Harnessing the complexity of large amounts of data is a challenge in itself. But Big Data is more than that: originally characterized by the 3 Vs of Volume, Velocity and Variety, it often requires dedicated computing solutions, which will be explored in this module.

**Prerequisites**
- Engineering Degree on Computer Science or a related domain (telecommunications, etc).
- Engineering Degree on another subject with a major on Computer Science.
- Work experience on Computer science.

**Learning objectives**
After completing this course, participants will be able to:
- Implement the distribution of simple operations via the Map/Reduce principle in Spark;
- Explain the difference between CPU and GPGPU computation;
- Connect on a cloud computing engine (e.g. Google Cloud Platform) and launch a simple task;
- Understand the usefulness of containers;
- Deploy a Docker container.

**Information and registration**
info.exed@isae-supaero.fr

---

**Highlights**
- Introduction to Big Data processing
- Virtualization & cloud computing
- Practical course

**Key elements**
- **Period:** February
- **Estimated duration:** 35 hours, 4 days
- **For whom:** recent graduates, jobseekers and experienced employees
- **Location:** ISAE-SUPAERO, Toulouse
- **Language:** English
Course Content

Distributed computing with Spark:
- History
- MapReduce paradigm
- Hadoop Stack
- Hadoop Distributed File System
- MLlib Machine Learning library

Virtualization and cloud computing:
- Different approaches to virtualization
- Economical models
- Technical benefits (snapshots, dynamic deployment and migration, failover...)
- Cloud engines (principles, deployment examples, node choices)

Docker:
- History,
- Fundamental differences w.r.t. virtualization
- Docker components
- Tools

Teaching methods

<table>
<thead>
<tr>
<th>Teaching methods</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lectures / tutorial</td>
<td>X</td>
</tr>
<tr>
<td>Collaborative learning</td>
<td></td>
</tr>
<tr>
<td>Flipped classroom</td>
<td></td>
</tr>
<tr>
<td>Blended learning (online and face to face)</td>
<td></td>
</tr>
<tr>
<td>Learning by doing</td>
<td></td>
</tr>
<tr>
<td>Project-based</td>
<td>X</td>
</tr>
<tr>
<td>Simulation</td>
<td></td>
</tr>
<tr>
<td>Case study</td>
<td>X</td>
</tr>
</tbody>
</table>

Assessment
- Hands-on evaluation on a computer (100 %)