

## **GENIUS KOM**

### **First day 04-12-2013**

9:30 – 11:00

Participants: Nuria Benitez (NB), Traian Branza (PO), Ricky Smart (RS), Nigel Hambly (NH), Gonzalo Gracia (GG), Anthony Brown (AB), Mark Taylor (MT), Daniel Hestroffer (DH), Frédéric Arenou (FA), Alfred Gil (AG), Frances Julbe (FJ), Jordi Torra (JT), Eduard Masana (EM), Xavier Luri (XL), Enrique Solano (ES), Lola Balaguer (LB).

Webex Connection: Andre Moitinho (AM), Dimitri Pourbaix (DP), Laurent Eyer (LE) and Nami Mowlavi (NM), Yoshiyuki Yamada (YY)

**Xavier Luri**, welcome and agenda (see slides on the Twiki page)

**Traian Branza, Project Officer** (PO) Guidelines for the good start of the Project (see presentation on the Twiki page):

- 550 people in REA
- 630M€ for space research
- 100M€ per call for space research.

The EC marks the policy, REA its implementation and ESA takes part in the policy, expert assistance and the technical review of the projects.

Moreover, REA is in charge of evaluation, negotiation and technical and financial management of projects.

On the Space work programme: 78% of the projects are rejected!, 12% in the reserve list and only 10% gets the financing

Good practices: The Coordinator should send an email to the PO every 3 months.

Important All Annexes of the Grant Agreement (GA) are in our Twiki (LB),

PO spoke about the Limits for Money reporting accounting and all the information on the Annexes of the GA. The PO will assess the timing of the deliverables in case of delays. There will be an Annual Technical Review Meeting (after 12, 24 and 42 months). In this meeting the PO will chose some External reviewers (still not known, probably from ESA).

Report: clear accounts and precise information. It is very important to keep timesheets. At any moment we can have an audit.

The reports should be clear informative, concise, transparent and on time.

Prefinancing already delivered. The amount received is a 53% percent (NB pointed out).

There is a Guarantee Fund (5%) of the money that will not be transferred until the final completion of the project and the reports. Once the completed reports, forms C and all deliverables are accepted, the payment will be transferred in a maximum of 105 days.

The amounts received by the parties can be transferred between different activities and even between partners, always upon prior agreement of the PO.

Questions and comments for the PO:

NB confirmed the money has been already distributed to all partners, after a small problem with Japan. YY confirmed they already got it.

XL reminds that the timesheets need to be filled by all partners

NH asks about VAT being not eligible. Are any other taxes included on that? Indirect tax can be claimed, but VAT is not eligible. Airport taxes are eligible. Only Added value tax is not (answered by NB).

MT asks about constraints on travel money. Can relevant meetings that are not directed to GENIUS but related, be funded? PO answered no (if it is not in the original proposal).

XL summarised: travel money strictly limited to the initial meetings planned in the budget. Travels are very limited (PO said). XL adds that the type of travels was already described on the original proposal. Not scientific meetings, whatever you said on the project must be followed (XL).

RS asked about justification of consumables. PO regarding the justification of consumables, hard discs, PC's, Laptop's are part of indirect costs and normally these costs cannot be accepted as consumables. Moreover, consumables are very limited for this project, only 667€ are planned by UCAM for hard disks. Been a quite low cost, this was accepted during negotiation. NB also says hard discs are not consumables. You need to ask for prior justification if the amount is high, if it is not much it should go with the depreciation rules in your institution. You will not get the total amount but the proportional amount covered by the depreciation rules of your institution. RS said in Italy hard disks are considered consumable. NB: as a general rule, anything that can be used for more than two years is not consumable.

**Nuria Benitez, UB – FBG Administration Issues** (see presentation on the Twiki pages)

NB introduces the UB and FBG roles, some figures about proposals and experience.

FBG is a foundation integrated and controlled by the UB to manage all the international FP7 projects. The main reasons for a foundation are the staff specialization of the work and the less amount of bureaucracy and more agility on the contracts and payments.

General framework: the FBG unit is divided in two: the promotion section dedicated to the applications. And the financial and management section which is in charge of the projects once they are successful.

This Project overview: Duration of the Project from October 2013 to March 2017, with three reporting periods: P1 12 months, P2 from 13 to 24 and P3 25 to the end.

Project Budget: Be aware of the need of reporting the total cost to get the EC contribution. There is only need to Audit for three partners (CNRS, UL, UB).

Annex I Work Packages. Every expense you do should go associated to a task. Distribution of effort: be sure you understand your tasks and effort assigned.

## Financial statements

Form C should be delivered 60 days after each period. EC has 105 days to pay..

The FBG will do an internal check of budgets every six months, just to check everything is running correctly, and the necessary changes are done with prior agreement. Activities and costs should be clearly related.

Controls: It is necessary to get a CFS for three of the partners. To be eligible cost, all bills should go on the name of the entities.

Reporting: Every partner should fill a Form C + use of resources + Audit (if necessary) and then sent it to the coordinator who should consolidate and transfer to the REA.

A final payment of the last 15% will happen at the end of the project once everything reported and accepted.

XL points out that this is an issue for the institution that should advance the full amount of money to the scientists.

All partners should do all the reporting periods: middle term and final.

Schedule of the finance. Your institution should pay your activities following your duties.

PO: the timing is very important: 42 months only. Extensions are not accepted mostly. Plan well your expenses. XL points out that the delays can come from the actual Gaia mission.

We have a new ECASS reporting system. The financial part of the new webpage does not exist until next week. PO: the new tools will not be available until beginning of 2014.

NB continues showing the old tools: remember each expense goes related to one work package on the format.

XL: we were beta testers for the submission of the proposal. It is the new system beta already consolidated? Hopefully, problems have been solved (PO, NB).

NB+LB: The originals of the final signed Consortium Agreement will be now given to each partner.

## **Coffee Break**

**Short presentations of the partners teams**, see slides on the Twiki.:

Xavier Luri: Presentation of UB team. Around 30 people working in DPAC including students.

Anthony Brown: Presentation of Leiden node: the institute (25 staff) and Gaia & GENIUS in Leiden. 2 EU financed PhD and postdoc (Risquez, Prodhomme)

Frederic Arenou (CNRS). CNRS umbrella for 3 institutes: Obs. Paris/GEPI, Obs. Paris/IMCCE, Besançon/Utinam. Description of institutes and involvement, their background in Gaia. Found some problems for recruiting people in Besançon so far. DH suggests to

share within the group information on vacant positions and any profile of potential candidates.

Nigel Hambly (UEDIN) presented the WP300 people (more than the UEDIN itself). Relation with DPAC work packages. Proposal to use Tasks sheets as encouraged by GG from the Project Office.

Ricky Smart (INAF) presents INAF team: Alessandro Spagna and RS at Torino and 4 more people. WP330: VO Dance to convert general data to make it VO compatible. Additions to TAP procedures. Testing on Gaia size database. WP240 cross multiwavelength archives. They will hire a new person. Two approaches suggested where they already have expertise: dense fields (globular clusters and Magellanic Clouds) and multiwavelength cross match.

Alfred Gil (CESCA) explained the history of CESCA and its foundation. Began as a supercomputer center and now it has a wider range of services to the community. They collaborate in simulations, data mining and web portal.

Mark Taylor (UBR) works on WP440 on the assurance for the compliance of Topcat tool with the Gaia data. He will devote his six first months to dedicate an effort of 50% from 01-06/2013.

Enrique Solano (CSIC): GENIUS activities are related to the SVO. Three WP: 3, 4 and 5. Identify Gaia standards to comply with VO.

Andre Moitinho (FFCUL Webex). U Lisbon people works through FFCUL foundation. Three people in Gaia. Visualization services and infrastructure. Already three projects in Visualization issues.

Yoshiyuki Yamada (KU Japan). Presentation of Jasmine project (Nano- Jasmine launch on Dec 2014, one year mission). The mission Small Jasmine is set for approval next Feb 2014. They have five members in GENIUS. Collaboration with Gaia began in 2007. Jasmine missions are similar to Hiparcos and Gaia to cover only bright stars (not observed by Gaia). Gaia can be improved with the combination of Hiparcos and Jasmine. Small Jasmine will be combined with SDSSIV (APOGEE). There will be a general meeting of GENIUS in Japan. RS ask if NanoJasmine is on time. It will be launched in mid 2015. XL points out there have been problems with the Brazilian launcher and that is the main reason of the delay.

Dimitri Pourboix (ULB Webex). No slides. Their work inside CU9 will be developed by the 40% of 1FTE on one year dedicated to validation tasks. And himself, Involved also in GREAT SC.

Laurent Eyer (UG) described their Institute and their participation in the Gaia DPAC CUs. Involvement in Validation and Visualization. Local funding for CU9 is still pending from Switzerland funds.

XL: Cambridge presentation not available, cause personal reasons (Nic Walton father passed away). UCAM will take care of the Alerts science. UCAM is involved in photometry, and other collaborations. NW is the leader of GREAT ITN. Condolences to Nic Walton from XI and the team-

NB: we have given away the Consortium Agreement except for the people not present. XL we will put the contact of NB on the web: [nbenitez@fbg.ub.edu](mailto:nbenitez@fbg.ub.edu)

XL question to PO: how are the review meetings on the three periods? Where do they take place? Is attendance mandatory for WP managers? It consists on the presentation of every WP from the managers and the Q&A with the externals reviewers. It is a one day meeting in Brussels. The four main partners should attend and maybe some of the others on a case by case scenario. XL questions if this is considered as management activity? NB answers that should be as stated in the proposal. LB: only UB has a budget for management so every other partner should do as stated on the proposal as RTD activity.

### **Lunch break**

**XL. Quick overview GENIUS in the Gaia context** mainly directed to our PO.

ESA builds and operates the satellite but does not do science with the data and does not do the reduction of the data. That was designed to be in the hands of a scientific community. Around 450 scientist and 25 institutes form the DPAC consortium led by AB. The consortium is regulated by a multi lateral agreement (MLA) that explains the structure of the DPAC, the Coordination Units CUs, data processing centers DPCs plus an executive DPACE and a Gaia scientist team GST.

The schedule:

- Launch of Gaia on the 19 December 2013
- First release for the Gaia data will be launch + 22 months
- Second release six more months (28 months after launch).
- Third release will be almost inside the GENIUS project.
- From 4<sup>th</sup> release and final in 2022-20023, GENIUS project will be already finished.

ESAC is speculating about possibilities for a future “cloud” system, but this is in the future.

February – March 2014 will be the beginning of nominal observations of Gaia.

Gaia CU9 Software Development Plan already includes GENIUS contribution.

XL comments probably experts for the project evaluations will be part of ESA. Some of ESA people have already been consulted about this project.

NH: there is a package to see how existing tools can be adapted for their Work Packages. XL: there will be a discussion on the needs of the database and NH contribution will be appreciated, from the point of view of the user but also from the technical point of view.

XL: there will be a prototype of the database for users to be able to test user cases, as a testbed for algorithms. Will keep NH updated on that.

### **Xavi Luri. WP100 Management.**

Regarding deliveries, the first is the KOM. No way of doing it on the first month, in our case was October but PO assures no problem with that.

We need to write a partnership agreement (different for CA) for only internal use. AB asks if we need to include legal advisors but XL answers that it is not necessary.

Gender issues: Be aware of the need of recruiting women.

NH asks for Openness and recruiting rules when you have already a team working. No way of changing their recruitment as no woman on the team.

Gender issues: work life balance. Try to travel minimum possible.

Organize conference with women speakers in the field. Promote women role on making an effort for inviting women.

Structure of GENIUS:

Need of appointing the External Advisory Board. Discuss the schedule.

DP points out that teleconference are not that reliable, maybe we need to spend some money to get a more reliable system of teleconference. XL says we are thinking into buying a Webex license. DP wants better microphones. DH says that we will need some physical meetings too. XL agrees and we should try to combine CU9 meetings with GENIUS meetings. For the review meeting DF offers the ULB facilities being in Brussels to organize a larger meeting.. XL: that would be a good idea to minimize travelling. NH insists that remote attendance is a problem with a bad sound. Find a good microphone is important: at least one per participant. There must be some simple cost effective solution for the sound problem. DP one microphone for all is difficult in a full room. NH surely there is a solution. XL noted. **Action** for DM: find a solution.

**Anthony Brown. WP200 Tailoring to the end user community:**

- WP220 Advance requirements to the Gaia archive, Requirements specification.
- WP230 Main idea is future proofing the Gaia archive and catalogue.
- WP240 Seamless data retrieval across archives: Gaia ESO, Plank, Euclid, LSST...Be able to get all of them without even thinking, more than cross matching. It should have the "feel" of a single archive.
- WP250 The living archive: to be able to update to new information after 2022.
- WP260 Reprocessing of archived data. Reprocessing of Gaia data with better information or getting to process different groups of stars together, Keep it readable for long time. Processing software should be available and able to be run.

Questions: DH: ESA should also keep intermediate calibration data. XL comment from last week meeting: discussion about the contribution by INAF, metadata catalogue with the link to other catalogues and cross matched, ESA also wants that. Jesus Salgado (ESAC) proposed to have a dedicated telecom (GG): How to build a tool for that.

**Nigel Hambly. WP300 Aspects of archive system design.**

The tasks of this WP fit into WP300 of CU9, in collaboration with ESA's Science Archive Team at ESAC. Therefore the task in this WP works around the archive system designed by this team, the Gaia Archive Core System. The SAT team cannot afford to coordinate a wide group of teams working on architecture issues, but it is open to establish clean interfaces for

contributions. VO will be one of the areas where GENIUS will contribute: VO tools, web 2.0 client-side interfaces, Infrastructural elements (TAP factory), generalised catalogue cross-querying using TAP. Enable cross database query: on-the-fly, transparent to the user. Intelligent Distributed query processing and OGSA-DAI. ESA will not be able to do all this. Everybody is using the TAP protocol nowadays. But it is necessary something more intelligent on the optimization and minimising for the cross match: on-the-fly distributed cross matching datasets. This is very ambitious but the kind of things we need to be looking to in a project like GENIUS. VO space stored somewhere (in the cloud) available as a prototype but need more enhancements for the Gaia archive system. WP934 to show the kind of things you can do now. Cross-match scaled up. You can do things like that in database stored locally but need to scale out for arbitrary catalogues. Now the limitation is the need to store locally the databases. And there is a limit on the number of databases that can be stored locally, so the need for a solution to avoid local storage. One solution is the pre-cross matching but the genuine solution need to be for a user that has its own database and the need for an on-the-fly cross matching of a billion scale sources. Scheme for a DQP distributed Query Processing to finish.

XL asks about the queries. Be sure to keep Jesus Salgado on the loop for this on-the-fly cross matching. NH: Sure, we should be able to envisage this in the near future with the power of computers. This is an issue that has been in VO community for a while. AB of course that is what the user requirements want. If you find ways to visualizing things in a higher number of parameters, people would ask for it. NH thinks typically the queries are in restricted areas of parameters/cases so it is possible to do it on-the-fly, even with user supplied data. AB. The queries would be expanded if the visualization/queries allow that.

DP will leave at 5pm.

### **Coffee break**

#### **Xavier Luri. WP400 Tools for data exploitation:**

It comprises four subpackages/tasks: Visualization tools, data mining, developing of VO tools, and development of tools for Grand Challenges.

WP420 Visualization (led by FFCUL). They have a very ambitious aim. Develop an object server that has visualization tool with direct access to the archive. You can have a 3D flyby on-the-fly and add more dimensions. Interactions with the requirements unit and with NH group on the architecture of the visualization. Object server, so several users could use the same visualization. And you can have the object server locally (pointed by GG). First step is to Identify already available tools. Define proper data models, servers with already calculated visualization.

WP430. Data mining (led by CSIC+UB).ESAC is already planning to buy a cluster that will be devoted to data mining, with hadoop or grid computing. Data mining into the archive is an option. Some associated centres that can host a copy of the catalogue could have also the data mining tools. GG says they already have a cloud system with the software. Requirements and architecture coordination. Seems hadoop covers most cases but some data mining could require differently. Need to identify user cases and come back with recommendations. Define how the data should be archived and some study on existing algorithms.

WP440 VO tools and services (CSIC+UBR). Adapt existing tools to the archive. MT following the first release, the budget from GENIUS will be finished but will have other money from CU9 funding. XL. We can participate on the testing before the archive is published, but will be from a pool of users more than developers.

### **Frédéric Arenou. WP500 Tools for data validation and analysis**

Work by CNRS mainly with collaborations from CSIC, KU, FFCUL, ULB, UG. GENIUS will do the implementation of the tests designed by the CU9 equivalent WPs. See presentation: Left side GENIUS, right side CU9. Detailed funding of each unit inside CNRS.

### **Jordi Torra. WP600 Support activities**

UB+CESCA+UCAM. See presentations. Introduction of the three main subpackages. FJ talks about the simulated data: GUMS10 and soon to come GUMS&GOG 11 with updated UM data. GOG including a Graphical User Interface that will be released shortly. XL: we should first define a data model. Tomorrow will have the discussion. First we need to move from the MDB the data we need to our local servers. And we need to work on the data model. Final data model should be frozen before the end of 2014. First thing in January we need to have a meeting and organize a data model (two working models in any moment, one to use, one to test), that is a technical issue. FA said we have to have one. XL has to be incremental but matched to the releases in ESA. FA asks how we can access the GUI. FJ you can download a jar file and use it. XL pointing out that the main simulation will run in Marenostrom. NH questions how to deal with large simulations problem: one possibility is to limit the scope of the attributes. FJ answers. Contributions of XL, AB. Matrix sizes problem (CU5 issue with photometric parameters).

EM (UB) gives the talk on behalf of Teresa Via (from CESCA) about the Webportal. They will send a questionnaire in the next weeks for opinions on the model and content of the webportal. It will be multilingual in accordance with the community. First webportal will be ready in January 2014.

GG asks who will have to answer the questionnaire. EM answered that is a closed survey with specific questions. XL it is your portal, so please say your wishes. LB points the need to comply with the commitments with the officer and EU funded projects on webpages.

### **Second day 05-12-2013**

Participants: Nigel Hambly (NH), Gonzalo Gracia (GG), Anthony Brown (AB), Mark Taylor (MT), Daniel Hestroffer (DH), Frédéric Arenou (FA), Teresa Via (TV), Frances Julbe (FJ), Jordi Torra (JT), Eduard Masana (EM), Xavier Luri (XL), Enrique Solano (ES), Lola Balaguer (LB), Dani Molina (DM).

Webex Connection: Andre Moitinho (AM), Dimitri Pourbaix (DP), Laurent Eyer (LE) and Nami Mowlavi (NM), Yoshiyuki Yamada (YY)

### **Dani Molina. Introduction to our Twiki Genius**

DM introduces Twiki words: consisting of two words with capital letter all together. Now everybody has a user for the GENIUS Twiki. On the top corner you have the Edit and Attach options, bottom left you have More actions. The editor is similar to other editors. And there is a Raw Editor that is the way Twiki stores the topic. Showing us how different options are done: Titles, subtitles, bold....and how to attach files, create a new topic and so on. We will have a Webex licence and create a related Google Calendar for the use of the Webex. Webex has also applications for Android, iPad, iPhone, and even you can ask them to phone you. Moreover we can record the meetings if we want. XL maybe we can do that for minutes. NH asks again about microphones, and Dani will check possibilities. Maybe there will be better options with personal microphones. DP says it is better to have individual microphones, so people get used to talk to the microphones. DH also agrees on the problems but maybe a table phone is better than individual microphones. NH points we should use the DPAC wiki instead to the GENIUS Twiki, at least for all scientific issues, only leave it for administrative and political issues. XL agrees with that, but we should have the specific information about documents for the EC reports and so on, documents intended for people specifically dedicated to GENIUS. For reporting will be much easier for us to have the info centralized here and will also be useful for confident issues. There is a section for each WP, we will have all the info for reporting in the part of the Management and we can put all the links to DPAC contents as we are all also DPAC:

XL suggests that the GENIUS meetings are recorded by default for minutes easiness and to be able to remember what was accorded. AB proposes that we can share this Webex facility with CU9 members but giving priority for GENIUS members. The maximum of people in a licence is 25 persons, but seems a reasonable number. The licence is personal and nominal and will go on the name of DM. **Action** to DM to get a Webex license and to look for microphone options.

If you have not received your Twiki user or you have any problem, please contact [dmolina@am.ub.es](mailto:dmolina@am.ub.es)

### **Joint discussion of GENIUS set up and detailed planning for the first year**

XL First a quick review of CU9 SDP schedule to see how it follows the GENIUS schedule.

Then WP200 will form a test group of users to check the archive with some simulations data (in coordination with ESAC SAT)..

CU9 plenary meeting proposed to take place in July (not before as the commissioning of the mission will be a very busy period). It would be a good time for Japanese colleagues to come to Europe, for updating us and increase the interaction with the Jasmine missions. The dates can be adapted to the Japanese colleagues' availability. **Action** for YY: before the New Year, tell us convenient dates for that meeting so the Japanese colleagues can attend. Maybe GENIUS can use this CU9 meeting, as a progress meeting too, separate part of the meeting for GENIUS and a part for Jasmine colleagues.

XL: There is a visualization meeting from the 9<sup>th</sup> to the 11<sup>th</sup> July in Vienna, maybe it would be possible to organize our meeting after or before that one to minimise travelling. FA asks if they could have a room for a meeting in validation. **Action** for AM: to speak with Joao Alves as organizer of the Vienna meeting. AM thinks it would be a good thing for their University too. XL wants to play around the dates and AM asks about how many days. AB suggested

two days for plenary and a third day for Validation and maybe splinters. AM: maybe better not to do parallel splinters as many people would like to go to more than one. XL. Lets see what the best dates for Japanese colleagues are and how can we arrange everything for July, in Vienna. MT asks if he is requested to be there as he has not received any notice. AM says sorry about that and you should absolutely come. MT will try to manage to be there.

Next one is a discussion about data model.

Then about Tools VO and CESCA testbed

Validation: massive statistics

Revisit the portal and concepts and simulated data and how to access this data.

First we go to the review of the SDP: CU9 Development Plan in DPAC, so everybody can get it through DPAC. There is a section on it about the detailed plan for cycles A, B and C. Planning until the first release has three cycles: A ending March 31, cycle B first public release of the archive by the end of September 2014, cycle C ending in January 2015 where we revalidate and test the archive for 8 months until the first catalogue release L+22 (October 2015). GG asks if 8 months is not too long for testing. XL says we should focus from August to October to validate the data before the release of the catalogue. So the system should be finished by August.

Then we go to the details of what each package has in the WP: first are issues related to management. Then we pass to WP920, why you have in Cycle A the CU9 provided tools should be documented. AB says that can be taken off. NH says that some draft documentation should be on place. AB says that is something they should do, not CU9. Data releases review: all documentation should be ready two months before the deadline.

WP930. Software Requirement Specification should be ready in Cycle A. One general SRS or one for each topic.

NH: In Cycle A the key thing is the requirements specification. After Java workshop last week, wants to ask several people to see if they will need some special specific hardware and software (like Visualization), to clear any misunderstanding what it would be required to implement this with the current resources.

XL asks if we need to write a separate document for scientific validation. GG thinks that should be separated. XL: it does not need to be decided now but it should be discussed. Either in the planning of the SDP should be cleared.

Cycle 3 we should have a working version of the archive.

WP940. XL asks if FA wants to comment on that. Cycle A contribution in SDP is clear. Clarify with NH if you want a separated SRS or included in the general.

Validation with GOG, we will need to fix with FJ the need for the non standard simulations (biased GOG to check validation tests) in cycle A. The model need to be defined, implemented and run so if it is need it for cycle A there is not so much time. That can count as Genius task on top of CU9 task.

In cycle C we will need two versions of the data model: one working version and one frozen. That has impact on the data model definition. At some point we need to freeze the data model, have a first model by early 2014. And then keep it frozen for the first release that should be end next year. The ideal is then to fix the data model as soon as possible, but in practice we can change it in case of emergency, we can have meetings for accepting or rejecting changes as we do. AB there is many things to take into account. XL wants to have a meeting to fix the model as soon as possible. NH asks how much of this work is GENIUS as it is not mention anywhere in the proposal. This is a CU9 thing. XL says it is important only to have into mind the two deadlines for GENIUS: fist data model early next year, and froze one end next year.

WP950 maybe GG want to comment. In first cycle only defining of the interfaces with all the others DPCs and who is in charge of accepting the software pieces. First couple of cycles we will need to decide this, as it is not very well defined for CU9.

WP960 Outreach. It would be important to check there are no inconsistencies GENIUS with CU9.

WP970 we will start to figure out how the data mining can be done, cycle B start testing the data mining with simulated data. We should clarify the spectra issue with GOG, we can for the time being do it just leaving empty the matrix. We can lower the level of oversampling. Please put that in the agenda to be discussed in DPACE. In cycle C a basic set of data mining should be ready.

WP980 Any special milestone in AM. AM: Do you have any question? We have these requirements that we can coordinate with the ones in GENIUS. We have some issue with the Gaia loads. The ITD need to be discussed, the exchange of ideas last week was very fruitful but I have not yet talked with the people who went (Albert Krone) to the meeting. NH that sounds good we can be in touch to clear. AM they were very ambitious but we need to clear what it is possible to do. We have been very general and we will have some more specifications when the SDP would be finished. GG they talk to have a meeting for Visualization telecom for Feb 2014, as before we need to have a meeting with Validation. XL: please check the compatibility of GENIUS with SDP schedule, that there is no contradiction between them

XL: Before coffee we will talk with AB as he needs to leave soon. GG can you share the Dictionary Tool? This was presented last week and it is used in DPAC for the Main Data Base but also for other data process. Each CU can go there and define the content of their section of the database. Several issues related to this: CU9 should find its data model using this tool. We need to create sections in CU9 folder describing what we have in data. There will be a group dedicated to this with a representative of each package. The process will be the same as with DPAC, mantis.... if you need something in the database from each WP, you need to inform this group. So please keep it in mind. Last week I raised some topics to discussion. Now we have some things that need to be removed from this section and go to the DPCs section. We are not creating product, we are receiving the data and doing things with that. Our model for the database will be a model for the DPC. FA asks if we can have flags. XL yes we can have flags, there is a difference if the flag was created during the processing of the data we will be able to read it and we can extract it and export them to our archive, but in the process of validation we can generate some flags too. We are writing in

the archive not in the MDB, we read MDB but we do our own database. NH what happen if we find a problem? XL we will never ever modify anything in the MDB, we report the faults and the DPAC will do the process, we only raise flags and problems. We can choose what we publish, we can decide if we do not publish data with problems or if we publish with the documented warning (AB said).XL: In the SDP this is explained. NH: yes that is fine, that is what I said, to move it to the DPC section. I noticed that there is a DPCI section that can potentially be used for some issues from the archive perspective. XL: if we are in the MDB section we are tight to their releases, but if we are in the DPC we can do our own releases and updated independently of the processing cycles, that is why we choose to go to the DPC section.

More points raised during the meeting in ESAC; Data model points: Data model should include specific sections for the external catalogues included in the archive. From the people at Italy we will have the part of catalogue data that not necessary will be published. Data model should be frozen well in advance of the actual release.

Need to synchronize. And be careful with stability! We should not depend in objects created in other CUs, as if they change their objects that will imply changes in us. Do it as stand alone as possible. NH this is understandable but sad because there are some fundamental reasons to use the dictionary tools. But it is a pure relational data model and not an object related data model. These are things to discuss on the engineering group. XL: should be a technical decision. The dictionary tool has several sections. I think we should make definitions on CU9 that do not depend on other CUs. NH thinks that is debatable. XL says that it makes releases more complicated. NH: that is life. There are dependencies as a fact of life. XL: we should have a meeting and my only awareness is stability and avoiding inconsistencies.

More points: is the data model for pre release the same as the released for users?? Do we need two data models, one for validation and then we extract what we need for the final. This is not relevant for the archive point of view as there will be everything but people does not need to see everything.

For the first release the data model will be very simple, but at the same time people will be working for a more complete release. Do we need a separate partition to be able to already work for the full database. GG: that means you are going to have frozen for the second version, frozen for the third version...

**Coffee break** AB and DP leaving.

WP400 VO Tools. We will have an open TAP interface for people to use the test archive. So you can already tell people that you can already connect to that through TopCat or SPEC.... with the simulation already there. The TAP interface should be ready early next year, and as soon as that happen to have tools to check. MT says that would be feasible but let me know to check everything works beforehand. NH asks about the current simulation + TAP already available. XL: that version is not open to the outside word. So they need to do it separately to be able to make it public. In fact something was done there locally. NH: just worried of a time schedule they need to get that open as soon as possible only for them to test and the developers (MT, ES and others) be able to access it. XL: there is an issue with making it public. MT asks which model the TAP is built. NH thinks is the CDS, MT that is quite a

sensible one. **Action** for MT and ES to know which TAP is implemented by ESA, ES works closer to Jesus Salgado so maybe he can ask directly about that. GG: they will be willing to have a small community to access it but not a big group. MT will write an email to Jesus (cc Enrique) with the open issues he believes there can be. Action for early next year.

Data Mining. Setting up a testbed in CESCO early next year. We need to have a telecom with Luis Sarro and pass CESCO a list of requirements of what we want. FJ: I would like to have it this Dec, I was talking with AG yesterday to have a first version this Dec. Check how to do this data so hadoop can read it. Meeting open to anyone on the GENIUS context.

Validation. FJ presented GOG and GAT (Gaia Analysis Tool) and that can be useful maybe for validation. That can help on the documentation need it for the release. And probably the checking of ranges and things like that are automatically done with GAT. Maybe it can be used as a data mining tool. So we agree the following: when we decide a version is good enough for release, then we run the extractor (Jesus done it) and ingest that in the database, and at that moment would be good to run this statistics tool GAT, so we can prepare an implementation suited to that used. We already used it with the IGSL catalogue and we ask validation for a list of things that you would like included. Please GG remind me to include it in SDP. FA asks where that will be installed. XL you can convert it to a doc or it can be access with the GAT tool. GG what we discuss it is to make part of it internal and part maybe public for the final release documentation. XL you get a set local sky pieces and a general of the whole sky. All the statistics depend on what is in the MDB (if there is Solar System data, we can put statistics related and so on). It can be used also for a filtering of the list of the flagged one as not usable, and check they have been removed and so on. The only problem is the list is very long and maybe too tedious to check.

WP600 XL introduced Teresa Via that will be the manager of the webportal, she will be the contact point for anything of the portal. The portal is GENIUS so we can do what we want with it! TV: please do send to me all your comments and suggestions. It will be useful to know if there is any local Gaia portal so I can see what you are doing. XL I add your email [tvia@cesca.cat](mailto:tvia@cesca.cat). TV we are working on the definition of sections and we will send you a small questionnaire on the contents of the webportal.

XL one point it will be nice to have all the info possible of Jasmine, webpages and so on.

Simulated data: XL would like to make it available as soon as possible. Soon we will submit a paper with the statistics of this simulated catalogue for A&A. Early next year we will have a paper ready to send you about the content and expected errors of the MDB. And maybe we can publish an internal one with systematic problems, to compare with the “non biased” one, to check if we can see the global effects on the data, and to design new checks on difference between the biased and no biased one.

AOB

NH in respect with financial matters: Do we have to do anything to begin with the grant? LB: just expend it. XL you can even spend more than the money received. LB: remember to spend it in the way you agree to do it on the proposal. Any change would need permission. XL: there is no problem to change but need to ask for permission. MT What if we do not spend? LB: you give it back. NH or give it to another partner, always asking for permission first.

GG: some of the topics are CU9 ones, very few are only for GENIUS: We should be very careful on sharing the info in CU9. XL this meeting needed to be only GENIUS KOM, but next meetings can be coordinated and shared with CU9. FA says there is some Milestones planned a year ago that need to be reviewed. NH can we do that without the permission of the EC. XL Review of Milestones (page 25 of GENIUS proposal): some will need to be delayed. We need to choose the external board but AB has left so we will do it by email. No major changes detected. External advisors could be: François Mignard, Lennart Lindegren, Gerry Gilmore. Please propose names, YY maybe you can give some names? DH how many advisors are expected? XL we do not have a limit, but we have a limit of budget so maybe 3 or 4.

Thanks for attendance. Wish you a successful next 42 months!

### **List of Actions**

- **Action** to DM: to get the Webex license and to look for microphone options.
- **Action** for YY: before the New Year, tell us convenient dates for the CU9 GENIUS Plenary Meeting so the Japanese colleagues can attend.
- **Action** for AM: to speak with Joao Alves as organizer of the Vienna Visualization Meeting July, 9 to 11, to check for a possible synergies to celebrate the CU9 – GENIUS plenary meeting.
- **Action** for MT and ES to know which TAP is implemented by ESA, ES works closer to Jesus Salgado so maybe he can ask directly about that. MT will write an email to Jesus (cc ES) with the open issues he believes can be. Action for early next year.
- **Action** to all: propose names for the GENIUS external advisors. Especially to YY to propose some names from Jasmine-related experts.