

LAVOISIER

DATA AGGREGATION FRAMEWORK



Nothing is lost,
Nothing is created,
All is <xs:transform>



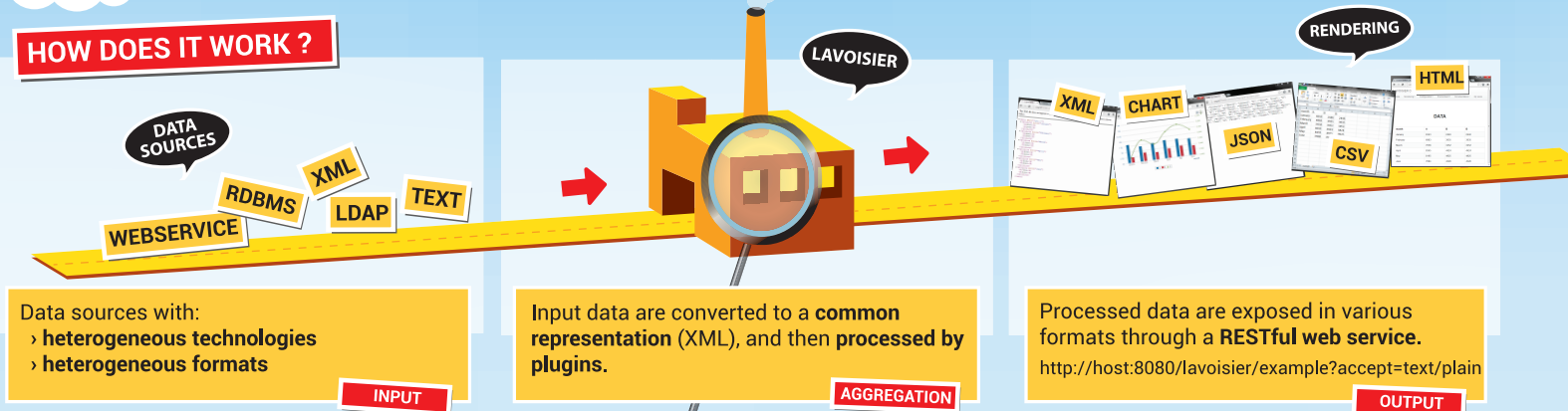
Collecting, transforming, and querying data

from heterogeneous data sources

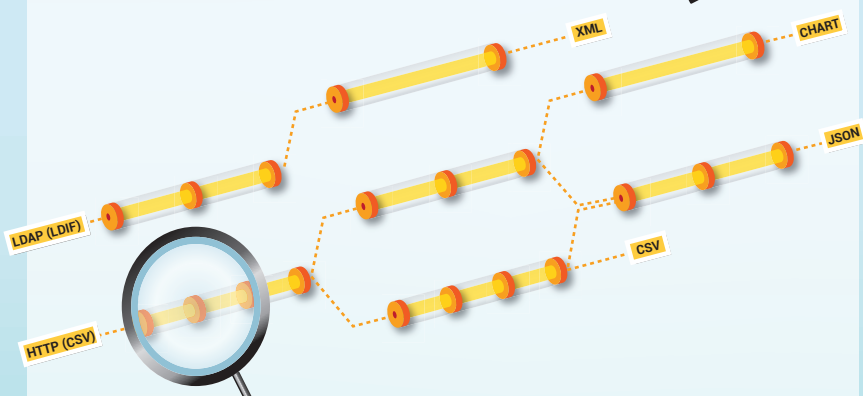
<http://software.in2p3.fr/lavoisier/>

SINCE 2005 !

HOW DOES IT WORK ?



WORKFLOW OF PLUGINS CHAINS



A Lavoisier application is described as a **workflow of plugins chains**. Each plugins chain generates an XML view from input data streams. Each XML view can be rendered to the output format chosen by the end user.

Optimizations :

- > Minimizing **memory consumption** by processing data streams rather than data structures.
- > Improves **data availability, consistency, response-time or network bandwidth usage** by tuning a cache for each XML view.

AGGREGATION

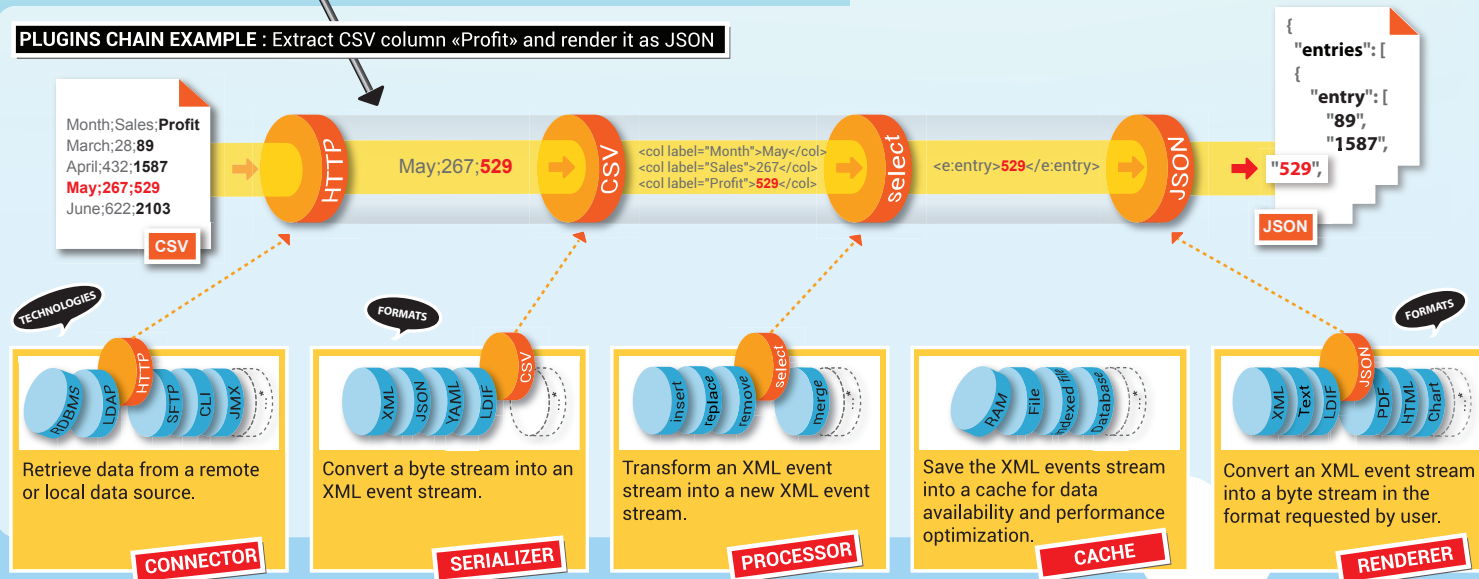
A PLUGINS CHAIN

```
<view name="example">
  <connector type="HTTPConnector">
    <parameter name="url">http://.../file.csv</parameter>
  </connector>
  <serializer type="CSVSerializer">
    <parameter name="header">true</parameter>
  </serializer>
  <processors>
    <processor type="SelectProcessor">
      <parameter name="namespaces">
        <entry key="r">...</entry>
      </parameter>
      <parameter name="nodes">*/r:rows/r:row
        /r:column[@label="Profit"]/text()</parameter>
    </processor>
  </processors>
  <renderers>
    <renderer type="JSONRenderer">
      <parameter name="text-as-pair">>false</parameter>
    </renderer>
  </renderers>
</view>
```

Process for generating an XML view is created by **chaining plugins**. A plugins chain is **described in XML format**.

CONFIGURATION

PLUGINS CHAIN EXAMPLE : Extract CSV column «Profit» and render it as JSON



(*) You can also develop your own plugins

LAVOISIER

FRAMEWORK D'AGRÉGATION DE DONNÉES

Rien ne se perd,
rien ne se crée,
tout se <xslt:transform>



Collecter, transformer et interroger des données
provenant de sources hétérogènes

<http://software.in2p3.fr/lavoisier/>

Depuis 2005 !

COMMENT CA MARCHE ?

SOURCE DE DONNÉES

WEBSERVICE XML RDBMS LDAP TEXT

Sources de données avec :

- Technologies hétérogènes
- Formats hétérogènes

ENTRÉE

LAVOISIER

Les données en entrée sont **représentées** dans un format commun (XML), puis **traitées** par des plugins.

AGRÉGATION

RENDU

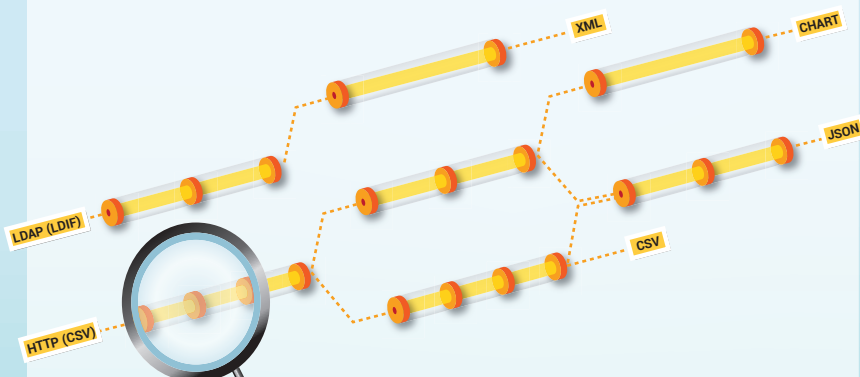
XML CHART JSON CSV HTML

Les données traitées sont exposées sous divers formats par un **web service REST**.

<http://host:8080/lavoisier/example?accept=text/plain>

SORTIE

WORKFLOW DE CHÂÎNES DE PLUGINS



Une application Lavoisier est décrite comme un **workflow de chaînes de plugins**. Chaque chaîne de plugins génère une vue XML à partir du flux de données entrant. Chaque vue XML peut être restituée dans le format choisi par l'utilisateur

Optimisations:

- Minimise la consommation mémoire en manipulant des flux plutôt que des structures de données.
- Améliore la disponibilité des données, leur cohérence, le temps de réponse ou l'utilisation de la bande passante grâce à une configuration de cache pour chaque vue XML.

AGRÉGATION

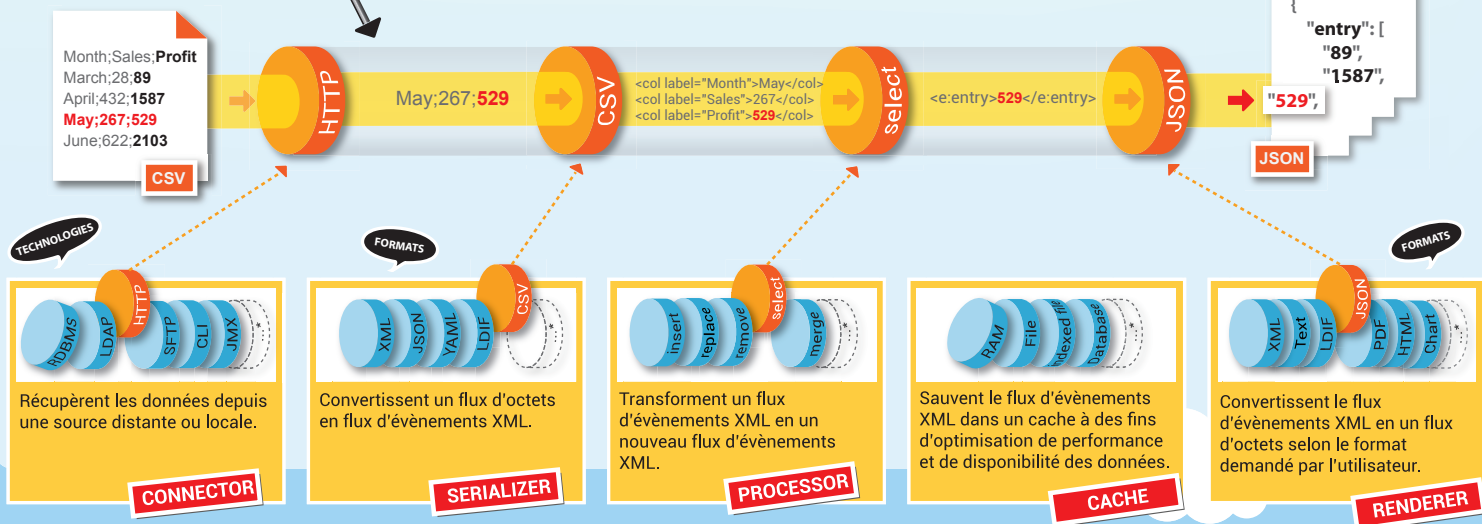
UNE CHÂÎNE DE PLUGINS

```
<view name="example">
  <connector type="HTTPConnector">
    <parameter name="url">http://.../file.csv</parameter>
  </connector>
  <serializer type="CSVSerializer">
    <parameter name="header">true</parameter>
  </serializer>
  <processors>
    <processor type="SelectProcessor">
      <parameter name="namespaces">
        <entry key="r">...</entry>
      </parameter>
      <parameter name="nodes">*/r.rows/r.row
        /r.column[@label='Profit']/text()</parameter>
    </processor>
  </processors>
  <renderers>
    <renderer type="JSONRenderer">
      <parameter name="text-as-pair">>false</parameter>
    </renderer>
  </renderers>
</view>
```

Le processus de génération d'une vue XML est construit en assemblant des plugins. Cet assemblage de plugins se décrit en XML.

CONFIGURATION

EXEMPLE DE CHÂÎNE DE PLUGINS : Extraction de la colonne «Profit» du CSV et restitution en JSON



(*) Vous pouvez également développer vos propres plugins