Ph.D. Objectives

• The work around proxies as part of the PhD of M. Martinez-Peck and C. Teruel [MPBF+15, TWDN15] for object encapsulation. [GN07] proposed different message sends to support ownership. To support ownership, researchers investigated ownership types [CNP01, BSBR03, MYS03, Mil06], introduced capabilities to restrain object properties to support isolation.

• Capabilities. Works such as the ones around E (the language), [MMF01, MS03, MYS03, Mil06], its entirety and in the context of the RGPD.

• Multi-facet data. A single piece of data can have multiple facets (kind, dimension, validation, trigger, ...). In addition, such facetted data should also handle the fact that a piece of data can be nested and/or resulting from the composition of multiple other piece of data. Such resulting complex data should support revocation and ownership/ownership of its subparts. Such facetted data may depend on the user and the user access rights: the same data in addition to its other properties may be invisible, read only, or writable.

We call Smart Data data having such properties (by opposition to big data). Such properties raise several challenges from an implementation point of view in terms of scalability (tracking changes or ownership), implementation (for example of warranty of immutability), and speed.

State of the art within the laboratory

Several international efforts have been carried around elements of the proposal but none of them on its entirety and in the context of the RSPO.

• Capabilities. Works such as the ones around E (the language), [MMF01, MS03, MYS03, Mil06], introduced capabilities to restrain object properties to support isolation.

• Isolation. Caja is an attempt to isolate Javascript applications [MYS03, TEM+11].

• Ownership. To support ownership, researchers investigated ownership types [CNP01, BSBR03, CDN05, COS05], for object encapsulation. [EKN07] proposed different message sends to support dynamic ownership in a dynamically-type languages.

Work in the team

Nevertheless none of them radically rethink object-oriented programming in presence of the properties supporting tractability, ownership, history and revocation. We performed some preliminary study of the topic and developed a first prototype in the context of the CPR Data - 2 [S0202]. The current proposal will revisit such background and take advantage of the knowledge accumulated by the following PhD defended in the team:

• The PhD of C. Teruel entitled "Adaptability and Encapsulation in Dynamically-Typed Languages: Taming Reflection and Extension Methods" and financed by the DGA was about the development of security constructs in dynamic languages [Arn13, ADD+10, ADDT15].

• The PhD of J.B. Arnaud entitled "Towards First Class References as a Security Infrastructure in Dynamically-Typed Languages" was about the development of security constructs in dynamic languages [Arn13, ADD+10, ADDT15].

• The work around proxies as part of the PhD of M. Martinez-Peck and C. Teruel [MPBF+15, TWDN15]...
RGPD is a law to help citizens to control their data. However, there is a challenge to help developers build systems exhibiting by construction such properties. The easier it will be to produce systems like that the more chance we will have that our data are safely managed. The goal of this PhD is to explore and design a language and execution engine where data exhibit properties that support RGPD and Smart Data.

The PhD will develop language constructs and execution engine to support Smart Data. The objectives are:

- Revisit and design object model and execution supporting the definition of multi-facet properties (frame KR-based).
- Define a model of elementary operations to compose, aggregate object while supporting warranty of no modification, history and origin.
- Revisit capability model to support revocation and the smart data properties.
- Define publication mechanisms with no modification warranty based on a block-chain back-end.
- Validate the results on real case studies. PhD will work on:
  - State of the art on revocation, capability, ownership, and support for RGPD.
  - Identify scenarios with sensible data and complex multi player multi access rights.
  - Evaluate first class instance variables and how they can contribute to the design of faceted data.
  - Explore design of object-oriented languages supporting smart data (origin, history, ...).
  - Identify a minimal change algebra supporting smart data properties.
  - Explore design on the light of scalability issues.

**Principales activités**

The PhD will work on

- Identification of key scenarios extracted from real life case studies. We will contact the partners mentioned in the next section.
- State of the art on revocation, capability, ownership, and support for RGPD.
- Identify scenarios with sensible data and complex multi player multi access rights.
- Evaluate first class instance variables and how they can contribute to the design of faceted data.
- Explore design of object-oriented languages supporting smart data (origin, history, ...).
- Identify a minimal change algebra supporting smart data properties.
- Design an architecture to support the execution of such platform.
- Access the proposed solution on the light of scalability issues.

**Collaborations expected**

At the moment of this writing we did not settle formal collaborations on the topic mainly because we want to be more advanced to be able to propose a first collaboration. Nevertheless the topic presented in this proposal has been designed based on discussions and workshops held with the following companies.

- Kertyx/Novartis Dominique Pahud is expert in medical and pharma data. These are the discussions with him that triggered this topic. We will reconnect him and use his expertise to validate our prototypes.
- ZWEIDENKER GmbH is a company interested in the resulting infrastructure (execution engine and results of the PhD). They are interested in applying it in the context of IoT data.

In addition, Utocat is expert in Blockchain and we would like to see how the solution we may propose could raise an interest. We worked with Utocat during two years in the past, but Utocat is a start up and we prefer be able to propose them a first solution to launch a possible collaboration.

**Selected list of publications**

- Jean-Baptiste Arnaud, Marcus Denker, Stéphane Ducasse, Damien Pollet, Alexandre Bergel, and Mathieu Suen. Read-only execution for dynamic languages. In Proceedings of the 48th International Conference Objects, Models, Components, Patterns (TOOLS’10), Málaga, Spain, June 2010
- Ronie Salgado and Stéphane Ducasse. Towards a smart data processing and storage model. In International Workshop on Smalltalk Technologies’19, August 2020
- Camille Teruel, Stéphane Ducasse, Damien Cassou, and Marcus Denker. Access control to reflection with object ownership. In Dynamic Languages Symposium (DLS’2015), 2015
  - Camille Teruel, Erwann Wernli, Stéphane Ducasse, and Oscar Niemiasz. Propagation of behavioral variations with delegation proxies. Transactions on Aspect-Oriented Software Development (TAOSD), pages 63–95, 2015

**References**

Jean-Baptiste Arnaud, Marcus Denker, Stéphane Ducasse, Damien Pollet, Alexandre Bergel, and Mathieu Suen. Read-only execution for dynamic languages. In Proceedings of the 48th International Conference Objects, Models, Components, Patterns (TOOLS’10), Málaga, Spain, June 2010.


Nicholas R. Cameron, Sophia Drossopoulou, James Noble, and Matthew J. Smith. Multi-tuple ownership.
In Proceedings of the 22nd annual ACM SIGPLAN conference on Object-oriented programming systems and applications (ODPSLA’07), pages 441–460, New York, NY, USA, 2007. ACM.

David G. Clarke, James Noble, and John M. Potter. Simple ownership types for object containment.


Donald Gordon and James Noble. Dynamic ownership in a dynamic language.


Mark Samuel Miller, Chip Morningstar, and Bill Frantz. Capability-based financial instruments.


Mark S. Miller and Jonathan S. Shapiro. Paradigm regained: Abstraction mechanisms for access control.


Ronie Salgado and Stéphane Ducasse. Towards a smart data processing and storage model. In International Workshop on Smalltalk Technologies IWSST’20, August 2020.

Camille Teruel, Stéphane Ducasse, Damien Cassou, and Marcus Denker. Access control to reflection with owner ownership.
In Dynamic Languages Symposium (DLS’2015), 2015.


Camille Teruel. Adaptability and encapsulation in Dynamically-Typed Languages: Taming Reflection and Extension Methods.

Camille Teruel, Erwann Wernli, Stéphane Ducasse, and Oscar Nierstrasz. Propagation of behavioral variations with delegation proxies.
Transactions on Aspect-Oriented Software Development (TAOSD), pages 63–95, 2015.

Compétences

Technical skills and level required:

- Languages : French and English
- Relational skills :
  - Autonomy
  - Good communication
- Additional skills appreciated :
  - Agile programming
  - http://www.pharo.org
  - Test driven development

Avantages

- Partial reimbursement of public transport costs
- Subsidized meals
- Leave : 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Access to vocational training
- Possibility of French courses
- Social, cultural and sports events and activities

Rémunération

1st and 2nd year : 1 982€ Gross monthly salary (before taxes)
3rd year : 2 085€ gross monthly salary (before taxes)