Trail Microservice

Introduction

This document outlines the development and implementation of the Trail Microservice. The document is sorted into theses sections: starting off with the background of the microservice and its place within the larger Trail application is a part of.

This is followed by the any Legal, Social, Ethical, and Professional (LSEP) considerations the project has faced, and the measures taken to address them, such as data privacy, integrity, and security. This is preceded by a detailed design section featuring UML diagrams and the entity relationship diagram (ERD) which will demonstrate the database these features are built off from. This will be built upon in the Implementation section, going into the technical detail alongside the technologies used to create this microservice. It will conclude with an evaluation, giving a thorough though of on strengths, weaknesses, and potential improvements of my implementation.

Below is the GitHub which contains the source code and the docker image that needs to run to access the microservice.

GitHub Repo: https://github.com/MxFrgsn/COMP2001-CW2

Docker Image: mxfrgsn/comp2001 cw2

Background

The trail microservice intends to be responsible for all CRUD operations of a trail, allowing people to view, edit, create and delete trails where necessary. This extends to trail attractions, the points of interests and activities the trail is best suited for like dog walking or a museum. In addition, a table and suitable CRUD operations for the location points across the trail – necessary as a trail is made up of a series of location points.

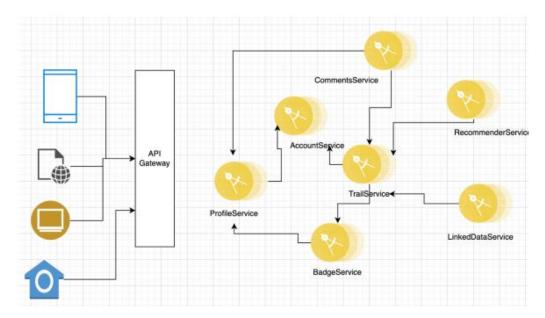
This is intended to be part of a larger Trail application. The Trail application intends to be a service to enhance user wellbeing, providing a reason to explore a particular area through the usage of trails to encourage physical activity and mindfulness. This intends to target people of all types, however the audience most likely to gain usefulness of this app is hikers, fitness and/or outdoor enthusiasts within the teenage to adult demographic.

The microservice primary functionality intends to play a crucial role of the app, by storing and handling the trail information through CRUD (Create, read, update and delete) operations. This will be done through a robust set RESTful API endpoints operation to allow the user to seamlessly interact with the database intending to at least meet these requirements:

- CRUD operation on trails.
- Anyone can view a trail but has limited view.
- Trails are a series of location points.
- Trails are owned by a user.
- As an admin I wish to create a new trail with information.
- As an admin I wish to edit an existing trail.
- As an admin, I wish to delete an existing trail.
- As a user, I wish to create a new trail.
- As a user, I wish to edit the trail(s) I own
- As a user, I wish to delete the trail(s) I own.

- To be protected from OWASP top 10 vulnerabilities where possible.

Here is how the entire Trail app would look, this document covers the trail service at the centre.



Legal, Social, Ethical and Professional (LSEP).

Legal

Legally, compliance with GDPR is essential, requiring transparency about what data is stored and where, ensuring only relevant data is collected. This microservice focuses on trial data, but the inclusion of authentication methods requires passwords and emails to be stored. Currently, these are not hashed, which would be critical for public release. The only personal data stored is names, a low-risk detail, with no age, address, or similar information, adhering to data minimization principles.

How location data is stored needs to be considered, especially as the application evolves to include features like suggesting nearby trails.

Social & Ethical

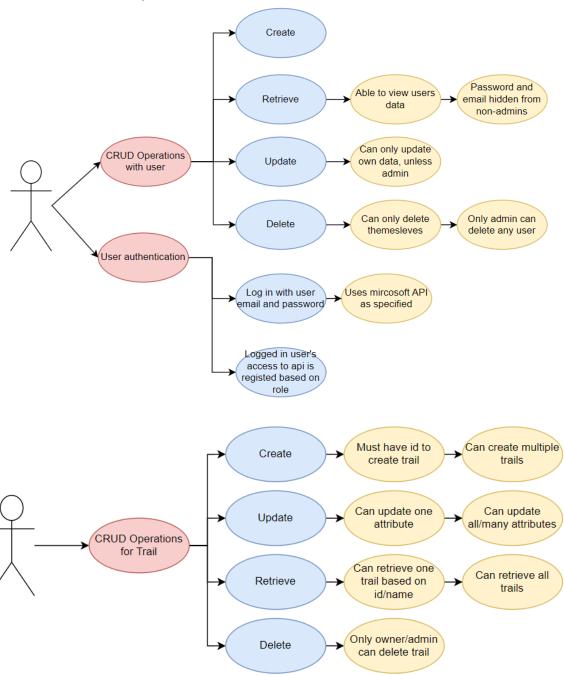
There are minimal ethical concerns aside from ensuring transparency about data store and manipulation. To address this, accounting should be introduced to improve oversight and trust.

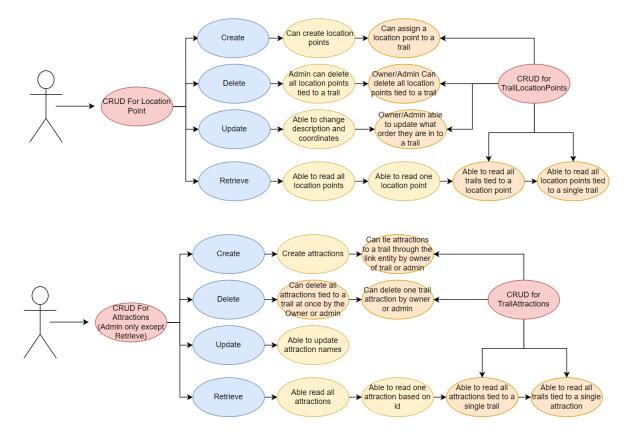
Professional

There are numerous professional issues to consider. Firstly, scalability and performance are always a key consideration. To ensure this, a normalized database and fast network ensure high performance but will need fine-tuning as this microservice integrates with others, particularly on software level. The use of ODBC and Python may limit portability due ODBC inflexible nature, causing potential issues if differing databases are used in other microservices. While the functions are internally tested, the absence of robust unit testing may overlook potential issues or performance bottlenecks in real-world use.

Design

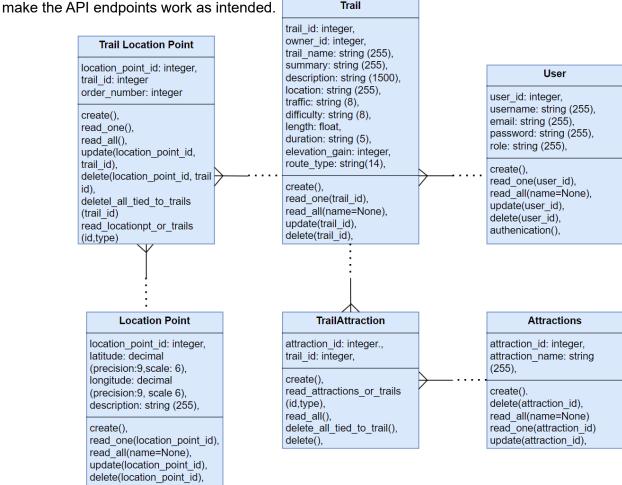
Main features of Implemented Service.

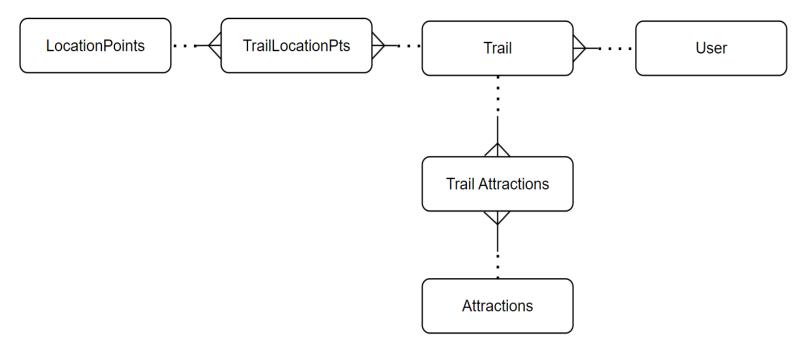




Class diagram

This shows all the attributes of each table used within the database, the exact data type used within Python along with every function being used by the swagger documentation to make the API endpoints work as intended.





Entities in database:

Entities	Description	Relationship
User	Holds user data, used for to know who's signed in and their permissions.	Many to one to trail
Trail	Holds trail data and who the trail	Many to one to both trail attraction and trail location point
Attraction	Holds attraction data that could be tied to a trail	Many to one to trail attraction
Trail Attraction	Link entity that ties attractions to trails	Many to one to attraction and trail
Location Point	Holds location points, used to create location points across a trail	Many to one to trail location point
Trail Location Point	Link entity tying what location points are along the trail and in what order	Many to one to trail and location point

Field definition grids of the entities.

Trail Entity

	,				Beerikk			
Attribute name	Description	Synonym (s)	Data type	Size (*=max)	Possible data values	Optional ?	Validation rules	Key ?
Trail ID	Unique identifier for a trail	N/A	Numeric	N/A	1-99999	N	Must be unique and sequenced.	PK
Owner ID	Identifier of trail creator	N/A	Numeric	N/A.	1-99999	N	N/A	FK
Location	City, State, Country of trail	N/A	Alphanumeric	255	City, State, Country,	N	N/A	N/A
Trail name	Name tied to trail ID	N/A	Alpha with spaces + special characters	255	Any	N	N/A	N/A
Route Type	How the trail is shaped	N/A	Alpha	14	"out and back", "loop", "point to point",	N	Must be "out and back" or "loop" or "point to point". Must be all lowercase.	N/A
Description	A description of the trail, fully detailed	N/A	Alphanumeric with spaces + special characters	1500	Any	Y	There should be at least 100 characters, if data populates it.	N/A
Summary	A short paragraph about features of the trail	N/A	Alphanumeric	255	Any	Υ	There should be at least 10 characters if any data populates it.	N/A
Traffic	How many people use trail	N/A	Alpha	8	light/moder ate/heavy	Υ	Must be "light", "moderate" or "heavy". Must be all lower case	N/A
Difficulty	What is the challenge level of the trail	N/A	Alpha with space	8	easy/mode rate/hard	Υ	Must be "easy", "moderate" or "hard". Must be all lowercase.	N/A
Length	How long the trail is	Distance	Numeric	6	000.00 – 999.99	N	Must use metric measurements, assume measured in km, floating point, rounded to 2 decimal points.	N/A
Elevation Gain	How far up the trail goes	N/A	Numeric	5	00000 – 99999	N	Must use metric measurements, assume measured in meters	N/A
Duration	How long will the trail take to complete	N/A	Alphanumeric + special characters	5	00:00- 99:59	Υ	Must be "extended" military time format – LHS must be <= 99, RHS must be <= 59	N/A

Attraction entity

	ttiraction chit	y							_
Attribute name	Description	Synonym (s)	Data type	Size (*=max)	Possible data values	Optional ?	Validation rules	Key ?	
Attraction ID	Unique identifier for attraction	N/A	Numeric	N/A	1-99999	N	Must be unique and sequenced.	PK	
Attraction Name	Nearby attractions_to see/do	N/A	Alpha with spaces	255	Any	N	Must be valid attraction e.g. "Wildlife", "Historical Site", "Waterfall" etc.	N/A	

Trail Attraction entity

Attribute name	Description	Synonym (s)	Data type	Size (*=max)	Possible data values	Optional ?	Validation rules	Key ?
Trail ID	Unique identifier for a trail	N/A	Numeric	N/A	1-99999	N	N/A	PK, KF
Attraction ID	Unique identifier for attraction	N/A	Numeric	N/A	1-99999	N	N/A	PK, FK

User entity

	tribute name	Description	Synonym (s)	Data type	Size (*=max)	Possible data values	Optional ?	Validation rules	Key ?
L	Jser ID	Unique Identifier for User	N/A	Numeric	N/A	1-99999	N	Must be unique and sequenced.	PK
	Name	Named tied to user id	N/A	Alpha with Spaces + special characters	255	Any	N	Any alpha, only accents and hyphens acceptable.	N/A
	Email	Email tied to user id	N/A	Alphanumeric + special characters	255	AAAA1111 @gmail.co m	N	Must fit standard email format, must contain "@" and ".", no spaces, must be unique	N/A
Pa	assword	Password tied to user id	N/A	Alphanumeric + special characters	128	Any	N	Must be at least 8 characters long, to enforce a degree of security	N/A

Location point entity

Attribute name	Description	Synonym (s)	Data type	Size (*=max)	Possible data values	Optional ?	Validation rules	Key ?
Location Point ID	Unique Identifier for location	N/A	Numeric	N/A	1-99999	N	Must be unique and sequenced.	PK,
Latitude	Latitude coordinate of for location point	N/A	Float	8	90.000000 °- +90.00000 0°,	N	Must be between 90.0000 and +90.0000, must always be a floating-point number, rounded to 6 decimal places	N/A
Longitude	Longitude coordinate of for location point	N/A	Float	9	- 180.00000 0° - +180.0000 00°	N	Must be between 180.0000 - +180.0000, must always be a floating-point number, rounded to 6 decimal places	N/A
Description	Important information tied to location point	N/A	Any	255	Any	Y	N/A	N/A

Trail Location Point Entity

Attribute name	Description	Synonym (s)	Data type	Size (*=max)	Possible data values	Optional ?	Validation rules	Key ?
Location Point ID	Unique Identifier for Iocation	N/A	Numeric	N/A	1-99999	N	Must be unique and sequenced.	PK, FK
Trail ID	Unique identifier for a trail	N/A	Numeric	N/A	1-99999	N	N/A	PK, KF
Order number	The order of which the location point goes across the trail	N/A	Numeric	N/A	1-99999	N	N/A	PK, KF

Implementation

Technologies used:

This microservice server-side code is written in Python, utilizing a Microsoft SQL Server for the database. Initial SQL setup operations are executed through a Jupyter Notebook using the Pyodbc library, leveraging ODBC driver. The Flask framework is employed to handle API routing, with Swagger providing documentation.

In conjunction, SQL Alchemy is utilized to manage data as an Object-Relational Mapper (ORM), mapping each table to a model in models.py, enforcing the necessary relationships and validation that match the SQL, which reduces the risk of SQL injections and maintains data integrity. Marshmallow is employed for serialization and deserialization between request bodies and Python objects, defining the schemas that control data visibility to the signed-in user.

The currently signed-in user is determined using Flask's built-in session management. By default, <u>user 1 is signed in</u>, set up with the admin role for testing purposes. Authentication occurs via the auth endpoint, changing the role based on the given role in the database if inputted data is verified by <u>authenticator</u>. This allows users to sign in using one of three accounts in the database, each having the user role with reduced permissions.

```
def authentication():
   auth_url = 'https://web.socem.plymouth.ac.uk/COMP2001/auth/api/users'
   user_data = request.get_json()
   if not user_data or 'email' not in user_data or 'password' not in user_data:
       return make_response("Missing email or password in request.", 400)
    credentials = {'email': user_data['email'], 'password': user_data['password']}
       response = requests.post(auth_url, json=credentials)
       response.raise_for_status()
       if response.json() == ["Verified", "True"]:
           logged_in_user = User.query.filter(User.email == user_data['email']).one_or_none()
           if not logged_in_user:
               return make_response("User not found in the local database.", 404)
           session['user_id'] = logged_in_user.user_id
           session['role'] = logged in user.role
           return make_response(f"Authenticated successfully. User ID: {logged_in_user.user_id}", 200)
       else:
           return make_response("Authentication failed. Invalid credentials.", 401)
    except requests.exceptions.RequestException as e:
       return make_response(f"Error communicating with authentication service: {str(e)}", 500)
    except Exception as e:
       return make_response(f"An unexpected error occurred: {str(e)}", 500)
```

To host the microservice, a Docker image is provided, as specified at the end of the introduction. Alternatively, you can run it locally by following the <u>GitHub README</u> instructions.

SQL and Models:

The database that has been constructed within the SQL server follows the entity relationship diagram and the field definition grids within the design section of the document. By extension, all the models in models py follow it. Below is a minor snippet of the <u>trail model</u>.

```
class Trail(db.Model):
   __tablename__ = 'Trail'
   __table_args__ = {'schema': 'CW2'}
trail_id = db.Column(db.Integer, primary_key=True)
   owner_id = db.Column(db.Integer, db.ForeignKey('CW2.User.user_id'), nullable=False)
   trail_name = db.Column(db.String(255), nullable=False, unique=True)
   summary = db.Column(db.String(255))
   description = db.Column(db.String(1500))
   location = db.Column(db.String(255), nullable=False)
   traffic = db.Column(db.String(8), nullable=False)
   difficulty = db.Column(db.String(8), nullable=False)
   length = db.Column(db.Float, nullable=False)
   duration = db.Column(db.String(5), nullable=False)
   elevation_gain = db.Column(db.Integer, nullable=False)
   route_type = db.Column(db.String(14), nullable=False)
   owner = db.relationship("User", back_populates ="trail_owner")
   linked_attractions = db.relationship("TrailAttraction", back_populates="trails_attractions_linked", cascade="all, delete, delete-orphan")
   trail_location_points = db.relationship("TrailLocationPt", back_populates ="linked_trail_points", cascade="all, delete, delete-orphan",)
   @validates('trail_name')
   def validate_trail_name(self, value):
       if len(value) < 5:</pre>
           raise ValidationError('Trail name must be at least 5 characters')
       return value
   @validates('summarv')
   def validate_trail_summary(self, value):
       if len(value) < 10 and len(value)!=0:</pre>
           raise ValidationError('Trail summary must be at least 10')
```

CRUD:

Each table has a CRUD operation implemented, uses the HTTP request methods. The code for each function is similar, so I'll focus on the implementation on <u>Trail functions</u>, as its the primary functionality of this service. There is, however, some differences for the link entities of trail attraction and trail location point which will be discussed later.

Create

This function is taking in the request body using the get_json() function within the request within flask. Then, using the built in flask session management, it'll take the role of who's signed in, if it's an admin, it'll take the owner id within the request body, giving them the ultimate control, but if it it's a user, not matter the owner id they inputted, it'll change it to the currently signed in user and then commit the trail the database. The SQL database and the marshmallow ensures that trail information submitted is valid before submission.

```
def create():
    trail_data = request.get_json()
    if session.get('user_id') is None:
        return make_response("User is not authenticated")
    elif session.get('role') != 'admin':
        trail_data['owner_id'] = session.get('user_id')
        new_trail = trail_schema.load(trail_data, session=db.session)
        db.session.add(new_trail)
        db.session.commit()
        return make_response(f"Trail created successfully", 201)
```

Retrieve:

The read one is taking in a singular id, and returning all the information tied to that id using the schema serialized using Marshmallow. The read all just returns all the trail data stored in the database, however with a "name" parameter that allows the user to query a trail on the

trail name, making it more user accessible. In both examples, the amount of information is limited if the signed in user is not an admin, hiding data backend data like the trail id.

```
def read_one(trail_id):
   trail = Trail.query.filter(Trail.trail_id == trail_id).one_or_none()
    if trail is not None:
        if session.get('role') == 'admin':
            return trail_schema.dump(trail)
        return limited_trail_schema.dump(trail)
    else:
        abort(404, f"Trail with trail_id {trail_id} not found")
def read_all(name=None):
    query = Trail.query
    if name:
        query = query.filter(Trail.trail_name.ilike(f"%{name}%"))
   trails = query.all()
    if session.get('role') == 'admin':
        return trails_schema.dump(trails)
    return limited_trails_schema.dump(trails)
```

Update

The update function takes in the trail id and similarly takes in the request body using the get_json function. First, it checks whether the trail exists in the first place, then checks against the signed in user to the owner of the id if they aren't an admin, ensuring only the owner can update their own trail. Then, as the PATCH method is used over PUT, to ensure that only the specified fields are updated rather than everything, the request body must be scanned for each attribute before commit the changed.

```
def update(trail id):
   trail_data = request.get_json()
   existing trail = Trail.query.filter(Trail.trail id == trail id).one or none()
   if existing trail is None:
        abort(404, f"Trail with trail_id {trail_id} not found")
   if session.get('role') != 'admin' and session.get('user_id') != existing_trail.owner_id:
       return make_response(f"Trail {trail_id} cannot be updated, currently authenicated user {session.get('user_id')} is not the owner of the trail.", 400)
   if 'trail_name' in trail_data:
       existing_trail.trail_name = trail_data['trail_name']
   if 'difficulty' in trail data:
       existing trail.difficulty = trail data['difficulty']
   if 'length' in trail_data:
       existing_trail.length = trail_data['length']
   if 'traffic' in trail_data:
       existing_trail.traffic = trail_data['traffic']
   if 'duration' in trail data:
       existing_trail.duration = trail_data['duration']
   if 'elevation gain' in trail data:
       existing_trail.elevation_gain = trail_data['elevation_gain']
   if 'route_type' in trail_data:
       existing_trail.route_type = trail_data['route_type']
   if 'summary' in trail_data:
       existing_trail.summary = trail_data['summary']
   if 'description' in trail_data:
       existing_trail.description = trail_data['description']
   if 'location' in trail_data:
       existing_trail.location = trail_data['location']
   db.session.commit()
   return make_response(f"trail with ID {trail_id} has been updated successfully.", 200)
```

Delete

The delete function also checks the who the owner of the trail is if the signed in user isn't an admin, before allowing them to delete the trail. As relationships are properly defined with models.py, this also means that any attractions/location points linked to the trail using the link entities are also deleted. This works similarly in other functions, with the main exception within user.py delete function, reassigning all the owner id of trails the user might have owned to the id 1, which is the id of the admin to maintain data integrity.

```
def delete(trail_id):
    existing_trail = Trail.query.filter(Trail.trail_id == trail_id).one_or_none()
    if session.get('user_id') != existing_trail.owner_id and session.get('role') != 'admin':
        return make_response(f"Trail {trail_id} cannot be deleted, currently authenicated user {session.get('user_id')} is not the owner of the trail.", 400)
    if existing_trail:
        db.session.delete(existing_trail)
        db.session.commit()
        return make_response(f"trail with ID {trail_id} has been deleted", 200)
    else:
        abort(404, f"trail with ID {trail_id} not found")
```

How link entities are different

The primary difference is the lack a "read one" function due to their compound key structure and absence of standalone data worth retrieving. Instead, a "get" function retrieves all IDs tied to a given input ID. For attractions, inputting an attraction ID returns all trails linked to it, and vice versa. This relies on specifying the ID type, crucial as IDs are auto-assigned integers via SQL's IDENTITY function. Additionally, there's no update function for trail attractions, as updating an entry is nearly identical to creating one with added steps.

Below is code used in <u>trail attraction.py</u>, a similar implementation used within trail location pt.py.

```
def read_attractions_or_trails(id, type):
    if type == 'trail':
        attribute = 'trail_id'
    elif type == 'attraction':
        attribute = 'attraction_id'

    trail_attraction = TrailAttraction.query.filter(getattr(TrailAttraction, attribute) == id).all()
    if trail_attraction:
        return trail_attractions_schema.dump(trail_attraction)
    else:
        abort(404, "No Trail Attraction found for the given ID")
```

Permissions Overall:

As an admin, you have access to all data, with permissions to create, read, update, and delete. The main difference for users is amongst two things: hiding some data such as trail id and the username and passwords of other users when viewing all users, by extension there are restrictions in place to ensure that users can only perform CRUD operations on their own trails. Additionally, the limitations extend to the Attraction endpoints, which are restricted to Admins for creating and deleting.

Evaluation

Overall, the project successfully achieved its objective of implementing the backend code for the trail microservice, fulfilling the requirements outlined in the background. However, several improvements could improve its functionality, security, and scalability:

Database Improvements

1. Location Table:

 Adding a table to store location information (e.g., city, state, and country) would prevent duplication and ensures consistency plus scalability.

2. Enhanced Data Retrieval:

 When requesting trail data, including location point data and attraction details would provide a more comprehensive response, making the service more user-friendly and useful.

Design Decisions

3. User Role Permissions for Location Points:

- Allowing users to create their own location points could enable more trails to be created entirely by users. However, this introduces risks, such as misuse and unnecessary data storage, potentially leading to increased long-term costs.
- In the current implementation, users can create location points, but a decision should be made to balance user empowerment and system efficiency.

Testing and Debugging

4. Default Session Role:

- For debugging and testing, the session defaults to the admin role, which simplifies testing but is unrealistic for a real-world scenario.
- Ideally, the session should default to the signed-in user or a guest role, but this feature is out of scope for this project and would typically be managed by other components of the microservice.

Permission and Privilege Management

5. Updating Attributes:

 Currently, certain attributes, like the owner_id, cannot be updated, even by an admin. While this limitation rarely impacts functionality, enabling admins to modify all data would align with their intended role.

6. Stricter Privilege Systems:

 Implementing a least-privilege system would improve security by ensuring users have access only to the data and functions necessary for their role.

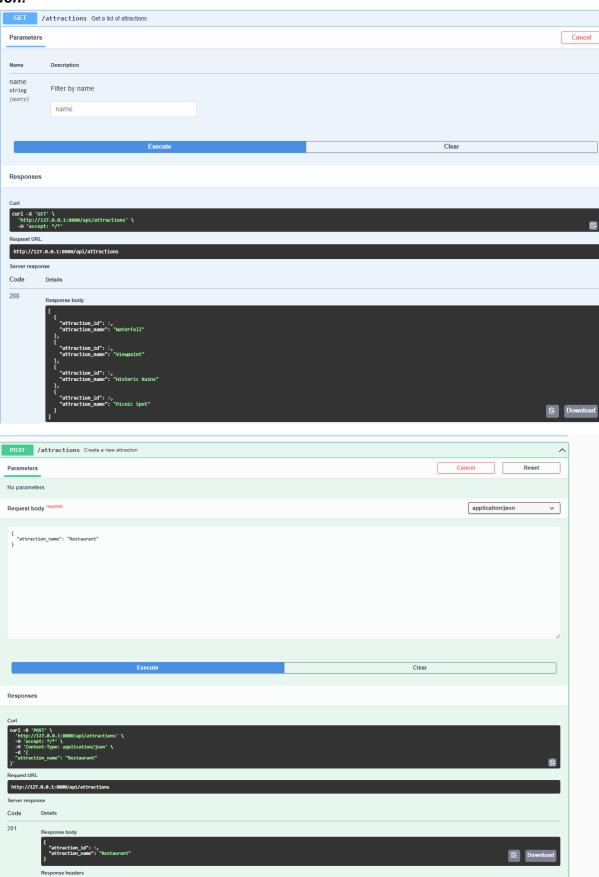
Security Enhancements

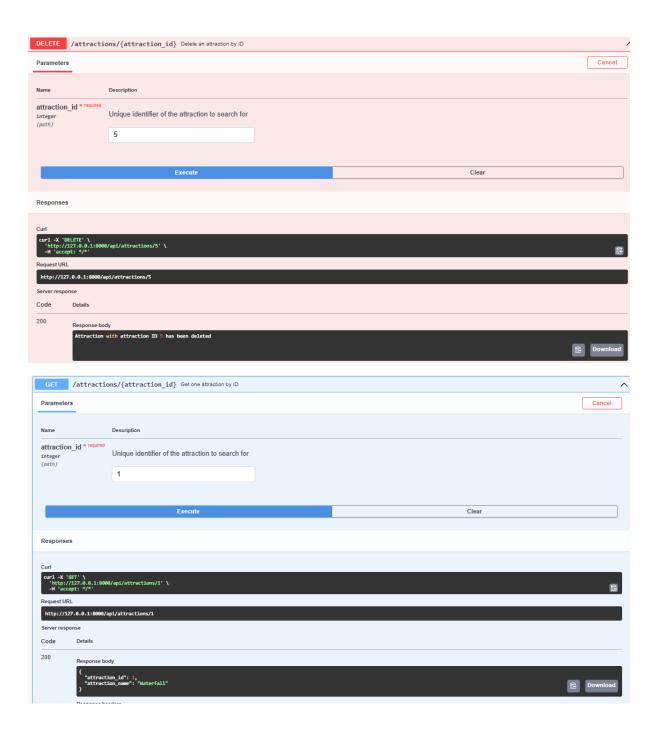
7. Data Security:

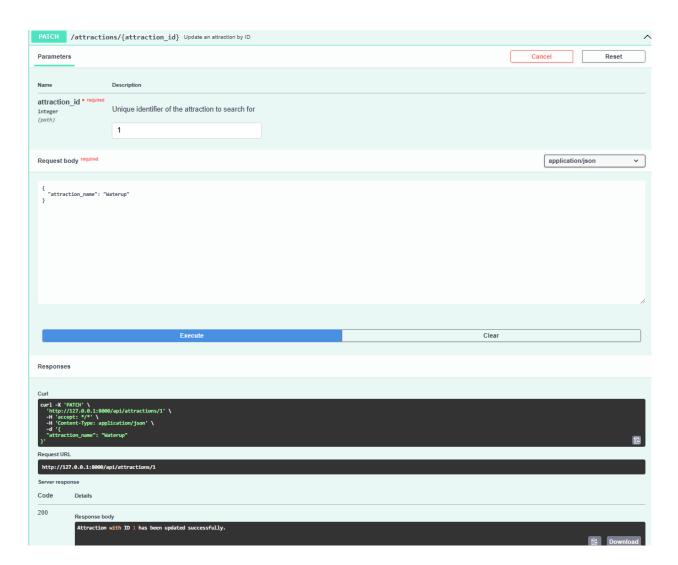
 Hashing sensitive data, such as passwords, and implementing encryption during data transmission would significantly enhance security.

Screenshots of testing

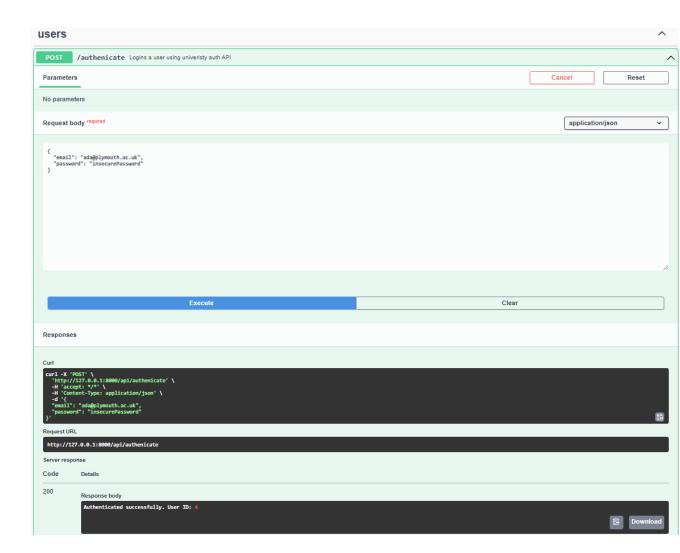
Attraction:

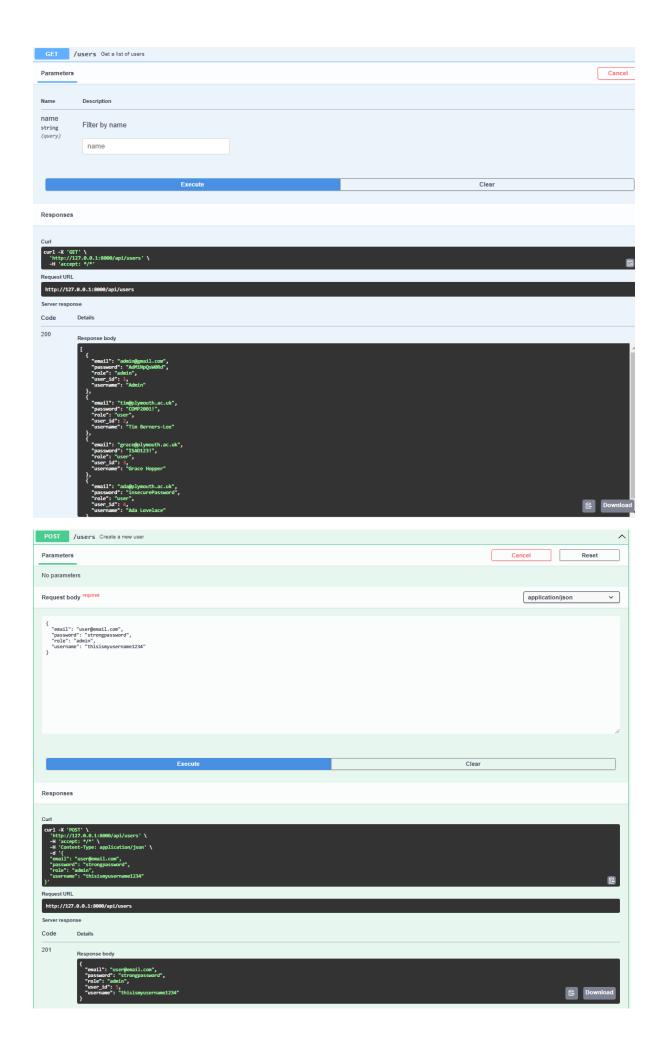


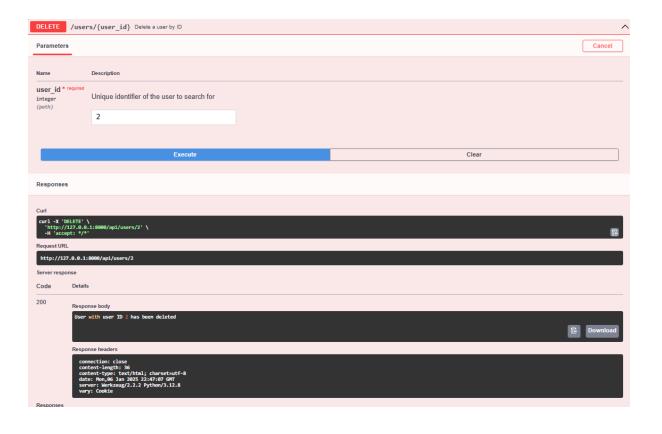


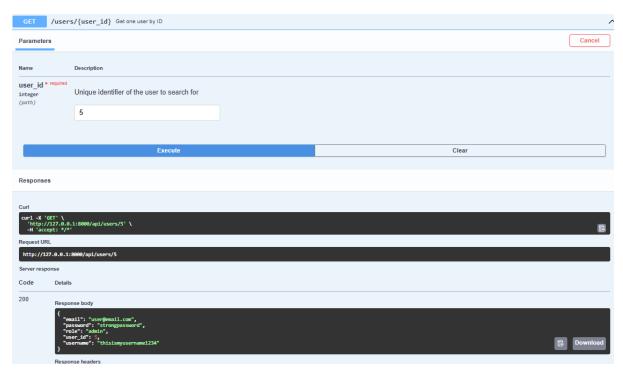


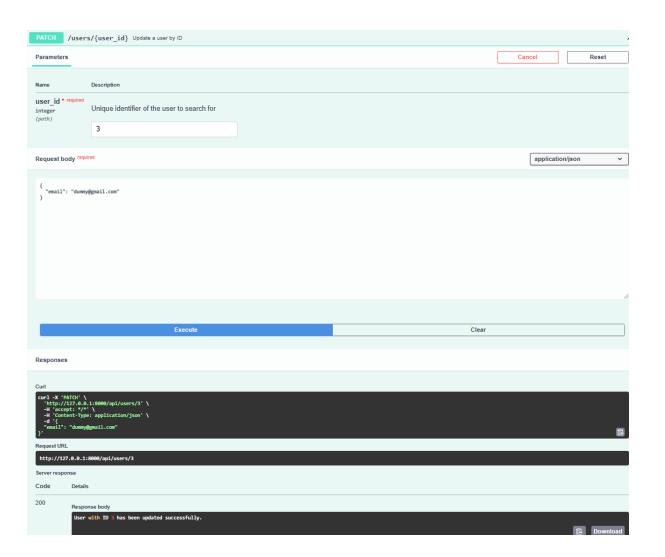
Users



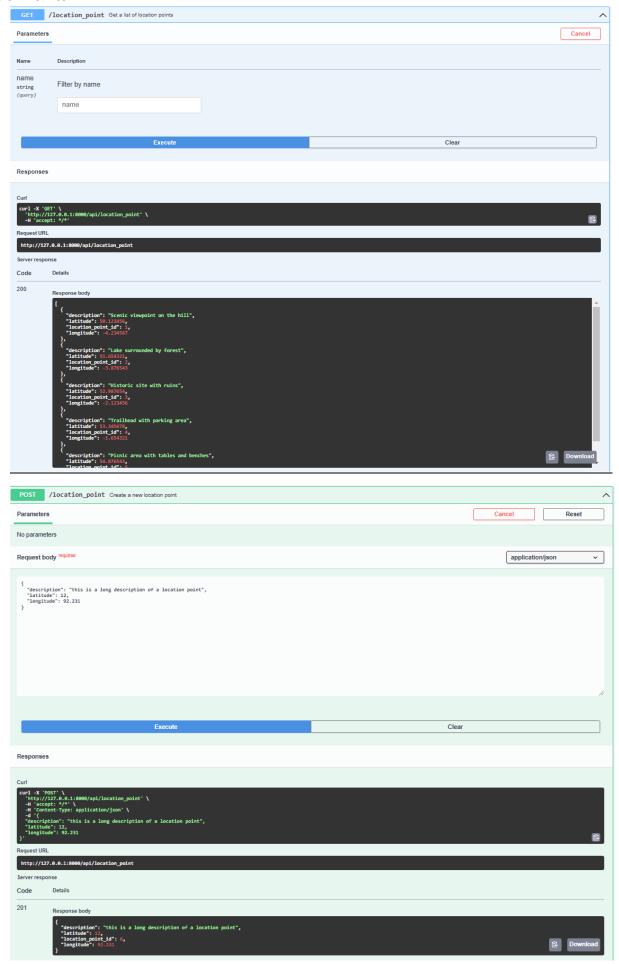


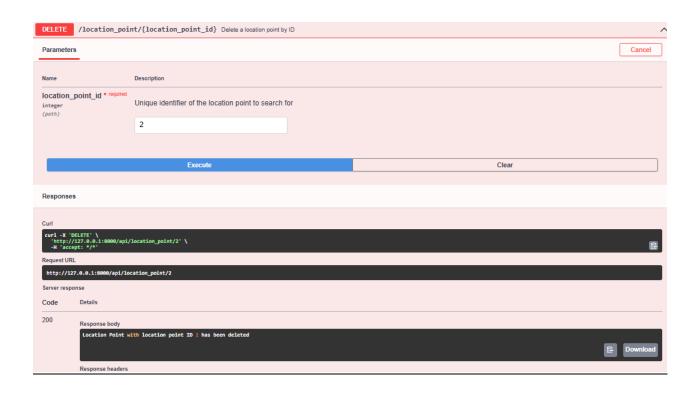


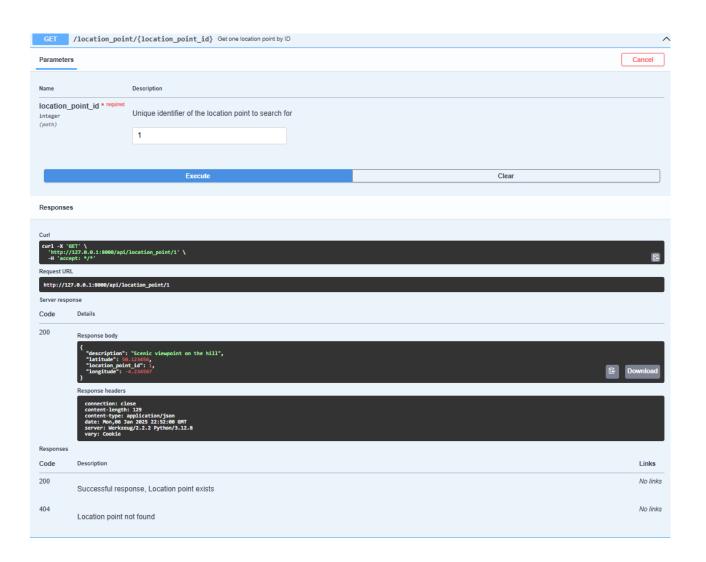


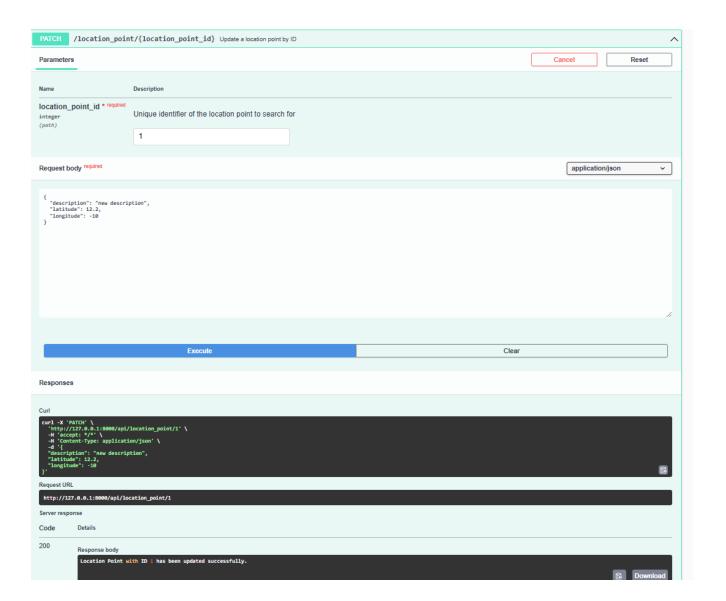


LocationPoints

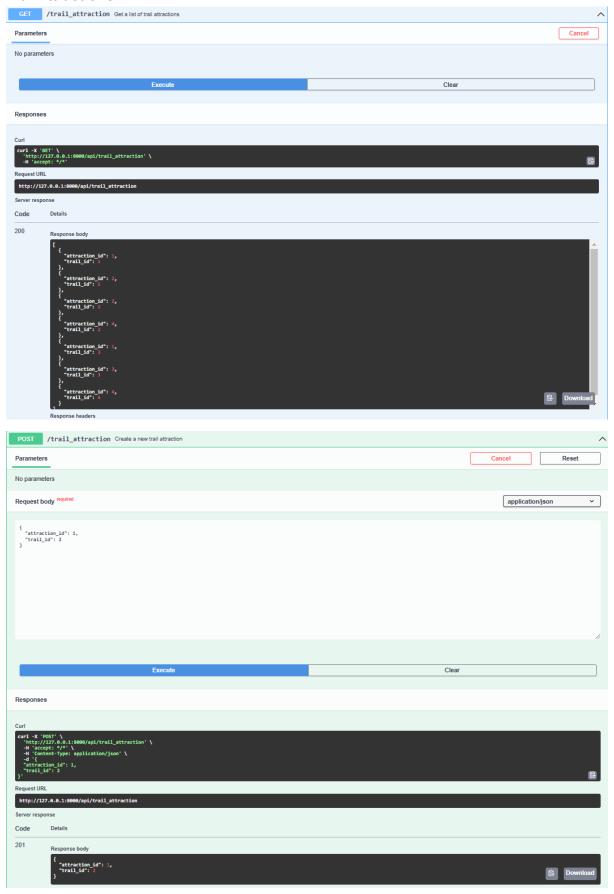


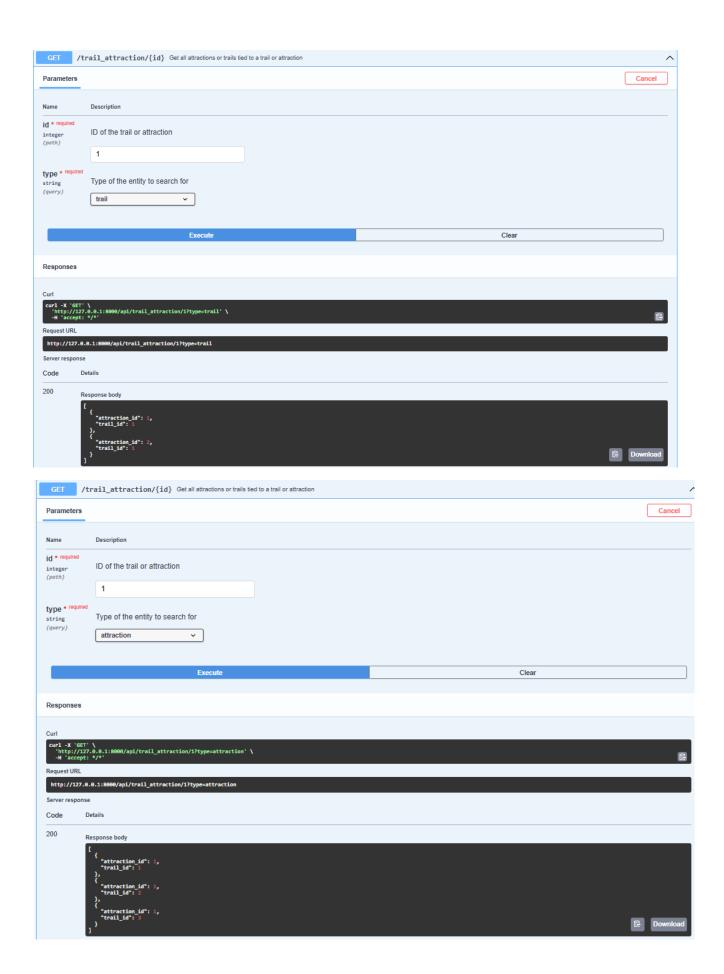


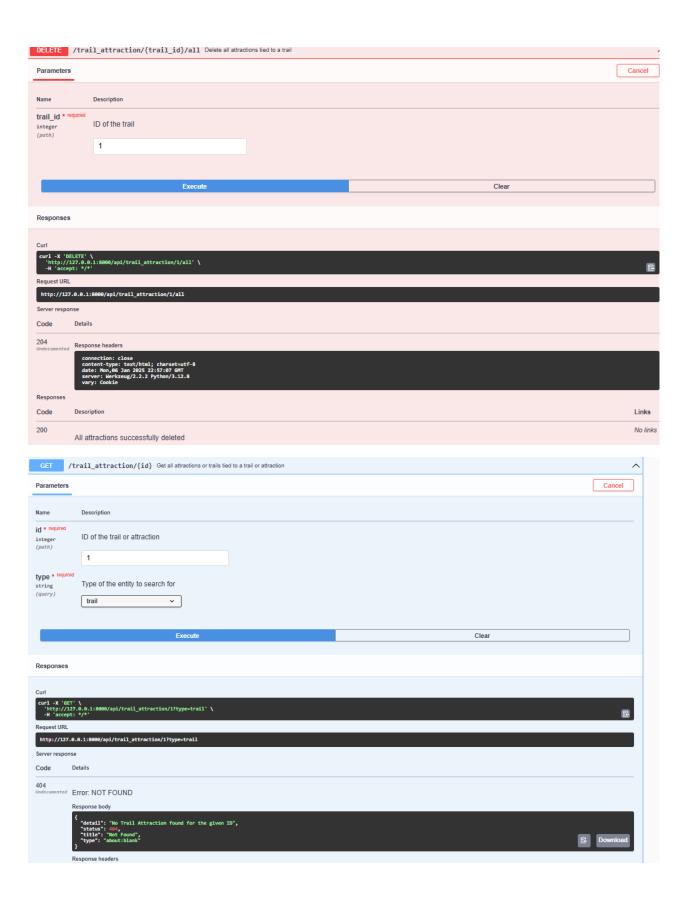


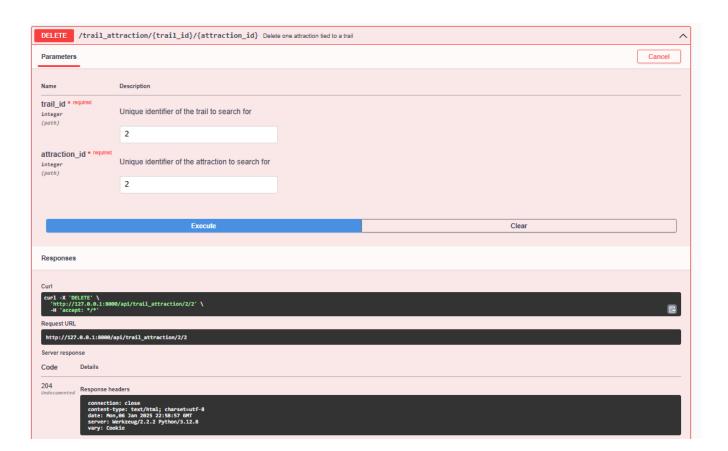


TrailAttractions

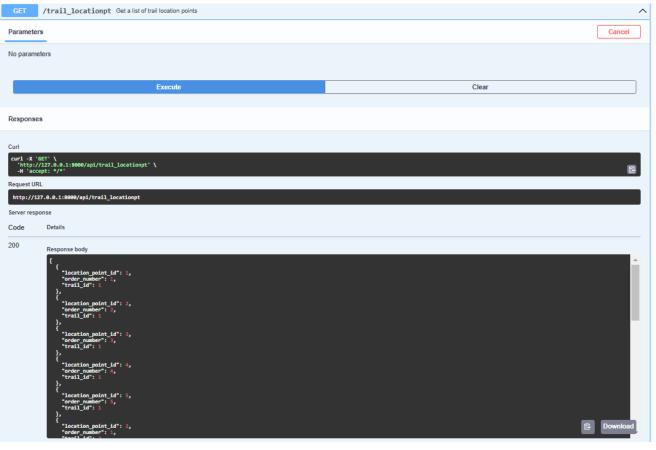


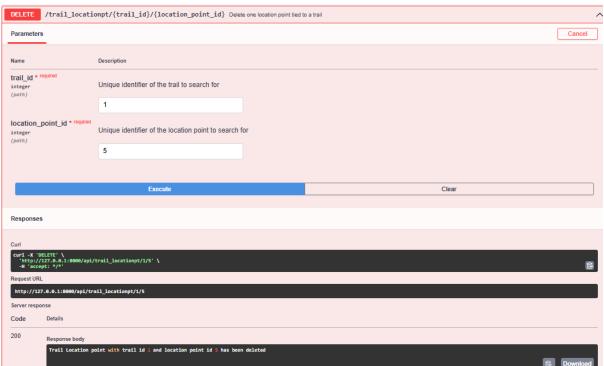




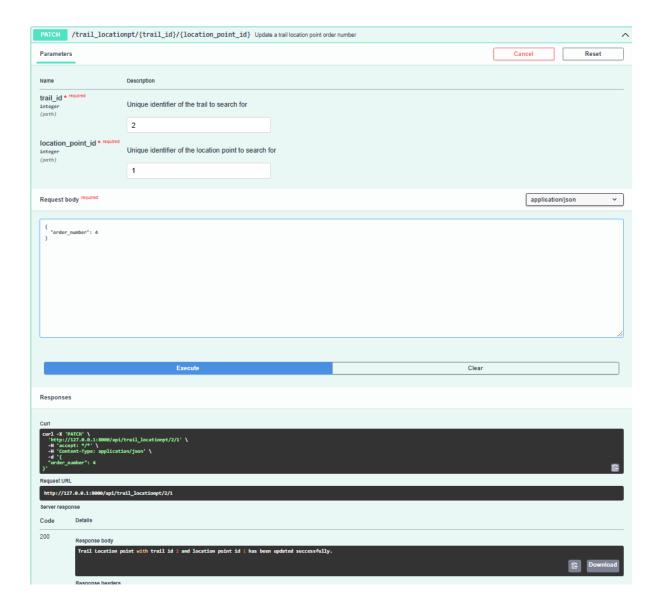


TrailLocationPoints









Trails

