Assignment

Random file

- You are only given the name of the file, and no other information.
- You need to figure out everything else about the file on your own.
- Supervisors only help with specific programming problems, not general questions about the file or its contents

What to do?

- Do some detective work documenting the
  - To help you displaying the file
  - To prepare for writing the report, which requires that the questions on the following slides are answered
  - Not all questions have answers, but we need to research to know as much as possible about the file
- Write code to display the file
  - Write an application that presents the file in an appropriate way (only tables is ok, like in the examples)
- If enough time, possibly some additional work
  - Add filtering
  - Add more data
  - Test with other files found on the internet
  - Extend the file in some way
- Write a report that details all of the above

Questions to answer

- What kind of document is it, database centric (like tables in a database with a simple and well known structure) or document centric (like a web page or a word processor document with unknown structure)
- Who designed the file format, and when was it designed?
- What is the format called and what is it used for?
- Are there any controversies surrounding the format?
- Is the format made for any specific type of application?
- Is the format an official standard (such as ISO) and/or is there a DTD or schema for the file format?
- Are there specific API-s already made for working with this type of file?
- Are there other related formats?
- Are there other versions of this format, if so what are the chief differences?
- Is there anything else that is interesting about this file?
- Which form of presentation and which form of API is best for this type of file?
- Is there any book or other form of documentation that gives any type of information about this type of file?
- Are there any other interesting things regarding this type of file that will be interesting for a reader of the report?

Documentation

- Give a brief account of the layout of the file and what type of elements that it contains
- If there is a DTD or Schema, can it be used to validate the file, and what was the result of the validation?
- What do we learn about the chosen API, after having made the implementation, was it a good choice?
- What do we learn about the contents of the file, if the file for example stores a contract, describe as much as possible about the specific contract that is stored in the file
- Write code that is as specific to the contents of the file as possible

Considerations

- Some files are large, so perhaps it is enough to focus on a part of the file
- If the structure of the file is simple, place more importance on the presentation, if it has a more complicated structure, focus on the navigation of the file
- Make the code general so that it accepts other files of the same format, not just the given file
- If you have time, make the presentation interactive (for example through filtering like in the first mini-assignments) or add graphics (like in the tips at the end of this lecture)

Programming tips

- The large assignment is for the most part a question of making a the same code as in the smaller applications, but for a more complex dataset.
- There are however some differences
- Namespaces
  - Some files use namespaces, which make programming more complicated
- Working with graphics
  - Generating box-graphics
  - Simple line graphics using html5

Namespaces

- Namespaces allow us to have finer control of the meaning of our tags
- For the XML file, there is little difference, we declare a namespace and use it either directly or by defaulting
- Defaulting name to the person namespace

```
<person xmlns="http://persondata.com/">
  <name>John</name>
</person>
```

- Namespace declaration

```
<root
xmlns:person="http://persondata.com/"
xmlns:car="http://cardata.com/">
```

- Namespace use

```
<person>
        <person:name>John</name>
</person>
```

Generating Boxes

Generating HTML5 code

- A precursor to the web programming course
- Some of the files contain data that can be represented graphically
- Either this data can be represented using text, by printing the information on-screen
- Or we could generate graphical primitives that use the data as a source for placement of graphics
  - Box graphics using ordinary html boxes
  - HTML5 code that draws lines directly