

Data protection Classifier

Opportunity for a 6-month internship

Security Research @ SAP Labs France
Sophia-Antipolis – France

SAP's security vision is built on 5 ideals to secure business: Defendable Application, Zero-Knowledge, Zero-Vulnerability, Security by Default, and Transparency.

SAP's security research group lays the foundation for realising the vision: The 30+ researchers of the Security Research unit focus on security engineering (e.g., the automation of the secure software development lifecycle), secure business execution (e.g., business process security and security in cloud based business applications) and secure operations (e.g., secure maintenance and support of complex and heterogeneous cloud IT landscapes).

[Security Research](#) proposes a 6-month internship in its Sophia-Antipolis offices (Mougins, France).

INTERNSHIP TOPIC

A report from 2014¹ claims that human errors account for over 95% of security incidents. New technologies such as machine learning can be leveraged for reducing the risk of human error. One such area where errors could be reduced is misclassification of documents. At SAP, document dissemination obeys strict policies based on their classification level, from 'Public' to 'Strictly confidential'.

Assigning the correct label to documents is a difficult task, as information tends to be less sensitive over time, or can be declassified when certain events such as public announcements are being made.

The goal of this internship is to develop a solution for auto-classifying documents based on their features, reducing the risk of mis-labelling. Some of the expected tasks include:

- Development of a tool able to download documents (MS Word, Powerpoint, PDF) from a repository and extracting its content
- Identification of relevant features for deducing a data classification from the content of the document
- Training of a supervised learning algorithm consuming the extracted content and suggesting a data classification label

Technologies/techniques will include: Bash, python, and Machine Learning

We expect that 40% of time will be dedicated to research activities, and 60% to development

CANDIDATE PROFILE

- University Level: Last year of MSc or less if the student has a good profile
- Good knowledge of one programming language (Python, Java are preferred)
- Good knowledge of versioning control systems like GIT or SVN
- Good knowledge of data science and machine learning algorithms
- Knowledge of textract is a plus
- Interest in research work
- Fluency in English (working language)
- Good oral and written communication skills

¹ https://media.scmagazine.com/documents/82/ibm_cyber_security_intelligenc_20450.pdf

INTERNSHIP CONTEXT

SAP

Over the past 45 years, SAP has grown to become the world's leading provider of business software solutions. With 12 million users, 96,400 installations, and more than 1,500 partners, SAP is the world's largest inter-enterprise software company and the world's third-largest independent software supplier, overall. SAP solutions help enterprises of all sizes around the world to improve customer relationships, enhance partner collaboration and create efficiencies across their supply chains and business operations. SAP group includes subsidiaries in over 180 countries and employs more than 84 000 people.

Security Research at SAP Labs France, Sophia Antipolis

Based at SAP Labs France Mougins, Security Research Sophia-Antipolis addresses the upcoming security needs, focusing on increased automation of the security life cycle and on providing innovative solutions for the security challenges in networked businesses, including cloud, services and mobile.

STANDARD INTERNSHIP PACKAGE

- *Salary*: depending on the length of the internship and your diploma.
- *Lunch*: SAP Labs France has a local cafeteria; interns contribute 2,40 €uro/lunch, like other SAP employees.
- *Holidays*: French Bank Holidays
 - January 1st; April 2nd, May 1st, May 8th, May 10th, May 21st, July 14th; August 15th, Nov 1st and 11th; December 25th
- *Travel*: no trip will be paid by SAP.
- *Accommodation*: SAP can propose an accommodation for the duration of your internship. The accommodation is subsidized by SAP: the intern pays half of the rental cost: 342€ for a 1-room apartment or 442€ for a 2-room apartment (Choice depending on the availability).

CONTACTS AND PROCEDURE

Please send **in English** your CV, a cover letter and any relevant documents to the following persons stating the title of the Internship in the subject: [Internship Application] **Data Protection Classifier**.

Supervisor

Cédric Hebert
cedric.hebert@sap.com
Tel. +33-(0)4-92286348

Administrative point of contact

Sylvine Eusebi
sylvine.eusebi@sap.com
Tel. +33-(0)4-92286477