

DAPS&CO

Data Access & Regulation, Module 3

Instructor: Robert Fahey <robfahey@aoni.waseda.jp>

Assignment 2 9 Feb 2021 (Tue)

Part One:

In today's class, I introduced some of the strengths and weaknesses of several different ways to store and access data:

- File Sharing
- SQL Databases
- NoSQL Databases
- Network Databases
- Cloud Platforms

For your assignment, I want you to write a couple of lines about each one of those data storage types, stating what kinds of project you think it would be most appropriate for.

For each one, think about both the **scale** of the projects it's appropriate for (i.e. is it just one person, a small team, or a large cross-national project?) and the **type of data** it would be suitable for (is it relatively small-scale, like survey data; medium-sized, like a large amount of newspaper or social media data; or a really huge data set, with tens of millions of items like Facebook posts or customer interactions?).

(If it helps, imagine that you're going to have an exam tomorrow about this lecture, and this assignment is your "cheat sheet" that you're allowed to carry into the exam. What would you need to write down about each data storage type to pass that exam?)

(Don't worry, there won't actually be an exam...)

Part Two:

For yesterday's assignment, you sent me information about two sets of data: your own project data, and an interesting dataset you found online.

Now I'd like you to think about each of those datasets again. Think about the kind of data they contain, and the kind of research you'd want to do with them.

In an email, write down which data storage approach you think would be most appropriate for each of these data types, and why you think so.

(These are hypothetical projects, so of course, there are no wrong answers.)

If you haven't submitted yesterday's assignment - the one describing those two data sets - then you may submit that assignment along with today's assignment, and on this occasion I won't penalise you for late submission.

Submission:

Email both of these data descriptions to me (robfahey@aoni.waseda.jp) before the next class (Wednesday, Feb 10th, 9am CET).

Late Submission:

Late submissions without a reason (illness etc.) will be accepted (with a grade penalty of 5% per weekday) up until the day of the module's final class (Tuesday, Feb 23rd). Any submission after that date will be rejected and your grade for this assignment will be zero.