

Working Group on Security in Big and Open Data (SBOD-WG)

Chair: Alessandra Scicchitano - GEANT

Vice-Chair: Ralph Niederberger – HBP, PRACE and EUDAT

Problem Statement and Scope

Nowadays, 'Big data' and 'Open data' are often-heard buzzwords. Big data refers to large datasets that are not always public. Open data refers to datasets that are not necessarily large but are available to everyone and can be used and republished without restrictions. There are big datasets which have to be accessible worldwide, replicable for security reasons (damage) or with high-speed access (available at different sites to spread download capacity), but accessible by everyone or by distinct people or working groups only. All these types of data may be different, but in some cases they are identical. Large datasets from scientific research sources are a clear example of the overlap between these areas because these datasets can be considered both big and open data.

The Security in Big and Open Data (SBOD) working group focuses on security issues that arise when dealing with big and open data especially within the e-infrastructures. Security issues in this context concentrate on confidentiality, integrity and availability. Confidentiality regulates access to the information, integrity assures that the information is trustworthy, i.e. has not been changed without authorisation, and availability guarantees access to the information by authorised people at any time.

SBOD intends to focus on high level security issues. Issues only specific / pertaining to CSIRTs (computer security incident response teams) are out of the scope of this working group.

WG Phases

The working group will work in phases. Each phase has a limited duration generally of 12 months. At the end of each phase SBOD will re-evaluate itself and prepare a proposal on whether the work should continue or not. The mentioned proposal will be presented to the WISE community and the WISE Steering Committee that will agree or disagree on it.

First phase (March 2016 - February 2017)

1. Work Plan

SBOD will start to list and discuss already existing studies and state of the art in order to build a clear overview of the issues already identified. This will be the starting point for the rest of the work.

Once the state of the art is complete, SBOD will work on a list of issues particularly important for e-infrastructures and on a set of recommendations on how to minimize the impact of these issues.

The SBOD-WG will produce the following items:

- White paper on Security issues in Big and Open data

2. SBOD Working steps:

- Provide a clear definition of what is big and open data for the WG.
- Create a short list of possible use-cases.
- Definition of what is in scope and out of scope of the WG.
- Analysis of all the different approaches used to handle data by the different e-infras and projects.
- List of the existing Security issues and related solutions.

- Analysis of on going projects about data flowing from one e-infra to the other.

All the steps listed above will provide inputs for the whitepaper. In any case each step will be documented where possible and the outcome published on the wiki of the working group: <https://wiki.geant.org/display/WISE/SBOD-WG>

3. Initial involved e-infrastructures and projects

GEANT, EUDAT, HBP, PRACE