**POINTING OUT THE PROBLEMS**

* Current data integration systems imply consuming and integrating data from data services, which deliver data under different quality conditions related to freshness, trust, cost, reliability, availability, among others.
* As known in traditional database approaches, integration tasks are computationally expensive.
  + The idea is that data integration tasks can take advantages from the elasticity and parallel processing provided by multi-cloud architectures.
* Multi-cloud environments bring new challenges to data integration due to different reasons:
  + Different actors are present: data provider, data consumer, infrastructure, and the data itself.
  + Each actor has its own constraints and characteristics that are important to the integration process.
  + The infrastructure, which is the cloud, is actor that unifies the others. The infrastructure is considered in the consumer point of view and in the producer point of view.
  + The consumer and the producer has different constraints over the infrastructure that are considered while processing the integration.
  + The constraints and characteristics of each actor are specified in different SLA contracts. Currently, the SLA contracts define restrictions only over resources. They do not cover these constraints and characteristics.
  + Data services can be deployed differently including different constraints and characteristics in the same cloud. This implies different types of contracts in the same cloud. Considering that data services can be deployed in different clouds, a variety of contracts can be found.
  + Thus, the integration process deals with different matching problems:
    - Matching the *query* with *data provider services* that can produce a result for the query;
    - Matching the *data consumer preferences and* the *quality guarantees* provided by the *data provider*. The *data consumer preferences* are defined over the constraints and characteristics of each actor. They can concern the data itself, the data  
      services and the type of subscription the user (data consumer) has with the clouds;
    - Matching the *data consumer preferences* and *data consumer’ type of subscriptions* - the data consumer may have several subscriptions with the clouds. Each subscription has an associated contract describing the resources limit, the budget, access permissions he has on the cloud, for instance. The integration process requires processing capacity, storage and a budget and the subscription must be in accordance with these requirements;
    - Matching the *data provider services* and *their type of subscriptions.* Similar as data consumers, the *data providers* have subscriptions with the clouds that have associated contracts. Data providers must also be in accordance with the integration requirements in terms of resources limits and budget
  + In summary, the integration process deals with a multi-dimensional matching problem concerning the constraints and characteristics of each actors that are now considering in this new vision of data integration.

**ARTICLE PLAN ACCORDING TO THE CONFERENCE EXPECTATION**

1) Introduction

Provide a clear problem statement,

Outline the research challenges that drive the proposed work.

2) Related works

Discuss the progress beyond the state of the art of the envisioned research outcome compared to current literature and approaches.

3) Approach

Describe the proposed solution and its expected impact.

4) Research plan

The expected research plan.

5) Preliminary results

The preliminary results.

6) Final remarks

Discuss how the work will be evaluated, stating how the suggested solution is different, new, or better than existing approaches to the problem.