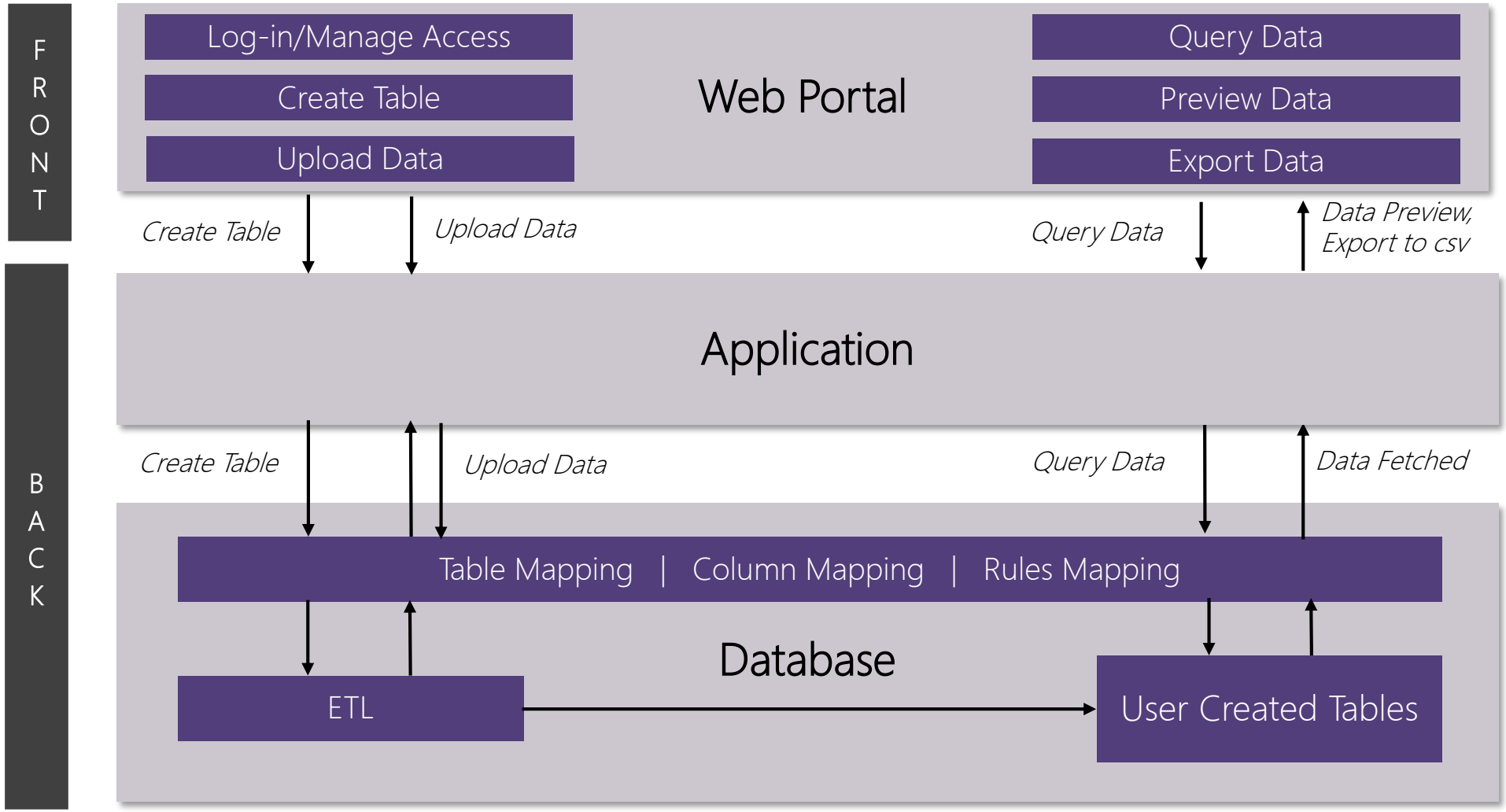
BACK



Example of join result from sites BRCR01 and BRCR02

Site	DO	Temp	Volt
BRCR01	421.25	7.527	-
BRCR01	418.16	8.347	-
BRCR01	419.49	6.467	-
BRCR01	425.32	6.788	-
BRCR01	420.54	7.321	-
BRCR02	399.54	5.253	4.3
BRCR02	386.34	6.123	4.4
BRCR02	401.67	5.795	4.3
BRCR02	396.79	5.367	4.4
BRCR02	370.55	6.032	4.5

System Diagram

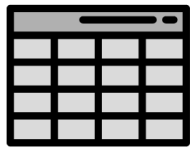


Demonstration

Carnegie Mellon University
Heinz College

LOG-IN AND MANAGE ACCESS

- ✓ User **click create table**, the page **redirects to log-in** page
- ✓ User **log-in** using the incorrect and correct username and password
- ✓ User admin **add and delete access** to the student

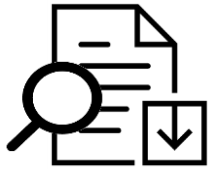


CREATE TABLE

- ✓ User **creates** table with **.txt file format** under **new** sensor type "DO_TEST"
- ✓ User **creates** table with **.txt file format** under **existing** sensor type "DO_TEST"
- ✓ User **creates** table with **.csv file format** under **new** sensor type "CON_TEST"

UPLOAD

- ✓ User **uploads** data from **Site 1** to the table under sensor type "DO_TEST" just created
- ✓ User **uploads** data from **Site 2** to the table under sensor type "DO_TEST" just created
- ✓ User **uploads observation data** to the existing form



QUERY/DOWNLOAD

- ✓ User **queries/downloads** the data **from 2 sites** under sensor type "DO_TEST"
- ✓ User **queries/downloads** the data **from different sensors** from Site 1
- ✓ User **queries/downloads** the data **from Observation Form**

chatham UNIVERSITY

4

Carnegie Mellon University
Heinz College

DOCUMENTATION

Documentation

Carnegie Mellon University
Heinz College

Limited Technical Resource

Technical Specification

- ✓ Introduction
- ✓ Functionality
- ✓ Architecture
- ✓ Design
- ✓ Code

User Manual

- ✓ Introduction
- ✓ Site Map
- ✓ Authorization & Authentication
- ✓ Create Table
- ✓ Upload
- ✓ Download

chatham UNIVERSITY

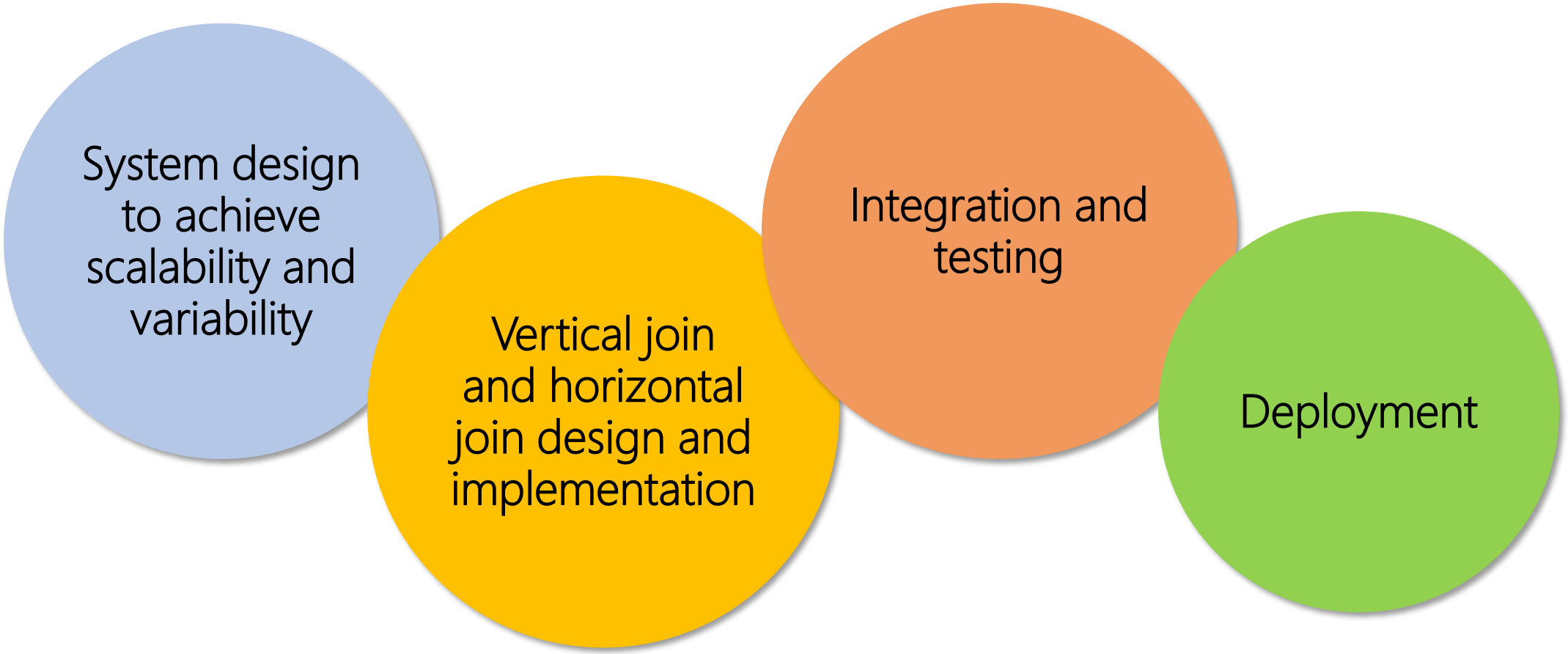
5

Carnegie Mellon University
Heinz College

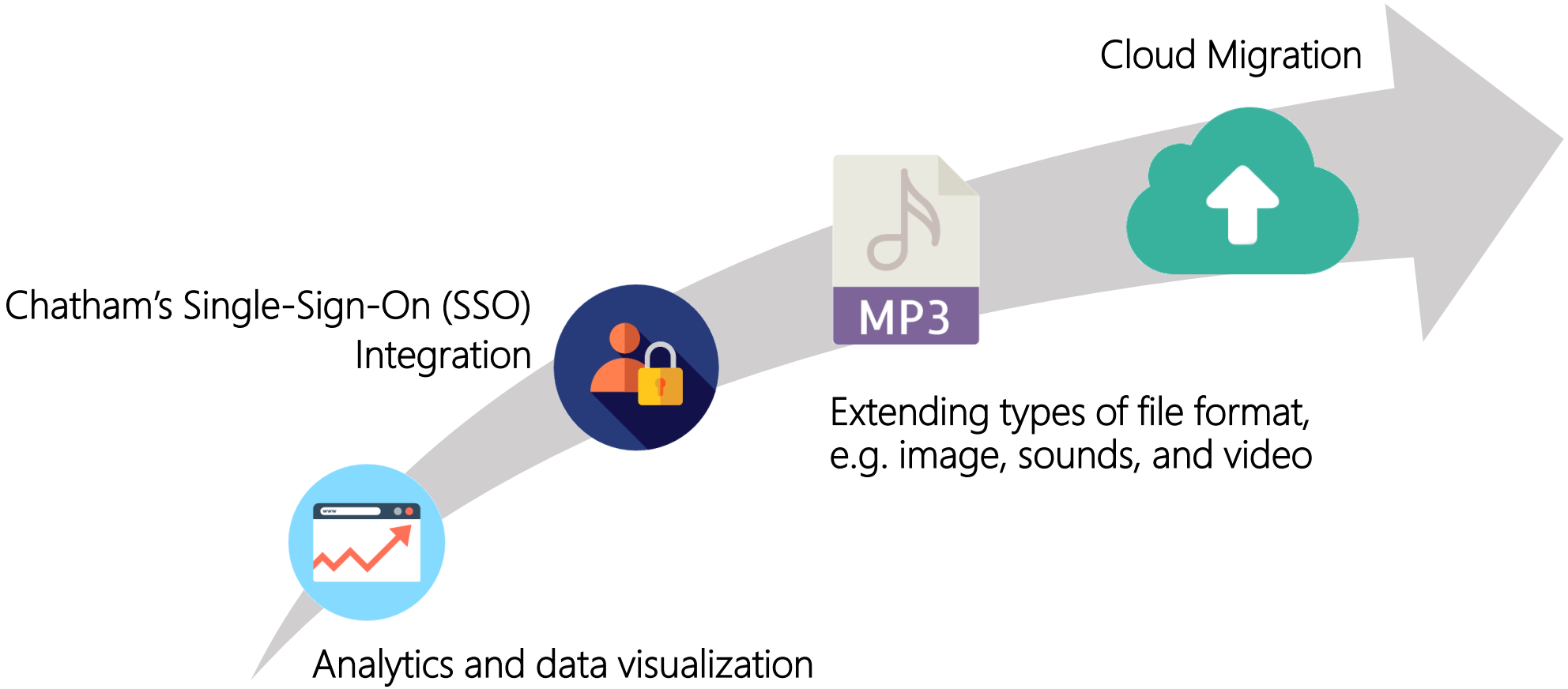
CHALLENGES & FUTURE WORK

Challenges

Carnegie Mellon University
Heinz College



Future Work/Recommendations



Carnegie Mellon University
Heinz College

chatham UNIVERSITY

Carnegie Mellon University
Heinz College

Thank You

chatham UNIVERSITY