
# Interrogator Software Requirements and Specification

prepared by: Daniel Tapiador and GAP
approved by: GAP
reference: GAIA-C9-SP-ESAC-DTP-001-0D
issue: 0D
revision: 0
date: 2011-09-05
status: draft

## Abstract

This is the requirements specification for the CU9 Interrogator software product. Please make comments in this wiki page (`http://www.rssd.esa.int/wikiSI/index.php?title=Talk:GAP:Development:ITG-SRS&instance=Gaia`).

# Document History

| Issue | Revision | Date | Author | Comment |
|---|---|---|---|---|
| D | 0 | 2011-07-29 | DTP | First draft. |

# Contents

# 1 Introduction

This document defines the requirements and specifications that any valid implementation of the Interrogator must comply with for the Gaia Catalogue and Archive (WOM-033). This task has been initiated within GAP (Gaia Archive Preparation) but will be taken over by CU9 once it is formed.

## 1.1 Objectives

The service interface of the Interrogator system will provide a way for different WP (visualization, advanced applications, etc) to access the Gaia archive (not only the catalogue but the observations associated, etc). On the other hand, it will allow different implementations to be developed in parallel (and later on benchmarked) so that proper decisions are made with regard to the technical solution(s) to be adopted for the Interrogator software product.

## 1.2 Scope

This document represents the requirements and specifications for the Interrogator system. It is subordinate to the Gaia Catalogue and Archive (WOM-033), the SMP, and the general requirements on all CUs in the DPAC SSS (WOM-018).

## 1.3 Assumptions

It is assumed from the very beginning that there will not be a single technology that will outperform the others in all the different use cases foreseen for the Gaia archive. Therefore, there will probably be more than one solution in operations so that reponse time for all questions to be made to the archive is reasonable, considering the amount of data to be published and the expectations of the scientific community.

Therefore, a benchmark with a set of TBD questions representing the most important use cases of the scientific community will be built and run against all available implementations of the Interrogator by using the service interface specified in this document. This will obviously be very useful for any client to decide which concrete implementation to choose based on similarity to benchmarked questions, allowing as well an automated redirection to the more suitable system for the questions made by the user.

## 1.4 Applicable Documents

Whenever applicable documents change, a change may be required in this document. The applicable documents are listed here for clarity - the full references are in Sect. 1.5.

| | |
|---|---|
| ESA/SPC(2006)45 | Gaia Science Management Plan (SMP) |
| WOM-017 | DPAC Project implementation Plan (PIP) |
| WOM-018 | DPAC Software System Specification |
| WOM-033 | Gaia Catalogue and Archive Software Requirements and Specification |
| WOM-011 | DPAC Software Engineering Guidelines |
| TL-001 | DPAC Product Assurance Plan |

## 1.5 Reference Documents

**[WOM-018]**, leaders CU1, D.C., 2011, *DPAC Software and System Specification*, GAIA-C1-SP-DPAC-WOM-018, URL http://www.rssd.esa.int/llink/livelink/open/2786798

Dowler P., T.D., Rixon G., 2010, *Table Access Protocol*, Tech. rep., IVOA, REC-TAP-1.0

**[ESA/SPC(2006)45]**, Gaia Project Scientist, 2006, *Revised Gaia Science Management Plan (SMP)*, ESA/SPC(2006)45, URL http://www.rssd.esa.int/llink/livelink/open/2720576

Graham M., R.G., Morris D., 2009, *VOSpace specification*, Tech. rep., IVOA, REC-VOSpace-1.15

Graham M., R.G., Morris D., 2011, *VOSpace specification*, Tech. rep., IVOA, REC-VOSpace-2.0

Harrison P., R.G., 2010, *Universal Worker Service Pattern*, Tech. rep., IVOA, REC-UWS-1.0

**[TL-001]**, Levoir, T., Damery, J., Hoar, J., et al., 2010, *DPAC Product Assurance Plan*, GAIA-C1-PL-CNES-TL-001, URL http://www.rssd.esa.int/llink/livelink/open/2439085

**[WOM-033]**, O'Mullane, W., 2009, *Gaia Catalogue and Archive Software Requirements and Specification*, GAIA-C9-SD-ESAC-WOM-033, URL http://www.rssd.esa.int/llink/livelink/open/2907710

**[WOM-017]**, O'Mullane, W., Drimmel, R., Mignard, F., et al., 2008, *Project Implementation Plan (PIP)*, GAIA-CD-PL-ESAC-WOM-017, URL http://www.rssd.esa.int/llink/livelink/open/2812481

**[WOM-011]**, O'Mullane, W., Hoar, J., Levoir, T., et al., 2010, *Software Engineering Guidelines for DPAC*,
GAIA-C1-UG-ESAC-WOM-011,
URL http://www.rssd.esa.int/llink/livelink/open/2760364

Ortiz I., D.P., Lusted J., 2008, *Astronomical Data Query Language*, Tech. rep., IVOA, REC-ADQL-2.0

Rixon G., G.M., 2008, *Single-Sign-On Profile: Authentication Mechanisms*, Tech. rep., IVOA, REC-SSO-1.01

## 1.6 Definitions, acronyms, and abbreviations

The requirements set out in this SRS follow the labeling scheme:

**CU9-*ITG*-*X*-*SCOPE*-*xxx***

where *ITG* is this WP number, the Software Product Label (e.g. TAP for "Table Access Protocol") or the Module in the case that there are multiple modules in the product, *X* is either S (for **S**cientific), T (for **T**echnical), Q (for **Q**uality Assurance), or M (for **M**anagement). *SCOPE* is a three or four letter scope specification of the requirement following the identified list of possible values of TL-001 whenever feasible. *xxx* is a monotonically increasing counter for every unique combination of *X*-*SCOPE*.

Each requirement is presented with its unique label and a number of attributes in accordance with TL-001 in the following form:

| *CU9-ITG-X-SCOPE-***020** | $C.v$ | Verific. | Status |
|---|---|---|---|
| Description | | | |
| Parent: Parent | | | |

with (see TL-001 for lists of allowed values, meanings and valid ranges):

| | |
|---|---|
| *CU9-ITG-X-SCOPE-xxx* | The unique identifier of the requirement (see above). |
| $C.v$ | Version number of the requirement composed of major ($C$) and minor ($v$) part. |
| Verification | Envisaged validation method of requirement - this will be either AUT for automated or MAN for Manual. The Software Test Plan defines precisely how the requirement is verified. |
| Status | Status identifier. |
| Parent | Higher level requirement or requirements, comma separated list. |

*SCOPE* is a three character scope specification of the requirement, one of :

REQ: concerning requirements,
SPC: concerning specification activities,
DES: concerning design activities,
IMP: concerning implementation activities,
INT: concerning integration activities,
VAL: concerning validation activities,
ACP: concerning acceptance activities,
MNT: concerning maintenance activities,
OPS: concerning operation activities,
FUN: concerning functional requirements,
PLN: concerning organisational requirements (including planning, phases, reviews, etc.),
SUP: concerning support requirements,
PRF: concerning performance requirements (including dimensioning, resource estimation, etc.),
RSK: concerning risk management requirements,
CTL: concerning verification and control activities,
CNF: concerning configuration management activities,
DIM: concerning dimensioning activities.

## 1.7 Acronyms

The following is a complete list of acronyms used in this document. The following table has been generated from the on-line Gaia acronym list:

| Acronym | Description |
|---|---|
| ADQL | Astronomical Data Query Language |
| API | Application Programming Interface |
| AUT | AUTomated |
| CANFAR | Canadian Foundation for AIDS Research |
| CSV | Comma-Separated Value (database output format, e.g., for MS Excel) |
| CU | Coordination Unit (in DPAC) |
| DPAC | Data Processing and Analysis Consortium |
| FITS | Flexible Image Transport System |
| FTP | File Transfer Protocol |
| GAP | Gaia Archive Preparations (DPAC WG) |
| HTML | HyperText Markup Language |
| HTTP | HyperText Transfer Protocol |

| | |
|---|---|
| ISO | International Organisation for Standardisation (Geneva, Switzerland) |
| IVOA | International Virtual-Observatory Alliance |
| JSON | JavaScript Object Notation |
| KPI | Key Performance Indicator |
| LRMS | Local Resource Management System |
| MAN | MANual |
| PBS | Portable Batch System |
| PIP | Project Implementation Plan |
| PaaS | Platform as a Service |
| RDBMS | Relational DBMS |
| REST | Representational State Transfer |
| SGE | Sun Grid Engine |
| SMP | Science Management Plan |
| SOA | Service-Oriented Architecture |
| SQL | Structured Query Language |
| SRS | Software Requirements Specification |
| SSS | Software and System Specification |
| STCS | Space-Time Coordinate serialized as String |
| SaaS | Software as a Service |
| TAP | Table Access Protocol |
| TBD | To Be Defined (Determined) |
| UCD | Unified Content Descriptor |
| UDF | User-Defined Function |
| URI | Uniform Resource Identifier |
| VO | Virtual Observatory |
| WP | Work Package |

# 2 General description and requirements

The Interrogator is the system that will provide access to the Gaia Catalogue. The system will be some kind of a fast parallel query engine (i.e. RDBMS, MapReduce, etc), highly tuned for the type of questions that will be asked, and with a uniform API (specified in this SRS).

As stated in Sect. 1.3, there may be the possibility that more than one technology coexist in operations for different types of use cases. Therefore, It is extremely important that clients of the Interrogator are designed in terms of the functionality specified in this SRS and that they do not make any assumption about the technology that will sit behind the API.

## 2.1 Context

As presented in WOM-033, the interrogation system will be the module accessing the physical storage area for answering questions made by the different clients of the system (final users, advanced applications, etc).

## 2.2 Decomposition

It has been agreed that a customized and possibly improved IVOA TAP service will be used for the early releases of the Gaia catalog. This TAP service will use the Astronomical Data Query Language (ADQL) for defining the way clients will make questions/queries to the system.

There will also be IVOA VOSpace services to export query results and import private data coming from users so that queries can also use custom data (external to the Gaia archive or coming from previously executed queries).

Data intensive workflows will be allowed to run within the archive infrastructure (high bandwidth, low latency data access, etc) by means of PaaS.

Security will be implemented for the appropriate use cases, i.e. authentication of users for importing data into their local space, etc.

## 2.3 General Requirements

| **CU9-TAP-T-DES-020** | 0.1 | MAN | Draft |
|---|---|---|---|
| Software-as-a-Service (SaaS) approach will be followed whenever possible to make the software portable. This will add, among other things, the possibility to run the system in the cloud (both private and commercial ones). | | | |
| Parent: CU9-ARC-M-FUN-020 | | | |

# 3 TAP+

IVOA Table Access Protocol (Dowler P. (2010)) has been identified as the early standard API for the Gaia archive Interrogator. As stated in the standard itself, it defines a service protocol for accessing general table data, including astronomical catalogs as well as general database tables. It also includes support for multiple query languages, including the Astronomical Data Query Language (Ortiz I. (2008)) and generic SQL among others. Special support is given for spatially indexed queries using the spatial extensions in ADQL.

TAP standard architecture follows the Representational State Transfer (REST) web services model which is considered to be faster in execution, but which must be used with caution, clearly defining the resources they can manage and how the data exchanged is represented.

Although current TAP recommendation is thought to be generic enough to allow different kinds of queries over the data published (which can be previously discovered by using the service itself), there are some considerations that are left to concrete implementations for inclusion. In this SRS, those considerations will be clearly set as requirements or not mentioned at all (no need to implement them).

Furthermore, it might be worth adding more functionality to the service so that the exposed functionality is in accordance to the data itself. Requirements have been laid out following this approach, and they specify some constraints which are not mandatory in the IVOA TAP recommendation, as well as areas where some improvements shall be added to the specification.

It is important to remark that *IVOA TAP recommendation will apply* unless otherwise stated in this SRS.

In addition, there are some fields where further investigation will be required once there is more experience in the archive usage by users. Some of these are:

- TAP is supposed to mainly serve relational data to users, but this might not be very well suited for Gaia data which comprises multidimensional data structures not so easily represented in the relational model. It is then needed further investigation in this area to find out whether TAP could be extended to support richer and/or more complex data models.
- TAP service might be seen as a composition/orchestration of several simpler/lighter services (SaaS over SOA). It might be needed to elaborate more on this taking into account the functional requirements laid out here, but as a first idea, there might be one service performing the query parsing and syntactical analysis, another one dealing with the infrastructure provisioning (virtualization services), one instance for building results in the format chosen, etc.

## 3.1 Requirements

| **CU9-TAP-T-FUN-020** | 0.1 | MAN | Draft |
|---|---|---|---|
| TAP+ shall be the (early) standard API for the Gaia archive Interrogator. | | | |
| Parent: CU9-ITG-T-FUN-020, CU9-ITG-T-FUN-140, CU9-ITG-T-FUN-200 | | | |

| **CU9-TAP-T-FUN-040** | 0.1 | AUT | Draft |
|---|---|---|---|
| User must be able to upload data to the archive system (or any other machine in the archive infrastructure) from their workstation or any other data service on the Internet (only VOSpace for now). | | | |
| Parent: CU9-ITG-T-FUN-100, CU9-ITG-T-FUN-180 | | | |

| **CU9-TAP-T-FUN-060** | 0.1 | AUT | Draft |
|---|---|---|---|
| User queries must be able to reference data previously uploaded to the local area available for users or any other data placed in any service accessible through the Internet (only VOSpace endpoints for now). | | | |
| Parent: CU9-ITG-T-FUN-120 | | | |

| **CU9-TAP-T-FUN-080** | 0.1 | AUT | Draft |
|---|---|---|---|
| The service shall allow third-party queries to other TAP or TAP+ services. These third-party queries might be specified within the main query sent to the only TAP+ service the user interacts with. This service would then analyse whether it needs to perform queries to other third-party services (i.e. useful for crossmatching) before actually running the main query submitted by the user. This would allow to select data and/or join several tables that are not being served by the same TAP or TAP+ service without forcing the user to deal with the intermediate results generated by the third-party queries. | | | |
| Parent: CU9-ADV-T-FUN-080, CU9-ITG-T-FUN-100 | | | |

In a scenario where the primary and secondary TAP or TAP+ services live in the cloud, this would reduce costs as no data would leave the cloud when performing the query (no I/O to the Internet for the execution of the query).

| **CU9-TAP-T-FUN-100** | 0.1 | AUT | Draft |
|---|---|---|---|
| Allow the possibility of creating the so-called sessions so that users can run their queries in isolated environments that might benefit from caching mechanisms or allow semi-interactive workflows to take place (the user may want to work with the data interactively from a shell after running the query, without downloading the entire query's result). | | | |
| Parent: CU9-ITG-T-FUN-100 | | | |

| **CU9-TAP-T-FUN-120** | 0.1 | AUT | Draft |
|---|---|---|---|
| The user might not know a priori the time it will take to run their query, so queries lasting longer than a determined amount of time (around one minute) in the interactive (fast) query engine will automatically be transfered to the batch (slower) query engine. That transfer does not mean inmediate execution in the asynchronous (batch) system. The query might be suspended for a while or even canceled and resubmitted if there are no resources available at that moment. | | | |
| Parent: CU9-ITG-T-PRF-080 | | | |

| **CU9-TAP-T-FUN-140** | 0.1 | AUT | Draft |
|---|---|---|---|
| Limitation of the maximum number of objects to be retrieved is currently in place in the IVOA TAP recommendation (Dowler P. (2010)), but the implementation shall also provide a way to specify the offset of the data to be retrieved (for *pagination* purposes). | | | |

| **CU9-TAP-T-FUN-160** | 0.1 | AUT | Draft |
|---|---|---|---|
| Users may specify where to store the results produced by the query. The options are the local workstation of the user, the local space available for users in the archive environment and any other data service (only VOSpace for now) possibly running on another machine (or even another administrative domain). They might also specify the partitioning of the results in several files for easing the download or i.e. for copying to a MapReduce infrastructure. | | | |
| Parent: CU9-ITG-T-FUN-100 | | | |

| **CU9-TAP-T-FUN-180** | 0.1 | AUT | Draft |
|---|---|---|---|
| The system shall provide support not only for VOTable input/output format but also for CSV, FITS, HTML and JSON. | | | |
| Parent: CU9-ITG-T-FUN-160, CU9-ITG-T-FUN-180 | | | |

| **CU9-TAP-T-FUN-200** | 0.1 | AUT | Draft |
|---|---|---|---|
| Users shall be allowed to “publish” data computed by them as a result of their processing. This might allow the concept of a living archive to some extent, as people might publish some of their results which would be available to the whole community. | | | |

| **CU9-TAP-T-FUN-220** | 0.1 | AUT | Draft |
|---|---|---|---|
| The implementation shall allow the possibility to send workflows of queries. This is achieved to some extent within standard TAP recommendation by using the *RUNID* parameter but it does not allow the definition of dependencies among the requests. | | | |
| Parent: | | | |

| **CU9-TAP-T-FUN-240** | 0.1 | AUT | Draft |
|---|---|---|---|
| The system shall allow ADQL and ADQL+ languages as well as standard SQL. | | | |
| Parent: CU9-ITG-T-FUN-020 | | | |

| **CU9-TAP-T-FUN-260** | 0.1 | AUT | Draft |
|---|---|---|---|
| */availability* endpoint should show information about the resources available in the data center (local storage available for the user, free slots for Virtual Execution Environments, etc). It might be worth defining a machine readable format for this info to allow automatic heuristics to be applied depending on data center load. | | | |

| **CU9-TAP-T-FUN-280** | 0.1 | AUT | Draft |
|---|---|---|---|
| */tables* endpoint might be expanded to */schema* to be able to show richer data models (not just relational ones). | | | |

| **CU9-TAP-T-FUN-300** | 0.1 | AUT | Draft |
|---|---|---|---|
| */capabilities* endpoint should produce a machine readable document to allow scheduling heuristics of queries/jobs. The contents shall specify the services available, back-end implementations in the data centre (RDBMS, MapReduce, etc) and any other information useful for learning more about the system. | | | |

| **CU9-TAP-T-FUN-320** | 0.1 | MAN | Draft |
|---|---|---|---|
| Root web resource must return a user friendly page with some basic functionality (i.e. text box for sending queries, etc) as well as samples and help contents. This page might be a link to the main page of requirement CU9-CIF-T-FUN-020 specified in WOM-033. | | | |
| Parent: CU9-CIF-T-FUN-020 | | | |

| CU9-TAP-T-FUN-340 | 0.1 | MAN | Draft |
|---|---|---|---|
| Unless otherwise stated here, IVOA Single-Sign-On profile (Dowler P. (2010)) shall be followed for implementing security mechanisms. Client certificates should be hidden as much as possible when any kind of security is needed (for authentication, authorization, etc), although they will be used for accessing certain services (overall those coming from Grid technology). Advanced/experienced users should be given the possibility of getting/downloading their own certificates and delete them from any other place to work only with proxies generated by them locally (the ideal way to work in terms of security). | | | |

| CU9-TAP-T-FUN-360 | 0.1 | AUT | Draft |
|---|---|---|---|
| Data upload to the archive where the user supplies the code that parses the data. This parser will have to obviously comply with a defined interface or so. This way, 'GBIN' would be a valid format for uploading data into the archive. The same for producing results in different formats. Parsing or results generation software could be provided as UDFs or MapReduce programs to be plugged in the appropriate backends. This would require a lot of standardization which might be dangerous in terms of complexity. | | | |
| Parent: CU9-ITG-T-FUN-160, CU9-ITG-T-FUN-180, CU9-ITG-T-FUN-100 | | | |

| CU9-TAP-T-FUN-380 | 0.1 | MAN | Draft |
|---|---|---|---|
| There should be a notification mechanism (by e-mail and/or any other more complex service notification system) for avoiding polling the status of asynchronous requests, allowing batch queries and workflows to be optimally executed. | | | |

| CU9-TAP-T-FUN-400 | 0.1 | AUT | Draft |
|---|---|---|---|
| Coordinates transformations shall be made by the service (as specified in the TAP protocol specification), allowing the user to specify the geometrical regions at their wish. | | | |

### 3.1.1 Performance Requirements

| CU9-TAP-T-PRF-020 | 0.1 | MAN | Draft |
|---|---|---|---|
| The minimum and maximum amount of physical or logical resources needed for the operation of the system, as well as the elasticity rules and Key Performance Indicators (KPI) shall be clearly identified. | | | |
| Parent: CU9-ITG-T-PRF-060 | | | |

| **CU9-TAP-T-PRF-040** | 0.1 | AUT | Draft |
|---|---|---|---|
| The system shall provide the user with the ability to suspend/resume query execution as well as all the actions specified in the Universal Worker Service Specification (Harrison P. (2010)) which any TAP service must comply with, like run, abort, etc. The asynchronous endpoint will then become a batch queuing system very similar as those provided by LRMS (Local Resource Management System) for job execution like SGE, PBS, Condor, etc. Indeed, some implementations might rely on them for providing this functionality. | | | |
| Parent: CU9-ITG-T-PRF-080 | | | |

| **CU9-TAP-T-PRF-060** | 0.1 | AUT | Draft |
|---|---|---|---|
| The service must allow the *permanent* redirection of a client to another service instance (through HTTP 300 responses or so). This would allow better load balancing, scalability and services composition mechanisms (for SaaS). This requirement is related to Req. CU9-TAP-T-FUN-100 in the sense that the user shall send the query to the frontend service, being permanently redirected to another isolated environment which will provide the session requested. This functionality would also give the possibility of redirecting certain types of queries to other service implementations that are more suitable for the query submitted by the user. The user shall have of course the possibility to disable this behaviour per query basis. | | | |

| **CU9-TAP-T-PRF-080** | 0.1 | AUT | Draft |
|---|---|---|---|
| The service must allow the user to specify the system that is going to run the query, i.e. RDBMS (like Intersystems Cache, Vertica or PostgreSQL), MapReduce, etc. Although most of the systems will provide only one implementation, there might be cases where two or more back-ends are put in place, or the service has knowledge of other services with different implementations so that different queries might be redirected to the most appropriate back-end in terms of performance or other service provisioning related matters (service availability, etc). This requirement will allow clients to specify the system implementation that should run the query. The service will return an error if that implementation is not available (locally and/or remotely in other registered services). Similarly, the user might specify the system for which the uploaded data must be available (i.e. it is not the same to upload data to a RDBMS or a MapReduce system). | | | |

# 4 ADQL+ and STC-S

IVOA Astronomical Data Query Language or ADQL (Ortiz I. (2008)) has been identified as the early standard query language to be used for specifying queries against the Gaia archive. ADQL is based on SQL-92 with some restrictions and extensions in order to support generic and astronomy specific operations.

It provides a convenient way for representing a wide number of queries, but there might be some scenarios not covered by the standard which might enhance the acess and analysis of Gaia data. Those will be specified through requirements in this section.

As in TAP+ section (Sect. 3), *IVOA ADQL recommendation will apply* unless otherwise stated in this SRS.

Areas where further investigation will be needed are:

- The language must allow the usage of relational data fields but there should also be a way to use multidimensional data structures contained in tables in the queries.
- Timestamp values (ISO 8601) accuracy is ms, but it will probably be needed more accuracy for Gaia data.
- ADQL query result must be a single table (in a single file). It will have to be investigated whether this is a problem for the Gaia archive use cases.

## 4.1 Requirements

| **CU9-ADQL-T-FUN-020** | 0.1 | AUT | Draft |
|---|---|---|---|
| ADQL+ shall be the (early) standard query language for the Gaia archive Interrogator. | | | |
| Parent: CU9-ITG-T-FUN-020, CU9-ITG-T-FUN-140, CU9-ITG-T-FUN-200 | | | |

| **CU9-ADQL-T-FUN-040** | 0.1 | AUT | Draft |
|---|---|---|---|
| The language shall contain spatial extensions to support 3D regions (at least sphere and 3D box) as well as functionality over them (*CONTAINS*, *INTERSECTS*, *LINEAR-DISTANCE*, etc). | | | |
| Parent: CU9-ADV-T-FUN-020, CU9-ITG-T-PRF-020 | | | |

| **CU9-ADQL-T-FUN-060** | 0.1 | AUT | Draft |
|---|---|---|---|
| The syntax shall provide a way to specify a query over different regions with *UNION* and *INTERSECTION* predicates (give me some data that comply with these constraints and is placed in these regions). This is already defined as *informative* within TAP (Dowler P. (2010)) "Use of STC-S in TAP" section. Spheres and 3D boxes shall also be valid geometrical figures as stated in Req. CU9-ADQL-T-FUN-040. | | | |
| Parent: CU9-ADV-T-FUN-020, CU9-ITG-T-PRF-020 | | | |

| **CU9-ADQL-T-FUN-080** | 0.1 | AUT | Draft |
|---|---|---|---|
| Gaia meassurements, magnitudes and columns shall be univocally identified with VO UCDs and proper units so that semantics are present in the different fields as much as possible. This metadata shall be available for querying, and therefore geometry functions will be able to undertake proper actions for converting coordinates to the relevant *coordsys* in case of need (see Req. CU9-TAP-T-FUN-400). | | | |
| Parent: CU9-DOC-S-FUN-040, CU9-ADV-T-FUN-080 | | | |

An open question is whether coordinates should be transformed if a target is specified in one epoch and there are other coordinates to crossmatch with in another epoch.

| **CU9-ADQL-T-FUN-100** | 0.1 | AUT | Draft |
|---|---|---|---|
| • *SELECT INTO "AnotherNewlyCreatedTable"* clause shall be permitted and this newly created table shall be stored in the local user space. The user shall be authenticated by using Single-Sign-On IVOA mechanisms (see Req. CU9-TAP-T-FUN-340) and all data stored in their limited local space area will be available later on for other queries.<br>• *SELECT INTO URI "VOSpace endpoint" WITH FORMAT "VOTable, FITS, CSV, etc"* might also be another use case to consider adding.<br>• *SELECT field1, field2, ...FROM "VOSpace endpoint" WITH FORMAT "VOTable, FITS, CSV, etc"* is the counterpart use case to the previous one which allows the user to upload data (in any format) to their local space area and make use of it automatically in the query.<br>• It is important to remark that inactive accounts might be tagged for deletion although the user should be informed about it beforehand. | | | |
| Parent: CU9-ITG-T-FUN-100, CU9-ITG-T-FUN-120, CU9-ITG-T-FUN-160, CU9-ITG-T-FUN-180 | | | |

# 5 VOSpace

Similarly to other previous sections (Sect. 3 and Sect. 4), IVOA VOSpace (Graham M. (2009)) has been identified as the early standard data service to be used for exchanging (downloading and uploading) data from/to the archive.

VOSpace is the IVOA interface to distributed storage. It specifies how VO agents and applications can use network attached data stores to persist and exchange data in a standard way. Although the first version of the standard has already been issued (Graham M. (2009)), at the time of writing, a new version (Graham M. (2011)) is underway, which will provide a more convenient REST interface.

*IVOA VOSpace recommendation (Graham M. (2009)) will apply* unless otherwise stated in this SRS.

Considerations of VOSpace implementations for Gaia taking into account current and future IVOA VOSpace recommendations will be laid out in this section as a set of requirements (i.e. Transport protocols implemented shall be HTTP, FTP, etc).

For the initial implementation of the Interrogator, local space for user tables is enough.

## 5.1 Requirements

No requirements have been defined yet.

# 6 PaaS

Considerations of Platform as a Service (PaaS) for “bring the software to the data” use cases.

It is foreseen (at the time of writing) a collaboration with CANFAR to evolve some of the services they have implemented so far, which may help develop the system required to allow the use case presented in this section.

## 6.1 Requirements

Requirements will be written here once it is clearer what we need.

# 7 Security

Security considerations (authentication, authorization, digital signature, etc) for Gaia Interrogator service. The basis for this should be the Single-Sign-On IVOA recommendation (Rixon G. (2008)).

There have been some problems recently reported by the VO community with regard to security. Apparently, standard TAP metadata queries/responses look like SQL injection attaks and trigger flags in security monitors. This will have to be further investigated so that a proper solution is found for TAP+.

## 7.1 Requirements

Requirements will be written here once it is clearer what we need.

# A Traceability to upper level requirements

The following table shows the traceability matrix from the Interrogator requirements shown in this document to the applicable requirements in the Archive SRS (WOM-033). It also shows the dependencies among requirements of this SRS:

| Parent Requirement | Requirements in this document |
|---|---|
| **CU9-ADV-T-FUN-020** | CU9-ADQL-T-FUN-040, CU9-ADQL-T-FUN-060 |
| **CU9-ADV-T-FUN-080** | CU9-TAP-T-FUN-080, CU9-ADQL-T-FUN-080 |
| **CU9-ARC-M-FUN-020** | CU9-TAP-T-DES-020 |
| **CU9-CIF-T-FUN-020** | CU9-TAP-T-FUN-320 |
| **CU9-DOC-S-FUN-040** | CU9-ADQL-T-FUN-080 |
| **CU9-ITG-T-FUN-020** | CU9-TAP-T-FUN-020, CU9-TAP-T-FUN-240, CU9-ADQL-T-FUN-020 |
| **CU9-ITG-T-FUN-100** | CU9-TAP-T-FUN-040, CU9-TAP-T-FUN-080, CU9-TAP-T-FUN-100, CU9-TAP-T-FUN-160, CU9-TAP-T-FUN-360, CU9-ADQL-T-FUN-100 |
| **CU9-ITG-T-FUN-120** | CU9-TAP-T-FUN-060, CU9-ADQL-T-FUN-100 |
| **CU9-ITG-T-FUN-140** | CU9-TAP-T-FUN-020, CU9-ADQL-T-FUN-020 |
| **CU9-ITG-T-FUN-160** | CU9-TAP-T-FUN-180, CU9-TAP-T-FUN-360, CU9-ADQL-T-FUN-100 |
| **CU9-ITG-T-FUN-180** | CU9-TAP-T-FUN-040, CU9-TAP-T-FUN-180, CU9-TAP-T-FUN-360, CU9-ADQL-T-FUN-100 |
| **CU9-ITG-T-FUN-200** | CU9-TAP-T-FUN-020, CU9-ADQL-T-FUN-020 |
| **CU9-ITG-T-PRF-020** | CU9-ADQL-T-FUN-040, CU9-ADQL-T-FUN-060 |
| **CU9-ITG-T-PRF-060** | CU9-TAP-T-PRF-020 |
| **CU9-ITG-T-PRF-080** | CU9-TAP-T-FUN-120, CU9-TAP-T-PRF-040 |

# B SaaS guidelines

Generic guidelines for building SaaS applications will be written here soon.