Deliverable D1.2.2

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 2013	
Actual delivery date:	01 10 2013	
WP leader:	Christoph Steinbeck	EMBL-EBI
Contributing partner(s):	*Reza Salek and Christoph Steinbeck*	

Authors: Reza Salek, Thomas Hankermeier and Christoph Steinbeck

Contents

1 Executive summary

The aim of this deliverable is to summarise up to 12 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 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

As previously was reported, the first COSMOS annual meeting was carried out at Leiden hosted by Co-Coordinator Prof Thomas Hankemeier and Dr Theo Reijmers from Leiden University and Netherlands Metabolomics Centre. A total of 20 persons attended the 11 & 12 of March 2013 annual COSMOS meeting in Leiden, which had the following agenda:

Monday 11th March (short term goals & deliverables):

10:00 Arrival & Coffee
10:30 Welcome & Introduction (**Thomas Hankemeier**, **Christoph Steinbeck**)
11:00 Update progress WP 2 (mzML (**Steffen Neumann**), nmrML (**Daniel Schober**))
12:30 Lunch
13:30 Update progress WP 3 (**Jan Hummel**), 4 (**Roy Goodacre**) & 5 (**Theo Reijmers**)
15:30 Coffee
16:00 Update progress WP 3 (**Jan Hummel**), 4 (**Roy Goodacre**) & 5 (**Theo Reijmers**)
17:30 Closure
19:30 Official Dinner

Tuesday 12th March (long term goals & deliverables):

10:00 Arrival & Coffee

10:30 Update progress WP 6 (**Leanoardo Tenori**) & 7 (**Ulrich Guenther**)

11:00 general discussion on

- Metadata subgroup discussion
- nmrML Workgroup discussion
- MS Workgroup discussion
- COSMOS & MSI role
- COSMOS & Met Soc 2013 Glasgow

12:30 Lunch

Continued discussions

16:00 Meeting Closed

Summary: During the meeting and by consultation with all COSMOS partners we planned the next 6-12 month of work activity within different subgroups, and discussed requirements towards the next deliverables and WPs overview. In addition, we created a tentative planed staff exchange between partners and detailed plans for the upcoming stakeholder meetings, finalizing the potential attendees and starting the invitation process, while creating a tentative agenda for the stakeholders meetings.

Other discussions included agreement on location of the next annual meeting (2014) to be held at Leucorea Wittenberg, Germany in late September.

A COSMOS stakeholder meeting was held on July 1st in Glasgow UK as planned, coinciding with the metabolomics 2013 meeting, held at the same location. By scheduling the stakeholder meeting during the Metabolomics Society meeting 2013 we have substantially saved on our costing, logistics and the time required to arrange the meeting. This was due to the fact that most of the stakeholders invited, were already present or planning to attend the society annual meeting. In total, we invited 52 (Including COSMOS) researches with diverse backgrounds and relevant expertise to metabolomics standards, metabolomics databases, including members of the metabolomics standard initiative, the proteomics standards initiatives, metabolomics journals, directories and members of the metabolomics society. We also had representation from NIH USA, MassBank database Japan, HMDB database Canada, GOLM Germany, PRIDE UK, NMC Netherlands and Metabolomics workbench USA that were present at the meeting.

The invited stakeholders were from diverse geographical locations all involved with standardization, databases and data repositories from across Europe, USA, Autralia and Asia (Japan and China). Details and the outcome of stakeholders meeting are reported in **D7.3.1.**

To date, as part of this deliverable, we have organized monthly teleconference meetings using Google hangout with the COSMOS WP leaders. Discussions and decisions were minuted using a Google Document that could easily be shared within the SC participants with a final and a second copy of outcomes stored on the COSMOS website Internal pages (COSMOS-FP7.EU), distributed among WP partners and announced via social media.

3.2.2 Report on Coordination effort with other WP

WP1 had been involved in a series on online discussions with the BioMedBridges partner as part of BioMedBridges Deliverable 5.1, building data bridges between biological and medical infrastructures in Europe. The aim of this discussions was to address the complex ethical, legal and regulatory issues resulting from data sharing or access to biomaterials. As a first step, a systematic in depth analysis of legal and ethical rules of sharing data and biological material between research infrastructures (RIs) on the European level was conducted. The objective was to provide a comprehensive analysis of the complete regulatory landscape for using data bridges, covering data protection according to Directive 95/46/EC, national data protection acts, GCP and GLP requirements, and considering implications by Directive 2001/20/EC, animal protection laws, security rules for bio-samples, laws concerning genetic data and stem cell research data, data access approvals by different forms of informed consent or by Hospital Boards and Ethics Committees. In addition, regulations for data connected to intellectual property and licence rights was discussed. The handling of ethics within our MetaboLights service was discussed. Metabolights is **an open source**, cross-species and cross-technique database for metabolomics experiments. Data is provided via user-submission **in ISA-Tab format (**http://isatab.sourceforge.net/**), linked to correct ontology terms**, which enables the data distribution and usage for dataset comparison. The metabolomics datasets of CERM in the WP6 use case will be managed within this

database. Additional sources of metabolomics data as part of WP6 are currently under investigation also as a part of a D4.3 deliverable from Jon Chamers and John Overington (EMBL-EBI).

WP2: WP2 with feedback from all partners worked on open source mass spectrometry formats. **D2.1** identified and addressed the limitations, which have so far slowed down the mzML adoption in metabolomics. For example, most GC-MS data is available either in non-open vendor formats or netCDF. COSMOS has now extended the existing exchange standards to capture metabolomics data in GC-MS experiments, detailed in **D2.1**. In addition we have evaluated data exchange standards developed by the Proteomics Standards Initiative (PSI) and added any required features to fully capture quantification data from Metabolomics experiments. WP2 has now developed and maintained exchange formats for raw data and processed information (identification, quantification), building on experience from standards development within the PSI initiative; more details are in **D2.3**. Finally we have coordinated efforts from multiple international groups who are working in NMR and metabolomics related software to design and establish the nmrML data XML based format, based on the experience with the PSI mzML format for mass spectrometry. As a result, the standards development work package WP2 has delivered the essential exchange standard for NMR based metabolomics raw data, more details in **D2.4**. We have organised two major workshops (at EMBL-EBI and second one at IPB Halle more details below) as well as regular biweekly online meetings (Google hangouts) to coordinate this effort with partners involved and nmrML format stakeholders

WP3: WP1 together with WP3 and COSMOS partners worked towards submission of standard metabolomics data sets to the MetaboLights repository (**D 3.3**). To date each partner has submitted or is working toward submission of the standard metabolomics data set to MetaboLights.

WP4 and WP5: With the consortium partners we have proposed a guideline for data deposition workflow between participating metabolomics databases and repositories, having in mind the potential collaboration with

the NIH work bench databases. This work was to ensure a coherent metabolomics workflow to run to its full potential, capturing agreed sets of metadata across different resources. The workflow definitions prioritised simplicity, usability, annotation quality and the plurality of metabolomics resources and databases to ensure a coherent connectivity between similar studies and to provide rapid matching results to end users; details of this are reported in **D4.1.** In addition we proposed a minimum set of agreeable COSMOS metadata format definition, formally agreed by the members of the COSMOS consortium to ensure correct and proper use, reporting and interpretation of the data by its owners and users. To achieve this common schema, a number of characteristics, or attributes, extracted from or associated with the original dataset were carefully designed, proposed and agreed by all COSMOS participants, stakeholders and publishers, as reported in **D4.2.**

WP6: Through working with WP6 we were able to optimize communications with BioMedBridges and with its partner biomedical sciences (BMS) research infrastructures with an interest in metabolomics. In addition, were able to maximizing the usefulness of the COSMOS activity for the current large-scale EU biomedical infrastructures and to optimize synergies between e-Infrastructures by providing input and receiving feedback from working groups addressing activities of common interest. Finally, were able to improve participation in the concerted activities and meetings related within the e-Infrastructures area. The details of this deliverable can be found in **D6.1.** In Addition as part of deliverable **D6.2** we established a NMR working group between the NMR metabolomics-working group of COSMOS and BioMedBridges. Furthermore, the requirements outlined previously within the deliverable **D6.1** of the current work package indicated that the action of the working

group should be oriented towards Biomedical Sciences in general and specifically in regards to biobanks.

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

- Metabolomics Society 2013 Conference SECC, Glasgow 30 Jun 4 July- http://www.metabolomics2013.org/page/workshops

10.00 - 12.00 Workshop 1B: Workshop on Metabolomics Data Dissemination, Standardization and Exchange (COSMOS/MetaboLights/NIH common funds Session) Background General Metabolomics Standards and Data Storage and Data Exchange in Different Standard Formats Reza Salek, EMBL-EBI, UK Hands on MetaboLights Tools and Data Submission using ISACreator Kenneth Haug, EMBL-EBI, UK, Philippe Rocca-Serra, EMBL-EBI, UK & Reza Salek, EMBL-EBI, UK Overview COSMOS (COordination Of Standards In MetabOlomicS) Initiative on Metabolomics Standards Christoph Steinbeck, EMBL-EBI, UK Introduction to the NIH Common Funds Metabolomics Consortium, Regional Comprehensive Metabolomics Research Cores" (RCMRC), Data Repository and Coordination Center (DRCC) Philip F. Smith, National Institute of Diabetes and Digestive and Kidney Diseases, USA Discussion with Users and Participants Moderator: Marta Cascante, Universitat de Barcelona, Spain & Christoph Steinbeck, EMBL-EBI, UK

- EMBL-EBI Metagenomics course 9-11 September: Managing, Analysing and Visualising Data Course, location Cambridge UK. http://www.ebi.ac.uk/training/course/metagenomics-course

- MetaboLights and the French Metabolomic and Fluxomic Network 10 - 12 june 2013 Amiens, MetaboLights and COSMOS talk and workshop - https://colloque.inra.fr/7_js_reseau_francais_metabolomique_fluxomique

- ISMAR 19 -24 May Brazil WW-NMR/COSMOS Workshop, Metabolomics Methods and New Standards, an introduction to WW-NMR and COSMOS

8:30 - 9:00 COSMOS, R Salek, C. Steinbeck, EBI - Progress in Metabolomics Standardization
9:00 - 9:25 Ulrich Guenther, Birmingham – Deciphering metabolic changes in cancer using metabolic fluxes
9:25 - 9:50 Claudio Luchinat, Florence. Intra- and inter-individual variability in metabolomic NMR profiles: challenges and opportunities

- The HUPO Proteomics Standards Initiative, Liverpool 15 -17 April 2013, COSMOS participated and presented http://www.psidev.info/presentations-hupo-psi-2013

- Biocuration 6th International Biocuration Conference Cambridge, UK, April 7-10, 2013 http://www.ebi.ac.uk/biocuration2013/

- 10th e-Infrastructure Concertation Meeting 6-7 March 2013 The Hotel, Brussels http://indico.egi.eu/indico/conferenceDisplay.py?confId=1217

- EMBO Practical Course on Metabolomics Bioinformatics for Life Scientists February 25- March 1, 2013 http://www.ebi.ac.uk/training/course/embo-practical-course-metabolomics-bioinformatics-life-scientists

- Cambridge University, Department of Genetics; Bioinformatics: Introduction to Metabolomics, metabolomics standard and databases.
  - Graduate School of Life Sciences, Fri 01 Feb 2013
  - Graduate School of Life Sciences, Mon 25 Mar 2013
  - Graduate School of Life Sciences, Wed 31 Jul 2013

COSMOS Workshop and partner exchange

Second EBML-EBI and IPB nmrML Hackthon, September 30 2013

Participants:

Reza Salek, Daniel Schober, Michael Wilson, Daniel Jacob, Philippe Rocca-Serra, Steffen Neumann, Jie Hao, Luis F. de Figueiredo, Ian Lewis, Kenneth Haug

Hackathon Focus:

1. Prepare for first release of nmrML and CV
2. Finalize use of cases/example implementations as required by D2.4 in Oct 2013
3. Start R package I/O readers for R based NMR software (batman and rNMR)
4. Initial technical specification of validator set-up, eventually with mock-up example validation (to be ready for Oct 2014)

Meeting venue:

Leibniz Institute of Plant Biochemistry
Dept. for Stress and Developmental Biology,AG Bioinformatics & Mass Spectrometry Weinberg 3, 06120 Halle, Germany

Presentations at IPB: [**Metabolomics Interest Group**] Talks and COSMOS Workshop on NMR data Standards at the IPB:

- Dr. Ian A. Lewis, Princeton University, Institute for Integrative Genomics: rNMR: open source software for NMR data analysis, 01.10.2013 @16:15 Kurt-Mothes-Saal, IPB

- Dr. Philippe Rocca-Serra, Technical Coordinator ISA-TOOLS.org, University of Oxford "Disseminating and Publishing Experimental Data while Complying with Community Requirements: The ISA experience", 02.10.2013 @10:00 Saal in Phytotechnikum

- Round table discussions with *local NMR groups (A Porzel Lab),* 01.10.2013 @16:15 Kurt-Mothes-Saal, IPB

Agenda:

Monday: (Informal, for people arriving early on Monday, arrival & early bird sessions)

- **15:45 welcome coffee & unofficial start of meeting**
  - Logistics: Internet connection, Update repositories, ...
  - Informal chat about objectives, goals and agenda
- 16:00 Pre-Opening R hacking session / R API design discussion
  - with Ian Lewis, Jie Hao and Steffen Neumann
- **17:30 Close**
- **19:00 Informal Group Dinner** at Peißnitzhaus or location near city center (to be anounced it raining)

Tuesday Oct 1st: (Intro and update on WP2 developments)

- **9:00 Official start at IPB Halle, please meet up at the IPB reception**
  15-20 min each participant, use template as provided for EBI hackathon
  Overviews and status-reports via short summarizing presentations of all participants
  - Michael Wilson: nmrML status, issues & coverage gaps
    - Enforce Kelvin for Temperatures
    - make sourceFile and pulseSequenceFileRef more consistent
    - flesh out the <binary> encoding
  - Daniel Schober: nmr CV status, issues & coverage gaps
  - Errors in the predecessor CVs
  - Reza & Ken: MetaboLights and nmrML/ nmrTab (mzTab) how they integrate
    Luis: nmr-fid-tool - library for NMR spectra processing
- **11:00-11.20 Coffee & Biobreak**
  - **11.20** Steffen Neumann: "The tools landscape"
    - Parser Status i.e. for tools like Batman and rNMR
    - R package, to be developed on github

    12.00 Validator discussions
    - Broad technical specification of validator software
      - Leveraging on PSI software not feasible as we use owl (OWL ontology import module and config needs reprogramming)
      - Using SWRL rules?
    - Define mock-up mapping rules for example validations
- **13.00-14.00 Lunch at local cantine**
- **14.00 Collective photo-shoot of COSMOS WP2 team**
- **14.30** Continue talks
  - Daniel Jacob: nmrML generator
- **15:00 Round table discussions with *local NMR groups***
  - *Presentation by Andrea Porzel providing an overview on IPB NMR data.]*
- **16:00** Presentations of **external speakers** (30-45 min per talk):
  - Ian Lewis: "rNMR: open source software for NMR data analysis", **Location: Kurt Mothes Saal**
  - Philippe Rocca Serra: presentation of choice
- **17:00** Jie Hao: Batman[Batman could via nmrML get access to Varian data]
- **18:00 Close**
- **19:00 Official welcome Dinner**

Wednesday: (nmrML XSD development issues)

- **9:00 Starting at venue place**
- **10:00 Talk** Philippe Rocca Serra: "Disseminating and Publishing Experimental Data while Complying with Community Requirements: The ISA experience", Location: Seminarraum in Phytotechnikum, IPB
- **11:00-11.20 Coffee & Biobreak**
- **11:20 nmrML XSD development issues**

- *DJ/PMB question (XSD Agenda part): is it plan to also integrate peaks list, as suggested in the UML use cases diagram for nmrML usage ? (i.e buckets for NMR)*
- *Peak lists*
- *Structure assignment*
- *Contact ontology*
- *Sample information*

- **13.00-14.00 Lunch at local cantine**
- **14:00 Visit to Porzel nmr Lab at IPB**
- **14:30 Release goals and policies**
- **15:00** Next Steps, get ready for upcoming WP2 Deliverables/28Nov EU meeting
- **16:00 Prepare memo for Bruker representatives and Varian and Jeol (at first release).**
- **17:30 Close**
- **19:00 Dinner** near market place (Cafe Schad http://www.zum-schad.de)
  - Alternatives: http://www.brauhaushalle.de/53429898d008ed51b/index.html
    - Diebels: http://dah.seg-unternehmensgruppe.de/anfahrt.html
    - Small city centre tour, or other socializing event

Thursday Oct 3rd: (deliverables and nmrML CV development issues)

Note: This is a Bank holiday in Germany. The IPB will be void at this day and the canteen is closed. We will rely on some local pizza service.

- **9:00 Starting at venue place**
  - Deliverable: Structure & implement upcoming WP 2 D.2.4 (nmrML + CV Prototype with examples)
  - General Documentation via wiki, websites, ...
    - Dissemination strategies, update NMR artefact and MI in Biosharing
- **11:00-11.20 Coffee & Biobreak**
  - Approve design principles for nmrCV
    - Top level ontology usage
      - BFO/ OBI/ BiotopLight ? ID policy
    - Discuss CV format (obo syntax vs. owl)
    - Discuss CV expressivity decisions (CV/Taxonomy vs. Description Logics semantics), Using Multiple asserted parenthood ? How are relations exploited ?
      - Have RDF, SPARQL and semWeb in mind (D2.8)
      - OWL to RDF ?
    - CV term deprecation policy
      - purgatory clean up
    - CV versioning and release policy
      - pre-release checks (script based?)
        - metadata completeness, e.g. definition ?
        - avoid redundancy ? NamingConventions-->OntoCheck ?
    - CV metadata policy
      - Review administrative and editorial Metadata
        e.g. like http://obi-ontology.org/page/OBI_Minimal_metadata, or using skos (DS, PRS)
- **13.00-14.00 Lunch via Pizza Express (Bank holiday today!)**
    - CV import and reference mechanisms
      - Interfacing with external ontologies/CVs (imports, mireoting)
        - float?, units, pato, ChEBI, Obi, IAO, Newt, …
      - How to deal with violating design principles i.e. DL class post-coordination vs. CV term pre-coordination ? (DS, PRS)
    - Agree on CV term-gathering policy
      - drawing new terms from external users' requests
      - drawing new terms from freet ext fields in xml ?
      - draw terms from ISA tools ?
    - Discuss CV publication & dissemination policy
      - Do we want to publish the CV in an openly accessible ontology library ? I.e. OBO Lib or is Bioportal or OLS enough ?
      - Biosharing publication

        - Will publish CV in OLS and bioportal as make Biosharing effort aware
    - CV evaluation strategy
        - Get Competency Questions for nmrCV & nmrML only
            - Get all spectra on 500Mz Bruker machine that do 1D H+ Spectra (on Human Urine samples for doping chemicals)
            - Get spectra generated via Bruker CryoProbe and D2O solvent
            - What samples with certain pH range
            - Get spectra according to decoupling method
            - Get spectra that were generated via open source nmr software
            - Get nmr Spectra that have been FT transformed and were smoothed via Gaussian smoothing
            - Show me reference spectra for xyz compound
            - What spectra used flow high resolution probe in the instrument?
        - Get CQ for nmrCV and IdentML CV
            - Get 1D spectra with doublets in ppm range x to y
            - Show me nmr spectra for changes in metabolites involved in Citrate Circle after Fat consumption in Human
            - How does the aromatic amino acid fraction differ in Hop plant variants ?
        - For future semantiWeb use case
            -
        - Evaluation tools? Onto-Check, OOPS?
- **15.00 CV content discussions**
    - Content gaps
    - Restructurings (avoid implicit redundancies)
- **16.00 Use case and example discussions**
        - are all completely covered by XSD and CV ?
    - Binding to MI standards (rules?)
        - On what layer is the MI enforced ? on XSD level? On CV/rule level? On both?
- **17:30 Close**

Friday: (use case discussion, reporting and wrap-up)

- **9:00 Starting at venue place**
- Examples and use case discussions (cont.)
- Add PMB bucketing use case: "special use cases, especially those where we use an novel binning approach, followed by a clustering step. (See <http://www.slideshare.net/danieljacob771282><http://services.cbib.u-bordeaux2.fr/SPECNMR/T06002>http://www.slideshare.net/danieljacob771282, *http://services.cbib.u-bordeaux2.fr/SPECNMR/T06002*)

- **11:00-11.20 Coffee & Biobreak**
- Crosstalk to other Cosmos WPs
- Planned and current papers
    - Reza on Roadmap paper
    - DanielS on WP2 standards paper
- **13.00-14.00 Lunch at local cantine**
- **14.00** Presentation of result
- Roundup discussions
- AOB & Action Items

15:00 Official Close

Staff Exchange Between University Bordeaux Segalen and Max Planck Institute Of Molecular Plant Physiology, April 22nd and May 3rd Benjamin Dartigues from Macha Nikolski's group

Birmingham - MetaboLights and COSMOS, May 1-2 Workshop

Agenda for EBI – UoB Visit and Discussions Thursday 2nd May 2013

EBI visitors: Chris Steinbeck, Reza Salek, Kenneth Haug
UoB organisers: Rick Dunn, Christian Ludwig, Ulrich Gunter, Mark Viant
Location: GEES 311 (Meet EBI visitors in Reception of School of Biosciences)

Potential UoB attendees for **morning training session (10am-12pm)**:
Rick, Mark, and their teams, Christian, Ulrich, Dan Tennant, Steve Young and their teams, include Kate Hollinshead, Andrew Filer (Sabina), Wiebke / Ange, steroidomics, John Easton, Shan He

10am – 12:00pm MetaboLights practical session (2 hours, all welcome, EBI to present)

- Introduction - past and present of metabolomics standards role of COSMOS
  - Minimum information reporting
  - Control vocabulary and ontology
- How do we capture metadata
  - Overview of ISATools
  - ISAconfigurator
  - Capturing metadata using ISAcreator
- MetaboLights Repository
  - Submission of data to MetaboLights
  - Reference layer in MetaboLights
  - Future plans – links to BML-NMR
- General Discussion – suggested topics
  - What "level" of metabolite ID's are reported (MSI, 4 levels)
  - We have focused on technology i.e. NMR and MS but what about specifics needs of both environmental and clinical data types?
  - Pipelining data between UoB resources and EBI e.g. BML-NMR

Break to get sandwiches/coffee (from Go Cafe) for working lunch

12:15pm Deep Metabolome Annotation (30 mins, limited attendees, UoB to lead)

- Introduce plans for extensive characterisation of metabolome, starting with Daphnia and moving to human samples (Mark and Martin to present), to include:
  - Why Daphnia (John Colbourne at UoB etc.)
  - Thermo TAP
  - Deep Annotation: Scale and form of the datasets we will generate, and timescales (assuming all goes to plan!)
  - Future: HTS with Daphnia and other biological systems
- Discussion
  - Can they be accommodated in MetaboLights? And/or in other databases (mzCloud) and how is data cross referenced? (note – some of this discussed in COSMOS below)
  - Submission pipeline of Daphnia and duration of knowledge layer
  - Integration with other 'omics'?

12:45 – 1:45pm COSMOS (60 mins, limited attendees, EBI to lead)

- Overview of COSMOS
  - Objectives and partners
  - COSMOS subgroups
- File formats to capture experimental metadata
  - mzML where are we
    - BIRMINGHAM questions/comments include:
    - Have standards been taken from proteomics w/o modification?
    - Handling different types of MS data?
    - mzCloud as (new) MS database, relevant to discussion
    - Can Birmingham help?
  - nmrML where are we

        - nmrML working groups
        - Standards of reporting NMR assignments
            - NMR binning
            - NMR concentration
        - Capturing the experimental metadata
        - Can Birmingham help?
    - Filling the gaps in CV and ontology missing for nmrML and mzML relevant to metabolomics

2 pm Natural Products (45 mins, limited attendees, both UoB & EBI presenting)

- Natural Products Chemistry research community (Greg and Mark to present)
    - their databases
    - their reporting standards: no formal ones for databases, but there are reporting standards for some journals
    - could we form a working group to define such standards? (Met Soc Task Group for metabolite ID – Rick / possibly the Met Soc Data Standards Task Group?) – High profile publication?
- Introduction to Chemical Entities of Biological Interest (ChEBI) **(EBI to present)**
- Discussion

Short break (others return)

3:00pm Metabolomics training & future collaborations (45 mins) (no slides)

- Review training currently offered by UoB (NBAF, academic courses)
    - Very brief overview of NBAF-Birmingham
    - ELIXIR UK Node – Metabolomics Sector Lead
- Review training currently offered by EBI
    - EMBO course
- Discuss potential for joint EBI / UoB Metabolomics Training Course
    - Course structure: division into (1) analytical/data generation/initial processing and (2) computational/informatics; including tools to assist with biological "interpretation" such as pathway enrichment analysis
    - Logistics: frequency, how to integrate the two parts
    - Can we translate aspects of this hands-on course to web-based course (are there MOOCs resources at EBI? UoB –
    - Funding: EMBO or EMBL? Metabolomics Society? Combination?
    - Sponsorship: e.g. Thermo (TAP)
    - Formal support for UoB PDRA liked to ELIXIR?
    - "Endorsement

3:45 pm Future collaborations and AOB

4:00 pm Departure

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.

- Metabo Interviews with Christoph Steinbeck
    - MetaboNews, December 2012: Metabointerview with Dr Christohp Steinbeck, highlights work of COSMOS, Metabolomics standards and databases such as MetaboLights
    - http://www.metabonews.ca/Dec2012/MetaboNews_Dec2012.htm#MetaboInterviews

- Metabo Interviews Tim Ebbels
    - MetaboNews, September 2013: Metabointerview with Dr Tim Ebbels, highlights work of COSMOS.
    - http://www.metabonews.ca/Sep2013/MetaboNews_Sep2013.htm#MetaboInterviews

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 Nov 2013
    - The Coordination of Standards in Metabolomics (COSMOS) consortium (http://www.cosmos-fp7.eu) over the last year has worked on extending the existing mzML and mzQuantML standards to encompass metabolomics experiments and data. Steffen Neumann of IPB-Halle leads these efforts that build upon developments by the Proteomics Standards Initiative (PSI). In addition, we have created examples for metabolomics data and augmented the controlled vocabulary required to capture the relevant terminology. Bioinformaticians writing metabolomics-related software can choose from different libraries to read/write these formats.

    - The popularity of the mzML inspired the development of a corresponding open file format for NMR, termed nmrML (check http://nmrml.org/ for development and documentations). We are eager to gather feedback on our first version of the standard from the NMR based metabolomics community. Currently NMR data analysis software such as batman (http://batman.r-forge.r-project.org/) and rNMR (http://rnmr.nmrfam.wisc.edu/) are involved to become early implementations of the format. We would like to invite other software developers and instrument vendors to join by contacting us at the nmrML mailing list: nmrml@googlegroups.com

- COSMOS and the MetaboLights database (http://www.ebi.ac.uk/metabolights/) together with the National Institutes of Health (NIH) Common Funds Metabolomics Initiative (http://commonfund.nih.gov/metabolomics/) and the Metabolomics Society are preparing to have a joint coordination meeting in early 2014. The tentative plan is to establish a network of partners to foster data flow within the emerging international network of metabolomics databases.

- Latest publication for metabolomics standard: The role of reporting standards for metabolite annotation and identification in metabolomic studies by Reza M Salek, Christoph Steinbeck, Mark R Viant, Royston Goodacre and Warwick B Dunn. *GigaScience* 2013, 2:13 doi: 10.1186/2047-217X-2-13 [PMID: 24131531]

News by University of Barcelona

- The UB: partner of the European project COSMOS to generate and promote new models to exchange metabolomics data
  News of Universitat de Barcelona, 04/2013
  (http://www.ub.edu/web/ub/en/menu_eines/noticies/2013/04/091.html)

- The UB: partner of the European project COSMOS to generate and promote new models to exchange metabolomics data
  News of Fundació Bosch I Gimpera, 04/2013
  (http://www.fbg.ub.edu/index.php/en/home-7/18-noticies/699-13-projecte-cosmos-eng)

- BioSilMeC jointly organizing the project COSMOS-FP7
  News of the XRQTC (Xarxa de Referència de Química Teòrica i Computacional), 02/2013
  (http://www.xrqtc.com/index.php/en/News/biosiimec-jointly-organizing-the-project-cosmos-fp7.html)

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 COSMOS imitative main publications, with all COSMOS partners included.

2. Submission of technical manuscript on the new nmrML open standard file format.

3. Workshop plan for the EMBO practical course on Metabolomics 2014 in, EMBL-EBI
4. Joint coordination meeting of COSMOS with the National Institutes of Health (NIH) Common Funds Metabolomics Initiative and the Metabolomics Society scheduled for early 2014. The tentative plan is to establish a network of partners to foster data flow within the emerging international network of metabolomics databases.

5. Schedule Stakeholders meeting at Metabolomics 2014

6. Organising the second annual meeting of partners for September 2014

7. Carrying out several COSMOS workshops nationally and internationally to promote standards in metabolomics

8. Towards joining resource meeting of PSI & MSI member scheduled next year on mzTab file format

9. Coordination of development and dissemination of MS XML formats; mzML, mzIdentML, mzQuantML and mzTab

10. Coordination of development and dissemination of NMR XML formats; nmrML, nmrIdentML, nmrQuantML and nmrTab

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

12. Coordination and dissemination of ontology and CV development

13. 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

1. 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.

2. Salek RM, Haug K, Steinbeck C. **Dissemination of metabolomics results: role of MetaboLights and COSMOS.** Gigascience. 2013 May 17;2(1):8. doi:10.1186/2047-217X-2-8. PubMed PMID: 23683662; PubMed Central PMCID: PMC3658998.

3. Salek RM, Haug K, Conesa P, Hastings J, Williams M, Mahendraker T, Maguire E, González Beltrán AN, Rocca Serra P, Sansone SA, Steinbeck C.**The MetaboLights repository: curation challenges in metabolomics.** Database (Oxford). 2013 Apr 29;2013:bat029. doi: 10.1093/database/bat029. Print 2013. PubMed PMID: 23630246; PubMed Central PMCID: PMC3638156.

4. Steinbeck C, Conesa P, Haug K, Mahendraker T, Williams M, Maguire E, Rocca-Serra P, Sansone SA, Salek RM, Griffin JL. **MetaboLights: towards a new COSMOS of metabolomics data management.** Metabolomics. 2012 Oct;8(5):757-760. Epub 2012 Sep 25. PubMed PMID: 23060735; PubMed Central PMCID: PMC3465651.

5. Jacob D, Deborde C, Moing A. **An efficient spectra processing method or metabolite identification from 1H-NMR metabolomics data.**Anal Bioanal Chem. 2013 Jun;405(15):5049-61. doi: 10.1007/s00216-013-6852-y. Epub 2013 Mar 23. PubMed PMID: 23525538.

6. Schober D., Mayer G., Moing A., Eisenacher M., Neumann S., **Ontological analysis of controlled vocabularies used in PSI/MSI supported XML standards**, Workshop: ODLS 2013, GI-Edition Lecture Notes in Informatics, *Proceedings of the Jahrestagung der Gesellschaft für Informatik 2013*, Matthias Horbach (Hrsg.),

Koblenz, Germany, 16.–20. September 2013, p. 1875-1888, https://wiki.imise.uni-leipzig.de/Gruppen/OBML/Workshops/2013-ODLS-en

7. Alejandra Gonzalez-Beltran, Steffen Neumann, Eamonn Maguire, Susanna-Assunta Sansone and Philippe Rocca-Serra. **The Risa R/Bioconductor package: integrative data analysis from experimental metadata and back again.** . *BMC bioinformatics, in press*

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	1		12
2:LU/NMC	1		12
Total	2	2	

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