[6:33:14 AM] \*\*\* Gerard Sylvester added Ahsan morshed, Gudrun Johannsen - FAO, Johannes Keizer, Prashanta Shrestha, Sachit Rajbhandari, Thomas Baker to this chat

\*\*\*

[6:33:33 AM] Gerard Sylvester says: Good morning! all in?

[6:33:58 AM] Armando Stellato says: y

[6:35:30 AM] Johannes Keizer says: hi

[6:36:25 AM] Gerard Sylvester says: Dickson - MIMOS STP Presentatino

[6:40:01 AM] \*\*\* Thomas Baker can't be added to this chat due to his/her privacy settings

\*\*\*

[6:46:45 AM] Gerard Sylvester says: Applications built on-top of the STP

[6:52:17 AM] \*\*\* Johannes Keizer added Thomas Baker to this chat

\*\*\*

[7:05:45 AM] Thomas Baker says: Dickson: Multimodel Semantic Browser

[7:06:29 AM] Gudrun Johannsen - FAO says: Multimodal semantic browser

[7:06:33 AM] Thomas Baker says: ...Visualize and navigate large knowledge bases

[7:07:07 AM] Thomas Baker says: Query built visually by dragging concepts and relaations

[7:07:18 AM] Thomas Baker says: Query Graph converted automaitcally into SPARQL

[7:07:50 AM] Thomas Baker says: ...Testbed is 1.3 billion triples

[7:09:15 AM] Thomas Baker says: ...Gather info from blogs, where are topics talked about?

[7:09:22 AM] Gerard Sylvester says: all queries are in parallel ; intuitive pre-fetch

[7:10:54 AM] Gerard Sylvester says: SOA - Architecture Presentation

[7:12:31 AM] Thomas Baker says: ...Allegro Knowledge Base server is indexed.

[7:12:48 AM] Gerard Sylvester says: Application <-> KBI <-> KBase Server (also creates index)

[7:15:06 AM] Gerard Sylvester says: Lookup service in SOA acts as UDDI for the services

[7:16:40 AM] Armando Stellato says: when out on the web, the delegator takes care of getting reg/req from distributed apps

[7:24:57 AM] Armando Stellato says: they wrote their own load balancing componennts, which measure directly the hardware load

[7:30:39 AM] Armando Stellato says: there's one Data Index (a file), then each KBS has its own local index with its indexed data. Each KBS can have different data, but their federated in someway

[7:32:56 AM] Thomas Baker says: Dickson: Allegrograph with free-text indexing the fastest

[7:35:00 AM] Thomas Baker says: ... NASA uses AllegroGraph

[7:35:07 AM] Prashanta Shrestha says: Allegro

[7:35:09 AM] Thomas Baker says: Up to 15 million triples, AllegroGraph is free.

[7:36:13 AM] Thomas Baker says: Dickson: should be two installations:

[7:36:23 AM] Armando Stellato says: free-text indexing is embedded inside allegrograph; they provide an extension for SPARQL which brings free textual search inside SPARQL as a dedicated operator (filter)

[7:36:34 AM] Thomas Baker says: ... one for public, one for knowledge managers - not in production mode

[7:36:52 AM] Gerard Sylvester says: migration policy to be put in place

[7:37:01 AM] Thomas Baker says: ...should be a process in place to migrate to production site

[7:39:33 AM] Gerard Sylvester says: Modelling - Knowledge Base - Architecture?

[7:40:09 AM] Gerard Sylvester says: [Keizer] : Modelling & KBase - Technology - User Interaction?

[7:40:35 AM] Thomas Baker says: ... Apps need to talk to KBI. Can be public domain.

[7:40:39 AM] Thomas Baker says: Johannes: Need three groups:

[7:41:27 AM] Gerard Sylvester says: [Keizer] (a) Modelling Group - AGROVOC Concept Scheme - What is the right approach to it?

[7:41:38 AM] Thomas Baker says: ...1) modeling group - need final agreement on future modeling of AGROVOC Concept Scheme - what is right approach? I think we are far on this and can agree. Feel more confident we can agree in this workshop.

[7:42:43 AM] Thomas Baker says: ...2) User interface - user requiremnets - work has gone into workbench - pretty good shape. This group: what are the things that have to be adjusted until we get working. What really has to be done. Another way to think user requirements - what would we like to do in a year - life-changing in handling this?

[7:42:44 AM] Gerard Sylvester says: [Keizer] (b) User Interface/Requirements Group - working in 2 direction - what are the things that need to be adjusted - what would we like to do in a year? Future?

[7:42:57 AM] Thomas Baker says: ... User Req that give completely new way of modeling?

[7:44:18 AM] Gerard Sylvester says: [Kezier] (c) Technology - what has to be done? transfer to something that works better (based on Armando presentation; AllegroGraph)

[7:46:36 AM] Thomas Baker says: ...3) Technology group: first, very crude discussion of technologyy as is - how to get it to work? Then: transfer to something that works better. Armando outlined yesterday. Then: integration into Allegrograph. Should we try to do now? Intrigued by benchmarking results! Users hate to wait. Then: how to manage mirror of development and production servers. Need to see open-source discussion - not ideologiical here - how to become sustainable? Sustainability maybe achievable in different ways besides open source. Fantasy: AGROVOC hosted at MIMOS, doing in collaboration with FAO for Agri community - stable solution - maybe decide not to embark on venture for open-source SKOS editor - getting open-source projects - most projects are dead - what is the balance?

[7:46:59 AM] Thomas Baker says: SKOS editor as open-source - useful - but not necessarily for us to do.

[7:47:45 AM] Thomas Baker says: AgriDrupal - we are collaborating - tens of thousands - Open Source community exists. Instead of embarking on open source - keep open mind for now.

[7:48:02 AM] Thomas Baker says: Where could this evolve if we are on MIMOS platform.

[7:49:24 AM] Gerard Sylvester says: [Tom] User Interface and Structural group into one group

[8:36:18 AM] Gerard Sylvester says: [Discussions from Group Modelling & User Interaction]

[8:41:42 AM] Armando Stellato says: find if there're really 43K synonyms rels in the original DB

[8:42:32 AM] Gudrun Johannsen - FAO says: acction: synonyms should be checked in AGROVOC database

[8:49:56 AM] Gerard Sylvester says: 8251 tuples hasSynonyms in the current AGROVOC MySql DB

[8:50:31 AM] Gudrun Johannsen - FAO says: is it only in English?

[8:50:54 AM] Ahsan morshed says: yes

[8:51:37 AM] Gudrun Johannsen - FAO says: ya it's possible

[8:52:45 AM] Armando Stellato says: [ACTION]: mege classes and related instancves

[8:53:56 AM] Gerard Sylvester says: hasMainLabel is it Pref label or Alt lable?

[8:53:57 AM] Armando Stellato says: [ACTION]: come up with an algorithm that uses the triple about isMainLabbe to determine whther it is a preferredLabel or altLbael

[8:55:11 AM] Gerard Sylvester says: then aso:isMainLable can be dropped after determining if isMainLabel determines if it is a prefLable or altLabel

[8:57:16 AM] Armando Stellato says: [ACTION]: have different classification scheme (for exammple, for AGRIS)

[8:59:22 AM] Armando Stellato says: [ACTION]: remove those c\_ and i\_ after removal of class/instance difference

[9:04:42 AM] Ahsan morshed says: [ACTION ]: to examine the datatype for example rdf:list . ..

[9:05:44 AM] Ahsan morshed says: [Action]: To check the has\_sub vocabulary

[9:05:45 AM] Armando Stellato says: [ACTION]: flag handling of termtypes in general as an issue, model has to be changed

[9:08:05 AM] Gerard Sylvester says: We have 15 scopeIDs in English

[9:08:30 AM] Armando Stellato says: [ACTION]: hasEditorialNote: is it significantly different from scopeNote?

[9:09:11 AM] Armando Stellato says: [ACTION]: we can use SKOS native properties for this

[9:10:25 AM] Ahsan morshed says: [Action]: FAOF status ?

[9:11:33 AM] Gerard Sylvester says: [ACTION] hasTermType needs to be reviewed - is it modelled as a data type?

[9:11:34 AM] Ahsan morshed says: [Action]: has\_termtype , need to be reviewed?

[9:12:21 AM] Gerard Sylvester says: SKOS termType can be related to hasTermType with a range and identified the URI instead of a string?

[9:13:51 AM] Armando Stellato says: [ACTION]: hasStatus: can bbe used any other standard property: FOAF uses some term\_status property which is prefixed as vs: though i did not find any vs prefix declaration

[9:20:27 AM] Armando Stellato says: [ACTION]: isPartOfSubVocabulary flagged: SKOS collection? scheme?

[9:24:18 AM] Gerard Sylvester says: 53,000 CONCEPTS and 500,000 TERMS (approx)

[9:27:42 AM] Armando Stellato says: [ACTION]: data is only 80%, but it just crashes at a certain time, so maybbe there are holes in it an data can seem even less. just we just have to generate a complete snapshot of itr

[9:32:09 AM] Armando Stellato says: [ACTION]: hasDefinition nature (first-lass citizen?) should be clarified

[9:34:43 AM] Gerard Sylvester says: Lexicalization (+ noun) = SKOS Label ?

[9:37:25 AM] Gerard Sylvester says: [Tom] no functional requirements, does not add any additional info, for noun...can it be dropped?

[9:40:17 AM] Ahsan morshed says: [Action]: to check it (noun)

[9:48:46 AM] Armando Stellato says: [ACTION]: replace lexicalization+noun --> SKOS-XL(abel))

[9:49:39 AM] Armando Stellato says: [ACTION]: discuss possible requirement of meta-vocabulary for linking external resources (to probvide homogeneous semantics for them)

[9:53:19 AM] Gerard Sylvester says: [JKeizer] : MeSH, WordNet, (+Euro WordNet) + other vocabularies

[10:03:34 AM] Armando Stellato says: [ACTION (already mentioned it before)]: different categorization scheme; intention was to integrate directly into the thesaurus, though maintaining like different FACETS on them (perspectives, havng separate classification schem)

[10:03:51 AM] Gerard Sylvester says: [JKeizer] AGRIS Caris classification scheme to be integrate into the Thesaurus and automatic linking to AGROVOC. AGROVOC CS facilitates extraction of any classification scheme as needed

[10:04:50 AM] Gerard Sylvester says: [JKeizer] flag AGROVOC concepts used in various classification schemes

[10:08:03 AM] Gerard Sylvester says: in SKOS a concept can belong to many concept scheme - same effect can be achieved by tagging concept as belonging to ex: <inScheme> AGRIS Caris classification or <inScheme> AGROVOC

[11:11:19 AM] Gerard Sylvester says: [Issue\_Flaged] : a number of properties are declared to have a range of xmls:string - is this the correct usage; in Dublic core RDFS:literal is used?

[11:19:19 AM] Armando Stellato says: other declared vocabularies should made available as separated linked data..

[11:26:59 AM] Armando Stellato says: [ACTION] some clustering properties: WBDatatypeProperty, conceptEditorialDatatypeProperty... better to not be put on linkedData (they are just used to drive the application through the model).

proposal: they could be used inside dedicated NamedGraphs (which would not be outputted to linkedData export)

[1:00:14 PM] Thomas Baker says: 15:30 - Johannes

[1:00:28 PM] Thomas Baker says: 1) What short-term changes needed to Workbench interface?

[1:00:37 PM] Thomas Baker says: 2) What kind of interface do we want for the future?

[1:00:58 PM] Thomas Baker says: ... for the breakout group on interface.

[1:01:47 PM] Thomas Baker says: 3) How could the transformation from Agrovoc from current model to SKOS take place? Are there content problems related to it?

[1:02:23 PM] Thomas Baker says: Technical breakout group...:

[1:03:00 PM] Thomas Baker says: -- what are the tech probs that must be resolved to fix for Open Archive by May - with working interface - taking into consideration that many triples may be eliminated

[1:03:47 PM] Thomas Baker says: ... If we cannot change structure by end of May we can do some interventions - renouncing some triples to improve performance

[1:03:57 PM] Thomas Baker says: ... Put everything into MIMOS system?

[1:07:45 PM] Armando Stellato says: johannes agrees on skipping aos from AGROVOC namespace

[1:08:25 PM] Armando Stellato says: agrovoc should \*reallY\* be in the URIO

[1:09:48 PM] Armando Stellato says: we have been heavily criticized by Neon evaluators for having non human-understandable IDs. In Jo's opinion, no language is globallly human understandable, and he prefers to keep original Agrovoc IDs

[1:29:08 PM] Armando Stellato says: [Tom] http://id.loc.gov/authorities offers example on a template for writing understandable URIs though maintaining codes: ...initialpartofNS/code#understandablename

[1:30:20 PM] Thomas Baker says: http://id.loc.gov/authorities/sh85097196#concept

[1:30:34 PM] Thomas Baker says: ... for the concept, resolves to the page: http://id.loc.gov/authorities/sh85097196

[1:31:29 PM] Gerard Sylvester says: Testing of the application by ICRISAT

[1:32:10 PM] Armando Stellato says: and, just for the hashVSslash issue (pro and contra), just some history: http://esw.w3.org/HashVsSlash

[1:44:19 PM] \*\*\* Gerard Sylvester added Michelle, Rowena to this chat

\*\*\*

[1:44:39 PM] Gerard Sylvester says: Michelle & Rowena you there?

[1:45:05 PM] Michelle says: yeap

[1:45:06 PM] Rowena says: yes

[1:46:30 PM] Gerard Sylvester says: [Gudrun] : Increasing the speed of the search functionality - prime importance

[1:48:40 PM] Gerard Sylvester says: Some features of Advanced Search like exact match, starting with etc to be incorporated into the Search

[1:53:23 PM] Thomas Baker says: Jo: Comparing search string with all labels...?

[1:54:37 PM] Rowena says: Clicked on the concept selected however it doesn't go the exact one.

[1:54:41 PM] Gerard Sylvester says: selecting concepts sometimes gives different results (wrong)

[1:54:45 PM] Thomas Baker says: Gudrun: Very often happens that you have to click three or four times - reload - takes a long time

[1:56:45 PM] Thomas Baker says: ... Or you get a list, often get "no results found".

[1:57:58 PM] Gerard Sylvester says: wrong results after a search query - consistently takes a lot of time to display search results

[1:57:59 PM] Thomas Baker says: ... Sometimes have to wait 3-5 minutes for result.

[1:58:16 PM] Gerard Sylvester says: [Gudrun] SEARCH highest priority!

[1:58:49 PM] Thomas Baker says: ... From my POV the highest priority. Takes ages until you can find the concept you want to relate to the concept.

[1:59:25 PM] Thomas Baker says: Gudrun: System seems not to remember the last action.

[2:00:06 PM] Thomas Baker says: ... System should know how to go back where you started.

[2:00:07 PM] Gerard Sylvester says: values not passed from one page to another through actions

[2:00:16 PM] Rowena says: Gudrun: The system doesn't remember what the search when after choose to click on "show URI" and also "non preferred term".

[2:00:58 PM] Thomas Baker says: Memory of where you came from.

[2:01:00 PM] Gerard Sylvester says: example : Search for Rice then select non prefered name -> search needs to be executed again

[2:01:17 PM] Thomas Baker says: Lavagna: Even while adding new relationships - finding concept to add.

[2:03:04 PM] Gerard Sylvester says: Instead of searching can we type directly the concept?

[2:04:36 PM] Thomas Baker says: Gudrun: As alternative to browse, should be possible to type "Rice" or "Forest" and if it shows me terms starting with "Rice" instead of waiting for everything to load.

[2:05:02 PM] Rowena says: Gudrun : suggest for a search with content to browse instead to wait for entire list to load.

[2:05:13 PM] Gerard Sylvester says: In creating a new relationship to another concept for Related Concept can we redefine the BROWSE to do a SEARCH with contains [text] and then select the relevant concept

[2:06:33 PM] Gerard Sylvester says: for move concept there is no validation

[2:06:41 PM] Thomas Baker says: Lavagna: Validation process - moving concept, need to do validation afterwards - but there is no validation process, and you don't see it in the list either - very dangerous - can damage system by moving to wrong place and not know.

[2:09:12 PM] Gerard Sylvester says: [Home] to be rechristen as [Recent Changes]

[2:15:25 PM] Gerard Sylvester says: If a terms exists - then a warning should be given while adding new terms : <XYZ> exists in this context do you want to create <XYZ> in a different context?

[2:19:36 PM] Gerard Sylvester says: Validator's role to make sure that there is no duplicate term; in any case there should be a warning otherwise too much of load on the validator

[2:19:41 PM] Rowena says: A validator to ensure terms are not duplicated.

[2:20:09 PM] Rowena says: Should there be a prompt out sign to highlight terms had been duplicated.

[2:21:47 PM] Thomas Baker says: Re: earlier discussion of URIs, see http://groups.google.com/group/pedantic-web/msg/fc698ddd51115d60, where Richard Cyganiak writes:

[2:21:52 PM] Thomas Baker says: In my experience,

getting content negotiation right is difficult, and it's a persistent

source of interoperability problems. At the same time, I perceive the

303 approach as more complex than the hash approach. With hindsight, I

think it's a bit unfortunate that “303 + conneg to HTML and RDF” has

become the “canonical” way of deploying RDF on the Web, because it's

complex and error-prone, and there are simpler methods. The focus on

“303 + conneg” has created a perception that deploying RDF is harder

than it actually is.

With RDFa, we're in the happy situation that there is a selection of

technologies that make it possible to do away with all the complex

server configuration around content negotiation and 303 redirects. So

RDFa plus hash URIs hits a sweet spot in many situations.

> I'm particularly interested in emerging

> best practice for publishing large, continually growing

> vocabularies.

I think that the question of “hash vs. slash” and “RDFa vs. HTML+RDF/

XML” is largely orthogonal to the question of the size or change rate

of the dataset. You could use <items/3234584235> which 303s to a

describing document specific to that resource; or you could use <items/

3234584235.html#it> without 303s and conneg. I would use the latter if

I wanted to keep my server setup simple and if I can live with the

slightly uglier URIs.

[2:35:02 PM] Gerard Sylvester says: Rights to be granted ONLY to people we know

[2:40:07 PM] Thomas Baker says: ... AGROVOC could be spammed.

[2:45:41 PM] Gerard Sylvester says: No term available in selected language - UI to be 'beautified' to hide this message?

[2:46:17 PM] Rowena says: Concept: No term available in selected language are shown at the top of the list whenever start up the system

[2:47:52 PM] Gerard Sylvester says: add if(isNull(termspell)) then don't display on the WB on SEARCH - based on language selected

[2:48:54 PM] Rowena says: Frustration occurred when a user derived such a long list of output but no able to view parent concept.

[2:49:11 PM] Gerard Sylvester says: Parent concept hovers above the child concept in the display??

[2:53:21 PM] Thomas Baker says: ACTION: In order to use terminology consistently, decide between:

[2:53:49 PM] Thomas Baker says: -- Label (pro: used in SKOS, con: some people expect "Term"

[2:53:54 PM] Gerard Sylvester says: Screen realestate to be looked into

[2:54:21 PM] Thomas Baker says: -- Lexicalization (pro: used in legacy AGROVOC model, con: AGROVOC-specific)

[2:55:16 PM] Thomas Baker says: -- Term (pro: some people expect this, con: different from SKOS terminology, which could lead to confusion in capacity building)

[2:57:46 PM] Gerard Sylvester says: Images links should be checked periodically

[2:58:28 PM] Rowena says: have to copy image for twice at least from internet explorer

[2:59:18 PM] Gerard Sylvester says: what happens if we find that a image has been moved/deleted?

[2:59:53 PM] Gerard Sylvester says: image template?

[3:00:35 PM] Gerard Sylvester says: option to add other resources also

[3:01:09 PM] Gerard Sylvester says: for the next version : Option for mapping

[3:02:57 PM] Gerard Sylvester says: Geographical location to be tagged with the concept? (like spatiallyIncluded in for countries)