Deliverable D1.2.4

Project Title:	Developing an efficient e-infrastructure, standards and data-flow for metabolomics and its interface to biomedical and life science e-infrastructures in Europe and world-wide	
Project Acronym:	COSMOS	
Grant agreement no.:	312941	
Work Package 7 description	FP7-INFRASTRUCTURES-2012-1	
Deliverable title:	COSMOS Project Report	
WP No.	1	
Lead Beneficiary:	1:EMBL-EBI	
WP Title	Management	
Contractual delivery date:	01 10 2014	
Actual delivery date:	01 10 2014	
WP leader:	Christoph Steinbeck	EMBL-EBI
Contributing partner(s):	*Reza Salek and Christoph Steinbeck*	

Authors: Reza Salek, Christoph Steinbeck

Contents

1 Executive summary

The aim of this deliverable is to summarise up to 24 months activity of WP1 on management and coordination with COSMOS partners:

1. Report on management activity
2. Report on coordination activity
3. Report outreach and dissemination activity mainly on workshops, external meetings and conferences related to COSMOS
4. COSMOS partner meetings and workshops

2 Project objectives

With this deliverable, the project has reached, or the deliverable has contributed to the following objectives:

No.	Objective	Yes	No
1	Report on COSMOS management activities	X	
2	Report on COSMOS coordination activities	X	
3	Report on COSMOS workshops and conferences	X	
4	Report on COSMOS outreach and Dissemination	X	

3 Detailed report on the deliverable

3.1 Background

This work package will provide the management infrastructure for the proposed work and coordination activities. It will make use of the existing electronic communication platforms of the Metabolomics Standards Initiative and the Metabolomics Society, and further develop them, in order to be used by the COSMOS consortium. We will also organize the annual COSMOS consortium and stakeholder meetings, as well as regular workshops and staff exchanges between the COSMOS partners. We will systematically document the decision-making process and decisions made in teleconferences, meetings and by email exchange. This will be compiled regularly into COSMOS consortium documentation.

3.2 Description of Work

3.2.1 Report on COSMOS management activities

We have carried out regular biweekly or monthly teleconference meetings using Google hangout with the COSMOS WP leaders and COSMOS partners. An agenda was provided at the beginning of each meeting with the discussions and decisions minuted using a Google Document. All the documents were shared within the COSMOS participants and a final copy stored on the COSMOS website Internal pages (COSMOS-FP7.EU).

The second COSMOS partner annual meeting was held on 25-26 September 2014 in Leucorea Wittenberg, Germany, hosted by Drs Steffen Neuman (IPB-Halle) and Dirk Walter (Glom, MPG). The final Schedule was as follow:

Wednesday, 24.9.2014

Afternoon/Evening: Participant arrival
20:00 Self-paid dinner in town
21:30 Informal networking and spontaneous subgroups

Thursday, 25.9.2014

8:45 Welcome
9:00 Morning session I: COSMOS - focus on the past months -Including discussion time of 5-10 min where appropriate

WP 1 9:00 Chris – overview introduction
9:15 Reza – WP1 detailed introduction
WP 2 9:30 Steffen – WP2
10:00 Daniel – WP2 (nmrML)
WP 3 10:20 Dirk – WP3

10:45 (Coffee)
11:10 Morning session II: focus on the past months -Including discussion time of 5-10 min where appropriate

WP 4 11:10 Elon (presented by Reza) –WP4
WP 5 11:30 Thomas –WP5
WP 6 11:50 "not present"
WP 7 12:10 Ulrich –WP7

12:30 (Lunch + coffee + hallway-session)
14:30 Afternoon session I: Plans and discussions from the WP leads
Upcoming Deliverables
16:20 (Coffee)
17:00 Afternoon session II: Plans and discussions from the WP leads
Parallel sessions: 1) Chris et al write manuscript 2) Project report "Story line" 3) MX publication in Google doc 4) mzTab introduction and discussion
Upcoming Deliverables II
18:50 End of official
19:00 Workshop dinner

Friday, 26.9.2014

9:00 Morning session I:
Discussion on metabolomeXchange.
metabolomeXchange Requirement

Discussion on mzTab requirement

10:30 (Coffee)
11:00 Morning session II:
Periodic reporting, Person-Months for current and remaining Deliverables (all WP leads)
12:00 (Lunch 11 people)
13:00 End of official workshop

List of participants:

Reza	Salek	EMBL-EBI
Christoph	Steinbeck	EMBL-EBI
Michael	van Vliet	LACDR/Leiden University
Kenneth	Haug	EMBL-EBI
Steffen	Neumann	IPB Halle
Daniel	Schober	IPB Halle
Dirk	Walther	MPG
Tim	Ebbels	ICL
Silvia	Marin	UB
Pedro	de Atauri	UB
Thomas	Hankemeier	Leiden University
Albert	Koulman	MRC
James	Smith	MRC
Ulrich	Guenther	UoB
Tobias	Schulze	UFZ

Meeting Aims: Within the COSMOS consortium we discussed current works and planed the final 12 months of work activity within different WP and different subgroups. We also discussed and planned future workshops and staff exchange between partners. We discussed the last and future upcoming stakeholder meetings and started strategic planning for the invitation process, it has been agreed tentatively to aim the final stakeholders meetings to coincide with the annual Metabolomics meeting in San Francisco in 2015 (http://metabolomics2015.org/). Other discussions would be agreement on time and location of the annual meeting in 2015. Summary of other agreements and discussions are as follow:

- 3 workshops were planned for next year
    - Workshop on flux analysis standardization and reporting requirements – Barcelona host Marta Cascanate
    - Workshop RDF/metadata and compression formats for XML data formats –EMBL-EBI by hosted Reza Salek and Chris Steinbeck
    - Joint PSI/COSMOS meeting in San Francisco
- Change the name MetaboStore to MetabolomeXchange for all reports, websites and communications
- Get in touch with the Chinese, Korean, Japanese and Indian (Asian) Metabolomics community to link their national (or large local) databases with MetabolomeXchange.

- Commence article on google docs with an opinions piece on MetabolomeXchange, targeted to NATURE Biotechnology
- Discussion and to collect tools for integrated Metabolomics data analysis from partners

3.2.2 Report on Coordination effort with other WP

WP2: WP2 with feedback from partners and stakeholder involved in developing nmrML format continue to progress. Reaching it main goal and milestone. This work package had two additional deliverables. In D2.5 we have coordinated efforts from multiple international groups who are developing tools and parsers for the nmrML format. WP2 delivered automatic converters that read in proprietary vendor raw data files (Bruker and Varian/Agilent) and generate schema compliant nmrML XML files either manually or in a high-throughput batch mode. These parsers and converters are available for multiple programming languages (JAVA and Python) and can be deployed as web applications, as part of existing software pipelines or as standalone command line tools. WP2 also delivered parser extensions for different established software frameworks such as R and Matlab based packages (e.g. Batman and rNMR), which allow for reading in nmrML files and make their content amenable to statistical analysis. We also had interest from the proprietary Chenomx NMR suite developers to support the format at a later stage. WP1 has been involved in coordinating various efforts and in invitation of partners out side consortium to join our work. Example NMR data file in nmrML format are continuously added to the website http://nmrml.org/examples/.

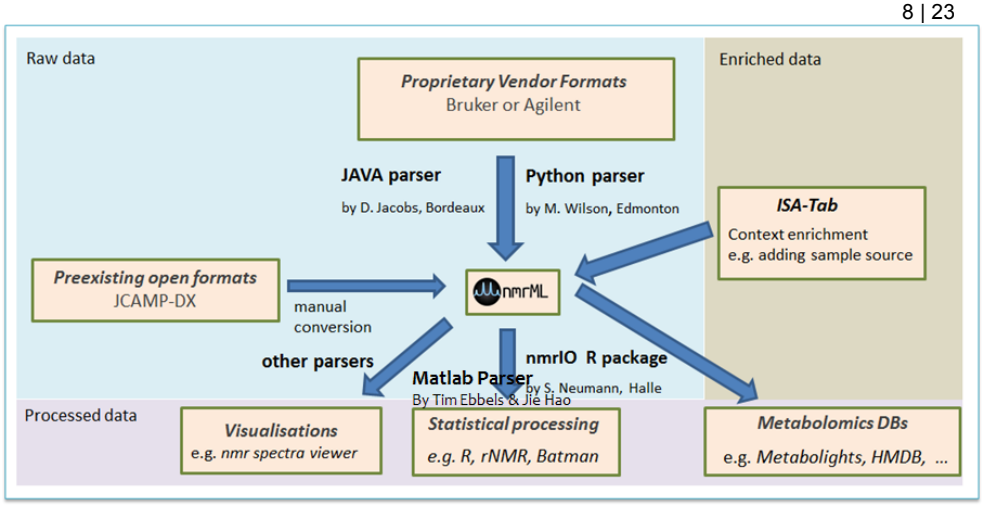


Figure 1: Illustration of NMR data management facilitation by means of the common nmrML standard

D2.8 examined semantic web standards that facilitate linked open data (LOD) and demonstrated their use for metabolomics data. While the technical standards (e.g. RDF and virtuoso server) already exist, WP2 has developed the "inventory" of terms and concepts required to express facts about metabolomics. WP2 has provided agreed-upon terminological descriptors, e.g. to characterize studies and digital objects in metabolomics. WP2 performed a survey of relevant data resources and existing LOD approaches to create, store and query semantic web data services for metabolomics. In addition to building RDF schemata to describe the LOD data content of established Metabolomics data providers, WP2 implemented several prototype resources, so called SPARQL endpoints, to test the RDF models, data conversions and querying. This culminated into a guideline document describing the current state, some best practices and future requirements for data service providers in metabolomics summarized in D2.8. WP1 has been involved thought out the discussion and developments with WP2.

WP3: This WP has collected 20 different computational tools and services that are relevant to data format standards. This included conversion tools, tools for spectra interpretation, tools for data processing, etc. All examples

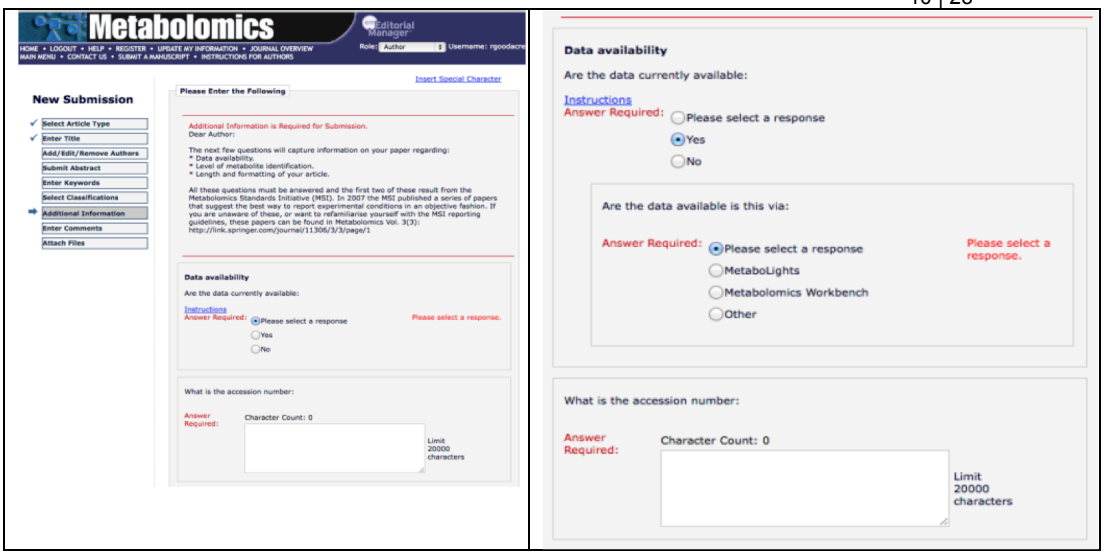
tools have adopted data standards or are partner in its developments (D3.2). All the packages are developed by the COSMOS consortium partners and made available to the metabolomics community (link available via COSMOS website - http://cosmos-fp7.eu/tools). WP3 together with **University of Bordeaux** continues to improve and establish software infrastructure that captures standardized experimental metadata. This software infrastructure was built based on the existing XEML-lab software suite that has graphics-supported descriptions of experimental designs along with standardized descriptions of experiment conditions via ontologies.

WP4: Had no deliverables for the period, however we have currently established recommendations of MetaboLights, MSI and links to COSMOS via:

- NATURE Scientific Data (http://www.nature.com/sdata/data-policies/repositories#omics)
- MDPI Metabolites, (see end of the document). (http://www.mdpi.com/journal/metabolites/instructions,
- The Metabolomics Journal (http://link.springer.com/journal/11306)
- EMBO http://emboj.embopress.org/authorguide#a3.4

We have also recommended that the "Data should be submitted in accordance with MSI (Metabolomics Standards Initiative) guidelines (http://metabolomicssociety.org/index.php/resources/metabolomics-standards)"

New proposed requirement by metabolomics journal (under review):



WP5: had two deliverables, In D5.1 WP5 and WP1 coordinated the efforts from multiple international metabolomics data providers (meeting setup by WP1) to make metabolomics data sets over their international data repositories searchable. WP5 designed and implemented a central online register called MetabolomeXchange to store meta-data of publicly available metabolomics data sets (Website registered and maintained by WP1). With this central register we provide a search interface for finding data sets of interest that are available in the different data repositories. Data sets can be added or updated by the individual providers by updating their local data feed. The provider feed is then read by the MetabolomeXchange update mechanism and processed accordingly. WP5 was able to connect all providers so far based on existing data feeds (XML/JSON) keeping technical and procedural changes to a minimum for the providers. To align the provider feeds WP5 wrote feed converters to adapt the original feeds to MetabolomeXchange compatible feeds. In addition to the basic search WP5 developed a ‘popular searches’ and ‘recent searches’ feature to improve the search experience.

For the deliverable D5.2, WP5 implemented a broadcast mechanism for MetabolomeXchange to inform the metabolomics community about new or

updated data sets. The broadcast mechanism is based on the RSS 2.0 specifications (http://www.rssboard.org/rss-specification).

WP6: This WP optimizes communications with BioMedBridges and its partner BMS infrastructures that have an interest in metabolomics (WP1 has been involved in BioMedBridges discussion). The deliverable 6.5 describes and summarizes a number of different activities that contribute to defining and supporting the role of metabolomics as a tool to validate pre-analytical procedures for sample collection and storage in bio-banks as well as to quantitatively assess the stability of stored samples over the years. Guidelines based on this information were described in 6.5 and have been formulated in a document currently under evaluation by CEN for approval as a European standard. This initiative started within the European project SPIDIA. To build on the above opportunities based on the development of metabolomics as a crucial technique, in addition to development of standard operating procedures and quality control in biobanks, a Expert Center for Metabolomics (EXCEMET) was formally established (http://www.excemet.org/). EXCEMET would present itself as a reference infrastructure for biobanks and has been described in a recent concept paper by the now established BBMRI-ERIC as a model of possible BBMRI-ERIC Expert Center.

3.3. Report COSMOS out reach and dissemination activity

3.3.1 Report on COSMOS Workshops and Partner Exchange

The Conferences attended or organised by EMBL-EBI

- **EMBL-EBI Agricultural-Omics course 17-21 February**: This course was setup to provide an overview of data resources, tools and analysis pipelines for 'omics datasets within the agricultural sciences. It was delivered using a mixture of lectures, practical computer sessions and discussions based on agricultural case studies and public domain data. Additionally there was opportunity to review the key opportunities and challenges within this field. On day five we provided and introduced hands on metabolomics databases,

resources and standards. The course was located at EMBL-EBI, Cambridge UK. http://www.ebi.ac.uk/training/course/agricultural-omics.

- **RIKEN-MPI Joint Workshop On Forefront of Metabolomics and Future Prospect RIKEN Japan 26 Feb 1 March.** Presented MetaboLights: Metabolomics data repository and the role of COSMOS" - COordination Of Standards In MetabOlomicS.
- **A workshop and hands on course: MetaboLights and COSMOS at Tokyo University 1 March -** presentation to participants and discussions on data standards.
- **EMBO Practical Course on Metabolomics Bioinformatics for Life Scientists** 17- 21 of March , 2014 http://www.ebi.ac.uk/training/course/embo-practical-course-metabolomics-bioinformatics-life-scientists
  - **Metabolomics International Data Exchange meeting**. COSMOS & MetaboLights. EMBL-EBI, UK, 2nd-3rd April 2014
  - **Workshop: Hands on data submission to MetaboLights using ISATools. And Maximizing '-Omics' interoperability by promoting metabolomics open standards data exchange formats: The biologist view**. Wednesday 21st May 2 hours during the 8th Meeting of the French Metabolomics and Fluxomics Network.
  - **EAWAG data management workshop, COSMOS & MetaboLights**. Switzerland, 25th June 2014
  - TMIC - Metabolomics Data Standards and Data Processing Workshop, COSMOS & MetaboLights. University of Edmonton, Canada, 15th-18th July 2014

Key areas covered during this meeting and workshop:

— Overview of the current metabolomics data/metadata infrastructure.

— Discussion of the key metadata items required for translation of experimental studies to online formats, which may be searched and analyzed

— Discussion of conversion of different data types to formats suitable for transfer to the DRCC

  - Raw data (NMR,MS, etc)
  - Analytical metadata (including protocols),
  - Associated biological and clinical data

— Final result datasets of quantitative/semi-quantitative metabolite values and appropriate substance identifiers.

— Discussion on data upload issues

— Discussion on online data presentation, search tools and statistical methods. (http://www.metabolomicsworkbench.org)

During this meeting we were able to have one-on-one discussions regarding COSMOS formats and data sharing as well as understanding the bioinformatics infrastructure of both project. We also agreed on next steps of collaboration in future.

- **Joint workshop with Biomed Central Gigascience Team at BGI Hong Kong, and the ISA Team at the University of Oxford e-Research Centre who organized a hackathon in Hong Kong,** June 2014

  In June 2014, also supported by a BBSRC UK-China partnering award (BB/J020265/1), the Biomed Central Gigascience Team at BGI Hong Kong, and the ISA Team at the University of Oxford e-Research Centre organized a hackathon in Hong Kong, with participation of young scientists and '-omics' data producers from local university and the UK metabolomics standards community (EMBL-EBI, COSMOS and Birmingham Metabolomics Centre). The goal was to establish common standards and curation practices for 'omics' data as well implement new ISA software (http://isatab.sourceforge.net), functionalities to facilitate deposition to MetaboLights repository (http://www.ebi.ac.uk/metabolights/), and support feature requests from journals using ISA formats, such as GigaScience (http://www.gigasciencejournal.com) and Scientific Data (http://www.nature.com/sdata/). The fruitful interaction between the participants over the duration of the meeting resulted in the delivery of an ISA-Tab viewing component for web browsers (http://dx.doi.org/10.5281/zenodo.11084 for the code and http://isatab.sourceforge.net/examples.html), and the conversion of MetaboLights ISA-Tab content to RDF. Finally, efforts to deliver an API supporting the programmatic creation of an ISA-Tab document are well under way. Owing to the success on this first ISA-thon event, the same teams will be organizing a follow-up meeting.

  Workshop was organized by COSMOS and the Data Standards Task Group of the Metabolomics Society during the annual Metabolomics Society conference in Tsuruoka June 2014

  Reza Salek (EMBL-EBI, UK) presented the current status and on-going developments worldwide on data standards, particularly the role of COSMOS (http://www.cosmos-fp7.eu). Steffen Neumann (Leibniz IPB, Halle Germany) gave an overview of data exchange formats in MS and how to get the users and instrument vendors involved. Philippe Rocca-Serra (Oxford e-Research Centre, UK) presented ISA-Tools for capturing metabolomics metadata, semantics and usage of RDF in metabolomics. In the second session, Scott Edmunds and Rob Davidson (GigaScience, Hong Kong) presented the Gigajournal point of view on publishing metabolomics data and workflows with the *GigaScience,* GigaDB and GigaGalaxy platforms. Susanna-Assunta Sansone (Nature Scientific Data) introduced NScD's role in placing importance on data and the role that the NScD is playing on putting value in your datasets. The last two presentations were by Reza Salek and Padma Maruvada (NIH/NIDDK, USA) giving an overview of metabolomics repositories, MetaboLights and NIH Metabolomics

WorkBench. We had plenty of interested discussion points on usage and promotion of standard, metabolomics data deposition and dissemination.

- **Joint COSMOS and HUPO PSI meeting April 2014**

  In April 13-16 2014 the COSMOS (COordination Of Standards In MetabOlomicS) is planning to participate and to have a joint meeting with the proteomics HUPO Proteomics Standards Initiative community. This meeting will take place in Schloss Reinhartshausen Kempinski, Nr Frankfurt, Germany. HUPO-PSI has defined community standards for data representation in proteomics and has facilitated data comparison, exchange and verification within the proteomics community. Many open source MS formats including: mzML, mzTab, mzIdentML and mzQunatML as well guidelines for minimum information reporting requirement for proteomic and peptide identification have been developed within this initiative. Working closely with the HUPO-PSI community should benefit the metabolomics community, particularly the COSMOS effort in development of open MS exchange formats for metabolomics. We also hope to contribute to the development of the MS based controlled vocabulary by PSI-MS by including the metabolomics community ontology requirement and terminology.

  Summer Schools: Data Acquisition and Analysis in Metabolomics (Pula, Cagliari), Italy, 15-20 September 2014, Topic covered Data standardisation in metabolomics and the role of COSMOS initiative, database and resources from EMBL-EBI for metabolomics data dissemination including data repository (MetaboLights) small molecular and metabolite (ChEBI) and pathways (Reactome). Other databases and open resource out there for metabolomics data handling, capturing results and data analysis were discussed. Also talks about a LIMS systems from collaboration work with Australia metabolomics for capturing metabolomics data and metadata that links to MetaboLights submission via ISA-tab format.

  MET-GR III workshop: Metabolic and Protein Network Analysis in Systems Biology, ***Conference and Cultural Center of the University of Patras, September 18-20, 2014.*** Metabolomics data repositories MetaboLights (http://www.ebi.ac.uk/metabolights) at the European Bioinformatics Institute and the Metabolomics Work bench by the NIH (http://www.metabolomicsworkbench.org/) in US and many more specialized repositories. Open exchange formats for data to make it as widely accessible as possible, vendor independent, enriched with terminological artifacts to ease exchange and query metabolomics experiments were discussed. Overview of COSMOS and its aims to develop and maintain exchange formats for raw data and processed information (identification, quantification), building on experience from standards developed within the Metabolomics Standards Initiative (MSI) and Proteomics Standards Initiative, (PSI) was presented. Also, the role that metabolomics repositories can play in adoption of such standards and making metabolomics results more extensively accessible.

COSMOS Workshop and partner exchange

- **First mzTab for metabolomics workshop, March 6-7 2014 -** COSMOS members together with the PSI community held the first "mzTab for metabolomics" workshop in Tübingen, Germany. Our aim was to drive reporting of metabolomic results further using a standardized, open, easy accessible and human readable tabular format. MzTab (version 1.0) already provides basic support for reporting small molecules that we plan to extend and harmonize with the more advanced reporting scheme for proteins and peptides available in MzTab.

3.3.2 Report on COSMOS in news and media

Information on COSMOS that has appeared in news articles:

MetaboNEWS

This newsletter is published in partnership between The Metabolomics Innovation Centre (TMIC, http://www.metabolomicscentre.ca) and the international Metabolomics Society (http://www.metabolomicssociety.org), and is intended to keep metabolomics researchers and other professionals informed about new technologies, software, databases, events, job postings, conferences, training opportunities, interviews, publications, awards, and other newsworthy items concerning metabolomics.

We now have a dedicated section in MetaboNews for monthly updates on Status of Data Standards: This new section within the Metabolomics Society News will be contributed regularly by Christoph Steinbeck (Chair of the Society's Data Standards Task Group) and Reza Salek from the EMBL-EBI, Cambridge UK.
MetaboNew Status of Data Standards –

- **MetaboNews July 2014:** Metabolomics Data Interoperability Interest Group (https://www.rd-alliance.org/group/metabolomics.html) has been formed as part of the Research Data Alliance (RDA, https://www.rd-alliance.org/). The RDA is dedicated to enabling open data sharing in science and is supported by, for example, the National Science Foundation (NSF) in the USA and the European Commission.

- The Metabolomics Data Interoperability Interest Group (IG) in the RDA aims to provide a forum to discuss all aspects of metabolomics data management and harmonisation. The initial membership of the Metabolomics Data Interoperability IG is drawn from the leaders of several internationally recognised research programmes that are developing metabolomics databases and from the international Metabolomics Society's Board of Directors. It encourages a broad group of researchers, technologists, and curators to join as members to address issues of data interoperability and sharing in metabolomics. Currently, the interest group is chaired by Christoph Steinbeck, whose team develops the MetaboLights database at EMBL-EBI in Hinxton, as well as Shankar Subramaniam, leader of the Metabolomics workbench project in the US.

- In particular, this IG will address interoperability of databases and the coordination of standards. In addition, there are obvious synergies with other existing technical and domain specific Interest Groups and Working Groups in the Research Data Alliance, including (but not limited to) the Big Data Analytics IG, Data Foundation and Terminology Working Group, or the Toxicogenomics IG.

- **MetaboNews August 2014:** A workshop was organized by COSMOS and the Data Standards Task Group of the Metabolomics Society during the annual Metabolomics Society conference in June held at Tsuruoka, Japan. Reza Salek (EMBL-EBI, UK) presented the current status and on-going developments worldwide on data standards, particularly the role of COSMOS (http://www.cosmos-fp7.eu). Steffen Neumann (Leibniz IPB, Halle Germany) gave an overview of data exchange formats in MS and how to get the users and instrument vendors involved. Philippe Rocca-Serra (Oxford e-Research Centre, UK) presented ISA-Tools for capturing metabolomics metadata, semantics and usage of RDF in metabolomics. In the second session, Scott Edmunds and Rob Davidson (GigaScience, Hong Kong) presented the Gigajournal point of view on publishing metabolomics data and workflows with the GigaScience, GigaDB and GigaGalaxy platforms. Susanna-Assunta Sansone (Nature Scientific Data) introduced NScD's role in placing importance on data and the role that the NScD is playing on putting value in your datasets. The last two presentations were by Reza Salek and Padma Maruvada (NIH/NIDDK, USA) giving an overview of metabolomics repositories, MetaboLights and NIH Metabolomics WorkBench. We had plenty of interested discussion points on usage and promotion of standard, metabolomics data deposition and dissemination.

- In June 2014, also supported by a BBSRC UK-China partnering award (BB/J020265/1), the Biomed Central Gigascience Team at BGI Hong Kong, and the ISA Team at the University of Oxford e-Research Centre organized a hackathon in Hong Kong, with participation of young scientists and '-omics' data producers from local university and the UK metabolomics standards community (EMBL-EBI, COSMOS and Birmingham Metabolomics Centre). The goal was to establish common

standards and curation practices for omics data as well implement new ISA software (http://isatab.sourceforge.net) functionalities to facilitate deposition to MetaboLights repository (http://www.ebi.ac.uk/metabolights/) and support feature requests from journal using ISA formats, such as GigaScience (http://www.gigasciencejournal.com) and Scientific Data (http://www.nature.com/sdata/). The fruitful interaction between the participants over the duration of the meeting resulted in the delivery of an ISA-Tab viewing component for web browsers (http://dx.doi.org/10.5281/zenodo.11084 for the code and http://isatab.sourceforge.net/examples.html), and the conversion of MetaboLights ISA-Tab content to RDF. Finally, efforts to deliver an API supporting the programmatic creation of an ISA-Tab document are well under way. Owing to the success on this first ISA-thon event, the same teams will be organizing a follow-up meeting.

We have regularly updated and promoted COSMOS via our several social media sites, including:

- Blogger - http://metabolights.blogspot.co.uk

- Twitter - #cosmosfp7

- Facebook - http://www.facebook.com/cosmosfp7

3.4 Next steps

1. Submission of technical manuscript on the new nmrML open standard file format.
2. Workshop plan for the EMBO practical course on Metabolomics 2015 in, EMBL-EBI
3. Schedule Stakeholders meeting at Metabolomics 2015
4. Organising the annual meeting of partner for 2015
5. Carrying out several COSMOS workshops nationally and internationally to promote standards in metabolomics
6. Coordination of development and dissemination of MS XML formats; mzML, mzIdentML, mzQuantML and mzTab
7. Coordination of development and dissemination of NMR XML formats; nmrML, nmrIdentML, nmrQuantML and nmrTab

8. Coordination on developments of tools, convertors and API for nmrML and mzML file formats

9. Coordination and dissemination of ontology and CV development

10. Interaction with vendors, software developers, Journals and databases to make COSMOS more inclusive and economically viable for them to participate in the development of the file formats, essential for the success of the initiative.

4 Publications

Orchard S, Albar JP, Binz PA, Kettner C, Jones AR, Salek RM, Vizcaino JA, Deutsch EW, Hermjakob H **Meeting New Challenges: The 2014 HUPO-PSI/COSMOS Workshop: 13-15 April 2014, Frankfurt, Germany**.Proteomics [2014]

Griss J, Jones AR, Sachsenberg T, Walzer M, Gatto L, Hartler J, Thallinger GG, Salek RM, Steinbeck C, Neuhauser N, Cox J, Neumann S, Fan J, Reisinger F, Xu QW, Del Toro N, Pérez-Riverol Y, Ghali F, Bandeira N, Xenarios I**The mzTab Data Exchange Format: Communicating Mass-spectrometry-based Proteomics and Metabolomics Experimental Results to a Wider Audience**. Mol Cell Proteomics [2014, 13(10):2765-2775]

Daniel Schober, Michael Wilson, Daniel Jacob, Annick Moing, Gerhard Mayer, Martin Eisenacher, Reza M Salek and Steffen Neumann: **Ontology Usage in Omics Standards Initiatives: Pros and Cons of Enriching XML Data Formats with Controlled Vocabulary Terms,** *ODLS Proceedings 2014*, Freiburg, www.onto-med.de/obml/ws2014/odls2014report_prefinal.pdf , page 36.

Hao, J., M. Liebeke, W. Astle, M. De Iorio, J.G. Bundy, and T.M.D. Ebbels, **Bayesian deconvolution and quantification of metabolites in complex 1D NMR spectra using BATMAN**. Nature Protocols, 2014. 9(6): p. 1416-1427.

Valcarcel, B., T.M. Ebbels, A.J. Kangas, P. Soininen, P. Elliot, M. Ala-Korpela, M.R. Jarvelin, and M. de Iorio, **Genome metabolome integrated network analysis to uncover connections between genetic variants and complex traits: an application to obesity.** J R Soc Interface, 2014. 11(94): p. 20130908.

Valcarcel Salamanca, B., T.M. Ebbels, and M.D. Iorio, **Variance and covariance heterogeneity analysis for detection of metabolites associated with cadmium exposure.** Stat Appl Genet Mol Biol, 2014. 13(2): p. 191-201.

Alcarraz-Vizán, G., Sánchez-Tena, S., Moyer, M.P. and Cascante, M. 2014. **Validation of NCM460 cell model as control in antitumor strategies targeting colon adenocarcinoma metabolic reprogramming: Trichostatin A as a case study. Biochimica et Biophysica Acta (BBA)** - General Subjects. 1840, 6 (Jun. 2014), p. 1634–1639.

Cacciatore, S., Luchinat, C. and Tenori, L. 2014. **Knowledge discovery by accuracy maximization.** Proc. Natl Acad Sci USA. 111, 14 (Apr. 2014), p. 5117–5122.

Cortés, R., Tarrado-Castellarnau, M., Talancón, D., López, C., Link, W., Ruiz, D., Centelles, J.J., Quirante, J. and Cascante, M. 2014. **A novel cyclometallated Pt(ii)–ferrocene complex induces nuclear FOXO3a localization and apoptosis and synergizes with cisplatin to inhibit lung cancer cell proliferation**. Metallomics. 6, 3 (2014), p. 622–633.

Gracie, K., Correa, E., Mabbott, S., Dougan, J.A., Graham, D., Goodacre, R. and Faulds, K. 2014. **Simultaneous detection and quantification of three bacterial meningitis pathogens by SERS**. Chemical Science. 5, 3 (2014), p. 1030–1040.

Martín-Bernabé, A., Cortés, R., Lehmann, S.G., Seve, M., Cascante, M. and Bourgoin-Voillard, S. 2014. **Quantitative Proteomic Approach to Understand Metabolic Adaptation in Non-Small Cell Lung Cancer**. Journal of Proteome Research. (Aug. 2014).

Salamanca, B.V., Ebbels, T.M.D. and Iorio, M.D. 2014. **Variance and covariance heterogeneity analysis for detection of metabolites associated with cadmium exposure.** Statistical Applications in Genetics and Molecular Biology. 13, 2 (Apr. 2014), p. 191–201.

Sánchez-Tena, S., Lizárraga, D., Miranda, A., Vinardell, M.P., García-García, F., Dopazo, J., Torres, J.L., Saura-Calixto, F., Capellà, G. and Cascante, M. 2013. **Grape antioxidant dietary fiber inhibits intestinal polyposis in ApcMin/+ mice: relation to cell cycle and immune response**. Carcinogenesis. 34, 8 (Aug. 2013), p. 1881–1888.

Salek RM, Steinbeck C, Viant MR, Goodacre R, Dunn WB. **The role of reporting standards for metabolite annotation and identification in metabolomic studies**. Gigascience. 2013 Oct 16;2(1):13. [Epub ahead of print] PubMed PMID: 24131531.

5 Delivery and schedule

The delivery is delayed: ☐Yes ☑No

6 Adjustments made

N/A

7 Efforts for this deliverable

Institute	Person-months (PM)		Period
	actual	estimated	
1:EMBL-EBI	3	2	24
2:LU/NMC	1		24
Total	4		

Appendices

1. N/A

Background information

This deliverable relates to WP1; background information on this WP as originally indicated in the description of work (DoW) is included below.

WP1 Title: Management
Lead: Christoph Steinbeck, EMBL-EBI
Participants: Christoph Steinbeck

This work package will provide the management infrastructure for the proposed work. It will make use of the existing electronic communication platforms of the Metabolomics Standards Initiative and the Metabolomics Society, and further develop them, in order to be used by the COSMOS consortium. We will also organize the annual COSMOS consortium and stakeholder meetings, as well as regular staff exchanges between the COSMOS partners.

We will systematically document the decision-making process and decisions made in teleconferences, meetings and by mail exchange. This will be compiled regularly into COSMOS consortium documentation.

Work package number	WP1	**Start date or starting event:**			Month 1			
Work package title	Management							
Activity Type	COORD							
Participant number	1: EMBL-EBI	2: LU/NC						
Person-months per participant	12	6						

Objectives
The consortium management activities will include

1. Coordination at consortium level of the 'technical' activities of the project.
2. The overall legal, contractual, ethical, financial and administrative management of the consortium.
3. Co-ordination of knowledge management, IPS and other innovation-related activities.
4. Preparing, updating and managing the consortium agreement between participants.
5. Maintaining communications with the Commission.
6. Overseeing the promotion of gender equality in the project.
7. Overseeing science and society issues related to the activities conducted within the project.

Description of work and role of participants

It is in the very nature of a coordination action to focus on communication between the participants for the sake of policy making, to document the outcome and spread the word to promote widespread community adoption.
We therefore wish to highlight the following:

Personal Communication
As part of this work package, we will organize monthly tele-meetings (Skype, phone, webex) of the COSMOS steering committee. Discussions and decisions will be minuted. We will invite international collaborating PI's to participate if needed.
Technical teleconferences of the work package participants will be held more frequently and likewise carefully documented.

Formal Communication
The policies, standards and workflows developed in this endeavour will be formally documented and published in the form of manuals, white papers and

recommendations. Any document created under this umbrella will be released under Creative Commons License to allow for barrier-free dissemination.
At the beginning of the project in month 2 we will deliver a project plan which will include a list of success indicators to monitor during the whole project, as well as the data we will gather that will help in assessing its impact. These indicators and metrics will be subject to change during the first review meeting and they will be reported at least in the annual reports.

Participants
The management work package will be coordinated by the EMBL-EBI, building on EMBL-EBI experience in the management of large consortia, for example in the BioSapiens, Embrace, and Felics projects. The Netherlands Metabolomics Center (LU/NMC) will be co-coordinator with their extensive experience in maintaining the largest national Metabolomics initiative in Europe and networking with an extensive set of international partners. In addition to EMBL-EBI and LU/NMC as the coordinators, all work package leaders are formal participants of this deliverable, due to the higher communication and reporting effort. EMBL-EBI have included the cost for an audit certificate under management subcontracting.

Deliverables

No.	Name	Due month
D 1.1	Project Plan	2
D1.2.1	COSMOS Project Report	6
D1.2.2	COSMOS Project Report	12
D1.2.3	COSMOS Project Report	18
D1.2.4	COSMOS Project Report	24
D1.2.5	COSMOS Project Report	30
D1.2.6	COSMOS Project Report	36