EDS 213: Metadata Standards, Data Modeling and Data Semantics

This course will introduce students to reuse and the preservation of environmental data

Contents

Instructor
Important links
Course description
Learning objectives
Curriculum
Session (subject to change)
Grading
Course requirements
  Computing
  Textbook (optional)
This course will cover the concept of metadata and how it can be leveraged for the integration of heterogeneous datasets into standardized data products. We will practice how to download data from data repositories both manually and programmatically relying on APIs. We will also discuss how to track the provenance of data, generate metadata integrating data semantics to increase data discovery, as well as archiving data products on data repositories to make them available to the broader community.
• Re-use data from others leveraging data repositories and metadata standards

• Rely on APIs to programatically access and download data

• Understand concept of metadata, data semantics and data preservation and why it matters in our path towards reproducible and open science.

Curriculum

EDS 213 is a half Fall quarter course (2 units) running from Monday 2021-09-27 to Wednesday 2021-10-27. The course is developed around two main themes:

• Re-using data from others

• Preserving your data

Session (subject to change)

<table>
<thead>
<tr>
<th>Day / Session</th>
<th>Session Type</th>
<th>Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday 9/27</td>
<td>Lecture</td>
<td>Re-using data – Data Preservation &amp; Data Repositories</td>
</tr>
<tr>
<td>Wednesday 9/29</td>
<td>Lab</td>
<td>Re-using data – Using dataONE to find environmental data &amp; Journal club</td>
</tr>
<tr>
<td>Monday 10/4</td>
<td>Lecture</td>
<td>Re-using data – Accessing data in a programatically way</td>
</tr>
<tr>
<td>Wednesday 10/6</td>
<td>Lab</td>
<td>Re-using data – APIs for environmental data</td>
</tr>
<tr>
<td>Monday 10/11</td>
<td>Lecture</td>
<td>Preserving data – Data modeling &amp; Metadata standards for environmental data</td>
</tr>
<tr>
<td>Wednesday 10/13</td>
<td>Lab</td>
<td>Preserving data – Developing your Data Management Plan</td>
</tr>
<tr>
<td>Monday 10/18</td>
<td>Lecture</td>
<td>Preserving data – Data Provenance and Semantics</td>
</tr>
<tr>
<td>Wednesday 10/20</td>
<td>Lab</td>
<td>Preserving data – Documenting things</td>
</tr>
<tr>
<td>Monday 10/25</td>
<td>Lecture</td>
<td>Preserving your data – uploading data to the KNB</td>
</tr>
<tr>
<td>Wednesday 10/27</td>
<td>Lab</td>
<td>Project presentations</td>
</tr>
</tbody>
</table>

Grading

The grading for this course is organized as follow:

• 20% Class participation
Course requirements

Computing

- Minimum MEDS device requirements
- Have a ready to be used GitHub Account (https://github.com/)

Textbook (optional)

- R for Data Science: https://r4ds.had.co.nz/