

**ESCOLA  
SUPERIOR  
DE TECNOLOGIA  
E GESTÃO**

**P.PORTO**

**Bruno Oliveira**  
2018/2019

**Apresentação de modelos estruturados, semiestruturados  
e não estruturados para o armazenamento de dados**

Processamento Estruturado de Informação

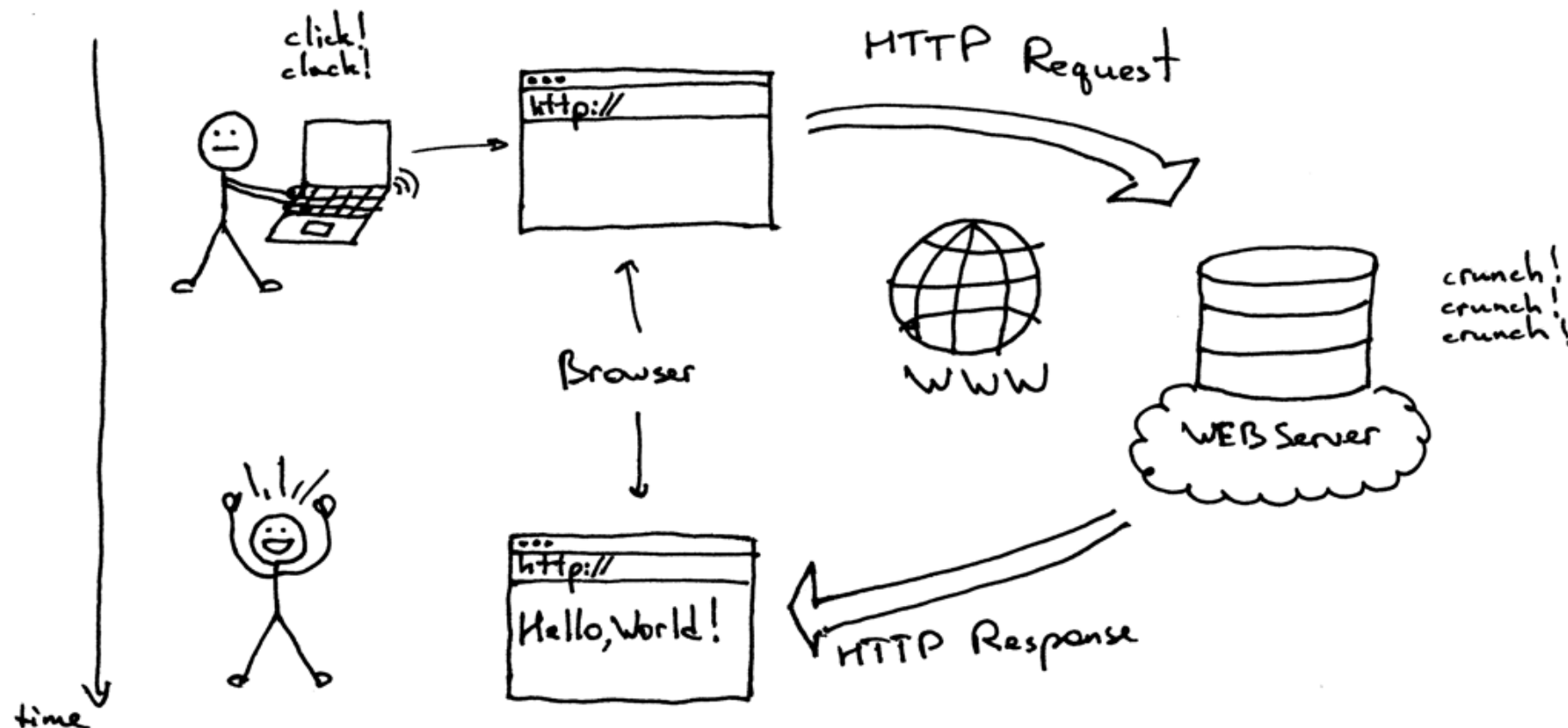
# Introdução

- Estamos numa era em que o bem mais precioso é a **informação**;
- Para enfrentar um cenários cada vez mais **competitivos**, é muito importante que as organizações possuam mecanismos para explorar os **dados** gerados;
- A **partilha** de dados entre aplicações de diferentes **domínios** e **organizações** é cada vez mais uma realidade;

## Armazenamento e transporte de informação na Web

- Em ambiente web, o **transporte** e **armazenamento** de informação assume um papel importante não só para suportar os requisitos aplicacionais mas também para suportar a comunicação entre aplicações;
- Em particular, o transporte de informação requer modelos simples e **flexíveis** o suficiente para se adaptarem a vários contextos;

# Armazenamento e transporte de informação na Web

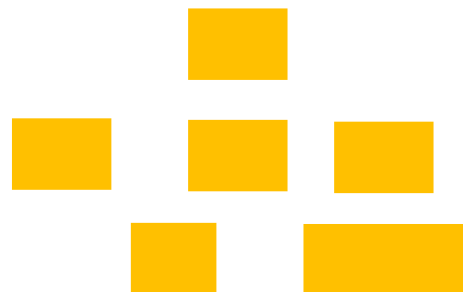


Fonte: <https://ruslanspivak.com/lbaws-part1/>

# Dados Estruturados e Não estruturados



- Dados **estruturados** são dados armazenados sobre uma estrutura de dados **pré-definida**.
- Significa que quando os dados são inseridos no repositório têm **obrigatoriamente** de seguir as regras de armazenamento
- Exemplo: Bases de dados relacionais



- Dados **semiestruturados** sobre uma estrutura flexível: diferentes representações dos mesmos dados podem existir (os dados descrevem a estrutura).
- Exemplo: XML, JSON

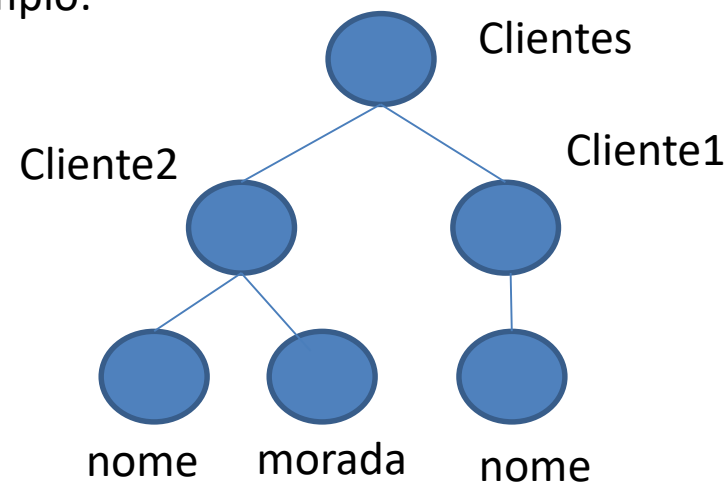


- Dados **não estruturados** são dados armazenados **sem uma estrutura** definida.
- Exemplos: ficheiros Word e PDF, imagens e vídeo

# Dados Semiestruturados

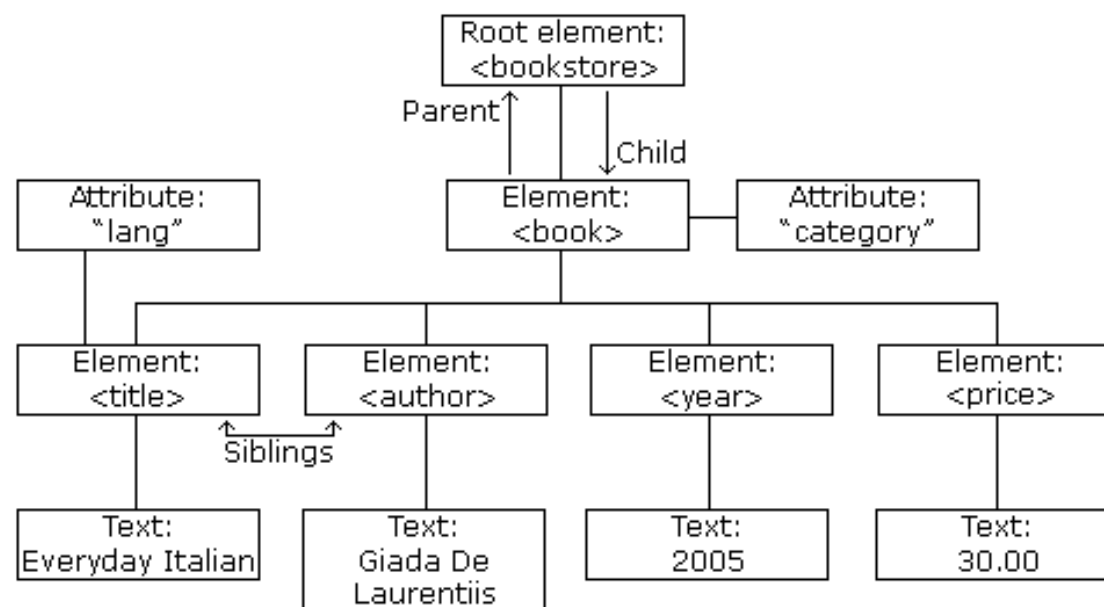
- Dados **semiestruturados** não estão organizados sobe uma **estrutura formal**;
- São compostos por **marcadores** que atribuem **valor semântico** aos elementos do modelo.
- Os **dados** permitem identificar a **estrutura** em que se encontram organizados.
- Os dados encontram-se organizados em **hierarquias**;

Exemplo:



# Exemplo de documento

```
<?xml version="1.0" encoding="UTF-8"?>
<bookstore>
  <book category="cooking">
    <title lang="en">Everyday Italian</title>
    <author>Giada De Laurentiis</author>
    <year>2005</year>
    <price>30.00</price>
  </book>
  <book category="children">
    <title lang="en">Harry Potter</title>
    <author>J K. Rowling</author>
    <year>2005</year>
    <price>29.99</price>
  </book>
</bookstore>
```



# Unidade Curricular de Processamento Estruturado de Informação

- XML (eXtensible Markup Language):
  - Estrutura em árvore do XML
  - Sintaxe do XML
  - Elementos XML
  - Atributos XML
- XSD (XML Schema)
  - O elemento Schema
  - Elementos e atributos XSD
  - Tipos de Dados XSD
  - Restrições XSD
  - Elementos complexos XSD



# Unidade Curricular de Processamento Estruturado de Informação

- Namespaces
- XPATH
  - Nodos
  - Sintaxe
  - Operadores
- XQuery e XSL

# Unidade Curricular de Processamento Estruturado de Informação

- Introdução a JSON
  - Sintaxe
  - Tipos de dados
  - Objetos
  - Arrays
- Gestão bases de dados e coleções utilizando JSON
- Documentos XML e a sua integração prática em ambiente Web

## Unidade Curricular de Processamento Estruturado de Informação

- Formatos XML de suporte a protocolos de serviços Web
- Introdução ao protocolo HTTP;
- Introdução a indexação, processamento e interrogação de documentos;

# Avaliação

- Avaliação contínua:
  - Trabalho prático (60%) + Exame escrito (40%)
  - Trabalho prático contém duas entregas programadas com 15% de ponderação cada, sendo que a apresentação do trabalho tem um peso de 70% (da nota do trabalho).
- Outras épocas:
  - Exame Escrito
- Nota mínima de 7,5 valores em todos os momentos de avaliação;

## Equipa docente

- Bruno Oliveira (aulas TP) - Regente da UC: [bmo@estg.ipp.pt](mailto:bmo@estg.ipp.pt)
- Marco Gomes (aulas PL): [mfg@estg.ipp.pt](mailto:mfg@estg.ipp.pt)

**ESCOLA  
SUPERIOR  
DE TECNOLOGIA  
E GESTÃO**

**P.PORTO**

**Bruno Oliveira**  
2018/2019

**Apresentação de modelos estruturados, semiestruturados  
e não estruturados para o armazenamento de dados**

Processamento Estruturado de Informação