File Formats

File Formats

```
<?xml version="1.0"
encoding="UTF-8"?>

<coffee objname="c1">
    <name> mocha </name>
    <shop> costa </shop>
    <price> 2.0 </price>
    <rating> 3.5</rating>
    <favourite> 0 </favourite>
</coffee>
```

When a simple file persistence strategy, CSV, Name/Value, YAML, XML & JSON are all potential candidate formats for file-based storage.

Common Formats

- Comma Separated Values (CSV)
- Name/Value Pairs
- YAML
- XML
- JSON

CSV

http://en.wikipedia.org/wiki/Comma-separated_values

- A comma-separated values (CSV) file stores tabular data (numbers and text) in plain-text form.
- Plain text means that the file is a sequence of characters, with no data that has to be interpreted instead, as binary numbers.
- A CSV file consists of any number of records, separated by line breaks of some kind; each record consists of fields, separated by some other character or string, most commonly a literal comma or tab.
- Usually, all records have an identical sequence of fields.

```
“mocha”, “costa”, 2.0,  3.5,  0
“americano”, “costa”, 3.0, 4.5, 1
“cappucino”, “starbucks”, 4.0, 1.5, 0
```

data.csv

Name/Value Pairs

http://en.wikipedia.org/wiki/Attribute-value_pai

- A name–value pair, key–value pair, field–value pair or attribute–value pair is a fundamental data representation in computing systems and applications.
- Designers often desire an open-ended data structure that allows for future extension without modifying existing code or data.
- In such situations, all or part of the data model may be expressed as a collection of tuples <attribute name, value>; each element is an attribute–value pair.
- Depending on the particular application and the implementation chosen by programmers, attribute names may or may not be unique.

```
db.url=jdbc:cloudbees://pacemaker
db.driver=com.mysql.jdbc.Driver
db.user=pacemaker
db.pass=pacemaker
jpa.ddl=create
```

application .conf

```
name="mocha"
shop="costa"
rating=3.5
price=2.0
favourite=0
id=1
```

data.conf

http://en.wikipedia.org/wiki/XML

XML

- Extensible Markup Language (XML) is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable.
- The design goals of XML emphasize simplicity, generality, and usability over the Internet.
- It is a textual data format with strong support via Unicode for the languages of the world.
- Although the design of XML focuses on documents, it is widely used for the representation of arbitrary data structures, for example in web services.

```xml
<?xml version="1.0" encoding="UTF-8"?>

<coffee objname="c1">
  <name> mocha </name>
  <shop> costa </shop>
  <price> 2.0 </price>
  <rating> 3.5</rating>
  <favourite> 0 </favourite>
</coffee>

<coffee objname="c1">
  <name> americano </name>
  <shop> costa </shop>
  <price> 3.0 </price>
  <rating> 4.5 </rating>
  <favourite> 1 </favourite>
</coffee>

<coffee objname="c1">
  <name> cappucino </name>
  <shop> starbucks </shop>
  <price> 4.0 </price>
  <rating> 1.5 </rating>
  <favourite> 0 </favourite>
</coffee>
```

data.xml

YAML

http://en.wikipedia.org/wiki/YAML

- YAML (YAML Ain't Markup Language) is a human-readable data serialization language.
- It is commonly used for configuration files, but could be used in many applications where data is being stored (e.g. debugging output) or transmitted (e.g. document headers).
- YAML targets many of the same communications applications as XML but has a minimal syntax
- It uses both Python-style indentation to indicate nesting, and a more compact format that uses [] for lists and {} for maps[1] making YAML 1.2 a superset of JSON.[2]

```
Coffee(c1):
    name       : mocha
    shop       : costa
    price      : 2.0
    rating     : 3.5
    favourite  : 0

Coffee(c2):
    name       : americano
    shop       : costa
    price      : 3.0
    rating     : 4.5
    favourite  : 1

Coffee(c3):
    name       : cappucino
    shop       : starbucks
    price      : 4.0
    rating     : 1.5
    favourite  : 0
```

data.yaml

JSON

http://en.wikipedia.org/wiki/JSON

- JavaScript Object Notation, is a text-based open standard designed for human-readable data interchange.
- Derived from the JavaScript scripting language, JSON is a language for representing simple data structures and associative arrays, called objects.
- Despite its relationship to JavaScript, JSON is language-independent, with parsers available for many languages.
- The JSON format is often used for serializing and transmitting structured data over a network connection. It is used primarily to transmit data between a server and web application, serving as an alternative to XML.

```
{
  "name":"mocha",
  "shop":"costa",
  "rating":3.5,
  "price":2.0,
  "favourite":0,
  "id":1
},
{
  "name":"americano",
  "shop":"costa",
  "rating":4.5,
  "price":3.0,
  "favourite":1,
  "id":2
},
{
  "name":"cappuccino lite",
  "shop":"starbucks",
  "rating":1.5,
  "price":4.0,
  "favourite":1,
  "id":3
}
```

data.json