MINUTES of NexusLinguarum + W3C BPMLOD teleconferences

Useful links/material

- W3C BPMLOD group website https://www.w3.org/community/bpmlod
- W3C BPMLOD group GitHub https://github.com/bpmlod
- W3C BPMLOD wiki (guidelines 2023)
  https://www.w3.org/community/bpmlod/wiki/Guidelines_and_best_practices_for_LLOD
- List of Meetings:
  https://www.w3.org/community/bpmlod/wiki/Meetings_of_the_community_group
- Expression of interest + github ids:
  https://docs.google.com/spreadsheets/d/1VZgOOV3bE_RLLRNu_A97ztpv0cuDuRp6I6Asy5t1fmA/edit#gid=1214801611
- Current Modus operandi

  https://docs.google.com/document/d/1EQPywP6_0qd9uiFJU_Wn4A1Yqx9z-9RbiTsgu2nKFtY/edit

Meeting 17/05/2024

Agenda and connection details

https://www.w3.org/community/bpmlod/wiki/Teleconference,_17/05/2024

Participants: (Please add your name)

Jorge Gracia (Univ. of Zaragoza)
Katerina Gkirtzou (Athena Research Center)
Victor Rodriguez (UPM)
Ranka Stanković, University of Belgrade
Thierry Declerck (DFKI)
Ineke Schuurman (KU Leuven (BE))
Ilan Kernerman (Lexicala)
Milan Dojchinovski (CTU Prague/INFAI)
Andon Tchechmedjiev (IMT Mines Alès)
Patricia Martín-Chozas (UPM)
Mike Rosner (UM)
Bev Corwin
Fahad Khan (CNR-ILC)
Christian Chiarcos (U Augsburg)
Gilles Sérasset (Université de Grenoble)

Penny Labropoulou (ILSP/ARC)
Giedre Valunaite Oleskeviciene (Mykolas Romeris University)
Patricia Martín-Chozas (UPM)
Thierry Declerck (DFKI)
Dagmar Gromann (University of Vienna)
Milan Dojchinovski (CTU Prague/INFAI)
Frank Abromeit

Regrets:
Mike Rosner
Penny Labropoulou (ILSP/ARC)
Andon Tchechmedjiev (IMT Mines Alès)

MINUTES - Check status on proposed guidelines and best practices:

Terminology generation (leader Patricia Martín-Chozas)

No representative

Bilingual dictionaries (leader Jorge Gracia)

Nearly finishes, some final touches are missing, after the end of the semester will be circulated again for final feedback, before summer break it would be publised (that's the plan)

Crosslingual linking (leader Mike Rosner)

- Regular telcos, last one one 26/04, next one scheduled for 28/05
- Progress on Guidelines drafting the first draft in proceeding well, soon we hope to have the first draft
- Paper is put on hold for the moment

Corpora annotation (NIF, Web annotation - leader Milan Dojchinovski)

No progress made due to low interest.
A resource with NIF for 10 languages is published from Ranka, Christian thinks it's too complicated for the guidelines.

Guidelines for Developing NIF-based NLP Services (leader Milan Dojchinovski)

No progress made due to low interest.

Wordnets (leader Fahad Khan) → Dropped

Decided not to proceed. Enough material from Global Wordnets Association.

Guidelines for LLOD aware services (leader Andon Tchechmedjiev)

Start of efforts awaiting progress on WP2 survey post Nexus Linguarum. Likely first drafting based on survey data extraction in next September or october.

LLOD for Under-resourced languages (Giedre Valunaite Oleskeviciene and Ranka Stanković)

The first draft for this topic: [BP LLOD for Under-resourced languages - Google документи](#)
Plan to proceed
- by circulation draft for feedback,
- publish a short version in a Journal and

Neuro-symbolic LLOD (leader Katerina Gkirtzou)

- Regular telcos, every month. Due to workload, May's telco has been cancelled, but June's meeting will be scheduled after LREC-COLING
- We have agreed on the definition of Neurosymbolic by reusing the definition from Hamilton's 2022 paper: integrating neural and symbolic approaches, but for our scope we will introduce also some relaxations to also include combined methods as well
- We decided on two different aspects: data aspect and approaches
- Data aspect: Find use cases and model them with FRaC and DataCube
- Approaches: Check papers about nesy and LLOD with the idea to fined how these individuated taxonomies are relevant to the LLOD community, which is the one targeted by the recommendation. For the moment on hold, waiting for proceeding of specific venues.

Licensing and metadata (leader Penny Labropoulou)

- No progress made due to other tasks; plans to restart in June.

Multimodal data (Ineke Schuurman) -> dropped for the moment

Sign languages (Ineke Schuurman and Andon Tchechmedjiev)

Regular calls, increasingly consolidated survey document, a subset of which could make a good base for a recommendation draft.

Lexicographic LLOD (Jorge Gracia + Francesco Mambrini)

- No progress since last meeting, plan to restart activities before summer break

Next Meeting:

In September. Doodle will be circulated at the beginning of September

Meeting 10/01/2024

Agenda and connection details

Participants: (Please add your name)

Jorge Gracia (Univ. of Zaragoza)
Katerina Gkirtzou (Athena Research Center)
Victor Rodriguez (UPM)
Ranka Stanković, University of Belgrade
Thierry Declerck (DFKI)
Ineke Schuurman (KU Leuven (BE))
Ilan Kernerman (Lexicala)
Milan Dojchinovski (CTU Prague/INFAI)
Andon Tchechmedjiev (IMT Mines Alès)
Patricia Martín-Chozas (UPM)
Mike Rosner (UM)
Bev Corwin
Penny Labropoulou (ILSP/ARC)
Fahad Khan (CNR-ILC)
Christian Chiarcos (U Augsburg)
Gilles Sérasset (Université de Grenoble)
Giedre Valunaite Oleskeviciene (Mykolas Romeris University)
Patricia Martín-Chozas (UPM)
Thierry Declerck (DFKI)
Dagmar Gromann (University of Vienna)
Milan Dojchinovski (CTU Prague/INFAI)
Frank Abromeit

MINUTES

Check status on proposed guidelines and best practices:

- Terminology generation (leader Patricia Martín-Chozas)

  In last month no specific effort itself. There are the meeting on the terminology module itself. Also there is an effort for survey on nlp and terminology task, with WG2 of nexus.

- Bilingual dictionaries (leader Jorge Gracia)

  Near finished guidelines, an advance draft was sent in the list for feedback. I got a single review from Ilan. Needs to finalize the metadata session. Once finished, will be published as a draft at first in bpmlod page, wait for a couple of weeks or a month before publishing officially.

- Crosslingual linking (leader Mike Rosner)
  - Progress on Guidelines drafting the first draft + Paper proceeding normally.
  - Last telco 18th December
  - Next telco scheduled for 15th Jan
  - The 1st draft will be soon out. Paper will take a little more time

- Corpora annotation (NIF, Web annotation - leader Milan Dojchinovski)

Had already one call to discuss the overall state.
Two unsuccessful tries to find a common slot for the next call (low number of participants). Will give it another try - doodle for a call until the end of Jan.

- Guidelines for Developing NIF-based NLP Services (leader Milan Dojchinovski)

Same as above.

- Wordnets (leader Fahad Khan)
  - No progress made (very busy period). But plan to work on this latter half of January & February

- Guidelines for LLOD aware services (leader Andon Tchechmedjiev)

Waiting on end of reading assignments for wg2 survey to start gathering the material that will go into the guidelines. Expecting draft by end of nexus

- LLOD for Under-resourced languages (Giedre Valunaite Oleskeviciene and Ranka Stanković)
  - Meeting started, Last telco 8th Jan
  - initial document prepared, bibliography collected
  - next meeting on 22nd January 10h CET

- Neuro-symbolic LLOD (leader Katerina Gkirtzou)

  - We meet once per month the 3rd Wednesday of the month at 13:00 CET. Next meeting is on the 17th of January.
  - Still in brain-storming mode: our discussions regards the focus of the topic and how to approach.
  - Start gathering material to study in github issues:
    - [Scope of the recommendation](#)
    - [Data models and metamodels for neurosymbolic learning on LLOD?](#)
    - [Potential applications and use-cases?](#)
    - [Examples of existing approaches](#)

- Licensing and metadata (leader Penny Labropoulou)
  - Inactive for a long time due to my other duties
  - Restarting work; doodle sent out for a new meeting (limited participation)

- Topic to discuss in the next meeting: Two ways to approach the topic: gather bibliography or more pragmatic by gathering resources
- Once decision is made: a first draft would be easy to compile

- Multimodal data (Ineke Schuurman) -> dropped for the moment

- Sign languages (Ineke Schuurman and Andon Tchechmedjiev)

There have been regular calls to identify existing efforts and material, rich document already in place. We won't have time to make new contributions by the end of nexus, so we will limit ourselves to surveying what exists and drawing perspectives for future directions. Both for the nexus final report, and for a first draft recommendation. Beyond nexus, we want to develop on some things like the sign language ontology and its integration with Ontolex, to make the recommendations actionable with actual RDF data.

- Lexicographic LLOD (Jorge Gracia + Francesco Mambrini)

Guidelines started, still early draft. Last telco skipped for health reasons

Other Issues

- Action: Jorge to check with the SO about possibility of merging the two Nexus Linguarum deliverables regarding Guidelines and Best Practices. → Approval by SO and MC

- JG: suggest to update https://www.w3.org/community/bpmlod/wiki/Guidelines_and_best_practices_for_LLOD -> **TODO:** KG

- Regarding the nexus deliverable on BPMLOD topic: a summary with a pointer to the best practices. We should try to have advance drafts in topic within the next few weeks

Next Meeting:

In February 2024

Meeting 31/10/2023

Agenda and connection details

Participants: (Please add your name)

Jorge Gracia (Univ. of Zaragoza)
Katerina Gkirtzou (Athena Research Center)
Victor Rodriguez (UPM)
Ranka Stanković, University of Belgrade
Thierry Declerck (DFKI)
Ineke Schuurman (KU Leuven (BE))
Ilan Kernerman (Lexicala)
Milan Dojchinovski (CTU Prague/INFAI)
Andon Tchechmedjiev (IMT Mines Alès)
Patricia Martín-Chozas (UPM)
Mike Rosner (UM)
Bev Corwin
Penny Labropoulou (ILSP/ARC)
Fahad Khan (CNR-ILC)
Christian Chiarcos (U Augsburg)
Gilles Sérasset (Université de Grenoble)
Giedre Valunaite Oleskeviciene (Mykolas Romeris University)
Patricia Martín-Chozas (UPM)
Thierry Declerck (DFKI)
Dagmar Gromann (University of Vienna)
Milan Dojchinovski (CTU Prague/INFAI)
Frank Abromeit

Regrets
Fahad Khan

MINUTES

Reminder: participation assignment update

1. Expression of interest + github ids:
   https://docs.google.com/spreadsheets/d/1VZgOOV3bE_RLLRNu_A97ztpv0cuDuRp6I6Asy5t1fmA/edit#gid=1214801611
   Please verify that the github id is correct, especially those marked in red background.

Check status on proposed guidelines and best practices:

1. Terminology generation (leader Patricia Martín-Chozas)

   Discussions with Rute for the Ontolex-Terminology module to start the first meeting the following next weeks. For the guidelines, the group will structure some guidelines with info from the survey paper of WG2 of nexus linguarum.

2. Bilingual dictionaries (leader Jorge Gracia)

   Update previous version.Update from old lemon to OntoLex vocabulary. Almost finished. Missing section regarding metadata. Jorge sent an email to the whole community for feedback. If needed, a dedicated telco will be set. No hard deadline for the feedback. Probably a couple of weeks.

   Github repo: https://github.com/bpmlod/Bilingual_Dictionaries_Report

   Draft available in github. Preview report at
   https://bpmlod.github.io/Bilingual_Dictionaries_Report/

   RS: Any specific tools used for the guidelines document, images, etc.?
   For the document see specs: https://github.com/w3c/respec/wiki and the Bilingual Dictionaries Report for the stylesheet, etc
   For images powerpoint and draw.io tools

3. Crosslingual linking (leader Mike Rosner)

   Regular telcos, different levels of cross-lingual linking, studying pros and cons per case. Also, a position paper will be written on the topic. Outline + skeleton of the guidelines available, writing to be started soon.

4. Corpora annotation (NIF, Web annotation - leader Milan Dojchinovski)

   CC and RS have written a paper and created a dataset with NIF. No regular telco yet. Ranka has approached Milan, Christian and Max, but not yet any telcos.

5. Guidelines for Developing NIF-based NLP Services (leader Milan Dojchinovski)

   Leader is not present.

6. Wordnets (leader Fahad Khan)

   The meetings haven't officially started for the WordNet Guidelines task but the task leader ahad Khan) has been in touch with members of the Global WordNet Association and has also had discussions with the lead on the Polish WordNet project (the approach of that WN is slightly different from the GWA schema). He plans to start meetings around mid-November.

7. Guidelines for LLOD aware services (leader Andon Tchechmedjiev)

   The topic is related with WG2 survey paper. Assignments have been set. Reading will finish by end of December, writing will start in January. Also in January we will start working on the draft for the guidelines.

8. LLOD for Under-resourced languages (Giedre Valunaite Oleskeviciene and Ranka Stanković)

No calls have started yet, they have some initial ideas how to proceed. They will start organising the first call in November.

9. Neuro-symbolic LLOD (leader Katerina Gkirtzou)

The first meeting for the Neuro-symbolic Guidelines took place on the 17/10/2023 and we discussed the following topics:

1. What do we mean with the term Neurosymbolic? Which tasks fall under this area? → Tentative consensus on differentiation of two aspects: data representation and neural services. Two documents of guidelines
2. How are we going to approach the topic (mode of communication)? →
   a. Decision not to split efforts into the two different aspects of data and services (interrelated topics)
   b. Github issues would be driving the discussion (asynchronous mode)
   c. Telco for coordination (synchronous mode) - Once per month.
   d. Git repo used for backing gdrive documents
3. Git issues with the next open questions to work on:
   a. [Data models and metamodels for neurosymbolic learning on LLOD?](#)
   b. [Potential applications and use-cases?](#)
   c. [Scope of the recommendation](#)
4. Open discussion: which aspect should we draft first? Voting pending until the next meeting, hoping the relevant effort in git issues would provide input to have an informed decision.

10. Licensing and metadata (leader Penny Labropoulou)

No update for last time. A new meeting will be organised soon.

11. Multimodal data (Ineke Schuurman)

Ineke no time to work on this topic per se for the months to come.
JG: Suggest to drop this topic. We said to give it a little bit more time
IS: +1
KG: +1
AT: +1 Adding it later on if people are interested.
Decision: drop the topic for now.

12. Sign languages (Ineke Schuurman and Andon Tchechmedjiev)

Efforts are progressing within the Multimodality task of WG3, after two calls with only 4 people, we have managed to expand participation by moving the regular slot to Fridays 14-15 CET.
There is a working document being enriched with examples and surveying work that has been done by Thierry.
https://docs.google.com/document/d/17bvdTXpXtDCL-DgLpGWUKPrtyxaGghgHyRC81pl1IT8/edit
Andon will fork the Sign Language Ontology created by Thierry (first partial version) and continue enriching the ontology to accommodate the examples and edge-cases identified.
Really under resourced, still a lot of work to be done.

13. Lexicographic LLOD (Jorge Gracia del Río + LiLa person?)

We haven't formally started yet. The first meeting will take place on November the 7th with Lila people and anyone else who is interested in

Other Issues

- Action: Jorge to check with the SO about possibility of merging the two   Nexus Linguarum deliverables regarding Guidelines and Best Practices. → Pending decision

Next Meeting:

Organize next meeting early December.

Notes from Nexus Plenary Meeting (07/09/2023) and W3C community day (12/09/2023)

- Missing leaders:
  - Sign languages: Ineke Schuurman and Andon Tchechmedjiev
  - LLOD for Under-resourced languages: Giedre Valunaite Oleskeviciene and Ranka Stanković
  - Multimodal data: Ineke Schuurman
  - Lexicographic LLOD: Jorge Gracia del Río + LiLa person
- NexusLinguarum deliverable on Guidelines:
  - Summary document with points to the BPMLOD guidelines
  - Jorge to check if a single deliverable (WG1+WG2) is ok
  - Guidelines for Encoding Domain Labels for Linked Data Lexical Resources in RDF : https://github.com/anasfkhan81/EncodingDomainLabelsRDF/blob/main/Guidelines.md will be part of the Nexus deliverable

Meeting 21/07/2023

Agenda and connection details

https://www.w3.org/community/bpmlod/wiki/Teleconference,_21/07/2023

Participants: (Please add your name)

Jorge Gracia (Univ. of Zaragoza)
Katerina Gkirtzou (Athena Research Center)
Victor Rodriguez (UPM)
Ranka Stanković, University of Belgrade
Thierry Declerck (DFKI)
Ineke Schuurman (KU Leuven (BE))
Ilan Kernerman (Lexicala)
Milan Dojchinovski (CTU Prague/INFAI)
Andon Tchechmedjiev (IMT Mines Alès)
Patricia Martín-Chozas (UPM)
Mike Rosner (UM)
Bev Corwin
Penny Labropoulou (ILSP/ARC)
Fahad Khan (CNR-ILC)
Christian Chiarcos (U Augsburg)
Gilles Sérasset (Université de Grenoble)
Giedre Valunaite Oleskeviciene (Mykolas Romeris University)
Patricia Martín-Chozas (UPM)
Thierry Declerck (DFKI)
Dagmar Gromann (University of Vienna)

Regrets:
Milan Dojchinovski (CTU Prague/INFAI)

MINUTES

General information and instructions regarding the modus operandi for contribution, participation assignment update

2. Repository with the current version of BPMLOD guidelines

   The current version in the report repository (https://github.com/bpmlod/report) has been tagged as agreed on the previous meeting

3. Current Modus operandi

   https://docs.google.com/document/d/1EQPywP6_0qd9uiFJU_Wn4A1Yqx9z-9RbiTsgu2nKFtY/edit

4. Expression of interest + github ids:

https://docs.google.com/spreadsheets/d/1VZgOOV3bE_RLLRNu_A97ztpv0cuDuRp6I6Asy5t1fmA/edit#gid=1214801611
Please verify that the github id is correct, especially those marked in red background.

Check status on proposed guidelines and best practices:

1. Terminology generation (leader Patricia Martín-Chozas)

No effort in this direction yet. The Terminology Ontolex module has started in the past couple of months, still in early stages.

2. Bilingual dictionaries (leader Jorge Gracia)

No progress yet. Jorge has limited availability at the moment and focuses on cross-lingual linking. Ideally, Jorge would like to have a first draft-suggestion.

3. Crosslingual linking (leader Mike Rosner)
- Developing BPMLOD guidelines, best practices, use cases for cross-lingual interlinking
- Our current ideas are inspired by exploring similarities between cross-lingual links and the translational equivalence relations that are made explicit in the well-known Vauquois Triangle model for describing machine translation systems.
- There is a paper in LDK regarding this issue from Frances Gillis-Webber (*Refinement of the Classification of Translations - Extension of the vartrans Module in Ontolex-Lemon*). To check the proceedings when they are published.

4. Corpora annotation (NIF, Web annotation - leader Milan Dojchinovski)

No active discussion to our knowledge yet.
An initial discussion started in Zagreb, during the datathon.

5. Guidelines for Developing NIF-based NLP Services (leader Milan Dojchinovski)

No active discussion to our knowledge yet.
An initial discussion started in Zagreb, during the datathon.

6. Wordnets (leader Fahad Khan)

No progress made, planning to work on this after the summer

7. Guidelines for LLOD aware services (leader Andon Tchechmedjiev)

No progress made. The guidelines is tied with the paper of WG2. In that task, papers have been gathered and gone through the process of first evaluation of adding or dismissing them. After summer break selected paper will be assigned to people and start working on them.

8. LLOD for Under-resourced languages **(Requires leader)**
- No information, nothing has started yet.
- Use the policy brief as a starting point.

- Giedre and Ranka volunteered to co-lead the effort
- [LREC paper](#) on cross linguality
- European Language Equality ([ELE](#)) project as a source for specific tasks that are pertinent to understanding the current status of different EU languages.
- ELE book: https://link.springer.com/book/10.1007/978-3-031-28819-7

9. Neuro-symbolic LLOD (Requires leader)

- Info: there is a new journal on [Neuro-Symbolic AI](#)
- Informal discussions between people who shared interest. Katerina will share a poll for a meeting in September (after LDK), maybe lead the task if there are no objections.
- Check out [SSHOC marketplace](#) for inspiration of use case
- Set up a brainstorming document until the planned meeting.

10. Licensing and metadata (leader Penny Labropoulou)

- initially focusing on licensing; defined scope (i.e. no legal experts' recommendations; focus on semantic representation of licences and licence assignment on resources);
- contacted CLIC (CLARIN Legal and Ethical Issues Committee) president for joining efforts (also collecting best practices); we can re-use/refer to CLIC resources (e.g. Legal Information platform)
- collecting use cases on Linked Data resources, how existing resources are licenced
- discussed on representation of licences (using ODRL)

11. Multimodal data (Requires leader)

- No progress made. No active discussions yet.
- Fahad wants to work on this, but does not have the time to lead the effort. Similar, for Dagmar and Andon.
- Ontolex multimodal effort focused mainly on sign languages.
- If no progress is made in the next few meetings, maybe we will drop them.

12. Sign languages (Requires leader)

We already sent email Ineke to take the lead in this effort, as she is a co leader in multimodal ontolex module.
Answer: She is the most competent on the question in nexus, would need to be assisted on LLOD aspects (co-leadership?). Will talk to some other people like Francis Bond about ongoing efforts to incorporate SLs in OMWN, see what could be coordinated.

13. Lexicographic LLOD (Requires leader)

- Jorge cannot lead this at the moment. If no one volunteers, maybe in mid-term he could lead it.
- Jorge also plans to communicate with LiLa project for input, guidance.

Other Issues

- How to integrate existing recommendation efforts by nexus members for lexicographic LLOD? Guidelines for Encoding Domain Labels for Linked Data Lexical

Resources in RDF :
https://github.com/anasfkhan81/EncodingDomainLabelsRDF/blob/main/Guidelines.md

- JG: As it is developed with the Nexus, it could perfectly moved as part of BPMLOD practices, either as they are, or as part of an existing effort. In any case they can be part of the delivered guidelines as they are now
- FK: Maybe use these guidelines as a first draft as it is fairly mature. And within BPMLOD an updated version can be created when needed.

Next Meeting:

Probably in September after LDK. A poll will be sent in early September.

Meeting 13/03/2023

Agenda and connection details

https://www.w3.org/community/bpmlod/wiki/Teleconference,_13/03/2023

Participants: (Please add your name)

Jorge Gracia (Univ. of Zaragoza)
Katerina Gkirtzou (Athena Research Center)
Ranka Stanković, University of Belgrade
Thierry Declerck (DFKI)
Ineke Schuurman (KU Leuven (BE))
Ilan Kernerman (Lexicala)
Milan Dojchinovski (CTU Prague/INFAI)
Andon Tchechmedjiev (IMT Mines Alès)
Patricia Martín-Chozas (UPM)
Mike Rosner (UM)
Bev Corwin
Penny Labropoulou (ILSP/ARC)
Fahad Khan (CNR-ILC)
Christian Chiarcos (U Augsburg)

Regrets:
Gilles Sérasset (Université de Grenoble)
Giedre Valunaite Oleskeviciene (Mykolas Romeris University)
Patricia Martín-Chozas (UPM)
Thierry Declerck (DFKI)

MINUTES

1) Consolidate the list of proposed guidelines and best practices - Status update

See https://www.w3.org/community/bpmlod/wiki/Guidelines_and_best_practices_for_LLOD

Check status of proposed guidelines:
- Terminology generation (TBX)
- Bilingual dictionaries
- Crosslingual linking
- Corpora annotation (NIF, Web annotation)
- Guidelines for Developing NIF-based NLP Services
- Wordnets.
- Guidelines for LLOD aware services
- LLOD for Under-resourced languages

- Neuro-symbolic LLOD
- Licensing and metadata
- Multimodal data
- Sign languages
- Lexicographic LLOD

On terminology

CC: May on guidelines for terminology generation to wait until the Ontolex terminology starts

JG: Maybe start in parallel with ontolex terminology meeting. The focus of the guidelines could be the minimal part of what exists and then evolve with Ontolex terminology vocabulary.

CC: Discussion on guidelines should be sent to both list (BMPLOD + Ontolex) so people be aware of what is going on.

Bilingual dictionaries

JG: I have a clear idea where to start. I will proceed in initiate the effort in April

Crosslingual linking

MR: I haven't a change to report on this. The next move should be to meet the people who showed interest (Joege, Katerina, Ilan, Gilles, Thierry) and proceed from there

Corpora annotation (NIF, Web annotation)

MD: There are some ideas, but nothing concrete yet. We will start in April.

RS: I have a question. I am preparing a dataset and I am confused about 2.0 and 2.1, which one to use and follow. On which version the guidelines will be on?

MD: We need to look the details, both go under the same namespace.

RS: From chat "https://nif-docs.readthedocs.io/_/downloads/en/latest/pdf/

nif:EntityOccurrence nif:confidenceProperty nif:entityOccurrenceConf ; nif:provenanceProperty nif:entityOccurrenceProv ."

MS: NIF 2.1 is an extension from 2.0. We need to see the details. Maybe more people face the same issues, so we need to elaborate on the guidelines

Chat: Ranka Stanković

https://nif-docs.readthedocs.io/_/downloads/en/latest/pdf/nif:EntityOccurrence nif:confidenceProperty nif:entityOccurrenceConf ; nif:provenanceProperty nif:entityOccurrenceProv .

Chat CC: so everything except for provenance and con fidence is just inmherizted from NIF 2.0?

AT: I suggest that this a good starting point to initial on the guidelines.

Guidelines for Developing NIF-based NLP Services

?

Wordnets.

FK: I have promised to contact a number of people that work on wordnets. I haven't yet, but I will. I will also initiate this on April or May.

Guidelines for LLOD aware services

AT: The process will be in parallel with WG2 task on the survey. We are currently filtering the paper in that task

LLOD for Under-resourced languages

JG: As far as I know, I don't anything has happen. I would be a contributor. We need a leadership position on this.

CC: What is the exact scope on this?

JG: The initial idea is to take the policy for under-resource language and convert it into some guidelines. The same from the LREC paper from Task 1.3

Chat CC: we could reach out to robert forkel on LLOD for underresourced languages

Neuro-symbolic LLOD

AT: We have some exchange with Christian with linking and embedding. No leadership yet. One representative from Ontolex-Frac and the guidelines to consolidate among the two communities.

Licensing and metadata

PL: No much progress on this yet. I had some email exchange with Victor an expert on ORDL licencing. He is willing to contribute. We need to see how to start working on it. It will be really helpful to have an idea of structure the guidelines

AT: Christian had the same question on the chat. The previous guidelines could be the guide. Today we could define the way we interact with each other and later on see on the structure itself.

JG: The deliverable itself will be a list of pointers to each guideline. I also like the idea of the reference cards from LIDER project.

Multimodal data + Sign languages

AT: What is the overlap of guidelines for multimodal data and sign languages?

FK: We are all waiting for Thierry who is in charge of this. I think there is a meeting on Wednesday. Anyone who is interested should attend the meeting on wednesday to push to the this direction of guidelines as well.

Lexicographic LLOD

FK: I got a STMS to work on dictionaries on LLOD. This might be a series of guidelines on specific task. I have started working on this by focusing on lexicographic resources.

DECISION: Unless there are other ideas, we can consolidate this list. There don't seem to be any major changes in the list.

2) Summary of chairs' discussion about the repository

We will be tagging the current version in the report repository (for easier reference) and we will move current master content into a tag specific branch, so that new guidelines could be a work in progress in the master branch

Main git repo : https://github.com/bpmlod/report

3) Consultation about repository architecture for new guidelines
   a) Single repository with independent folders for each recommendation. Shared git issues space across all guidelines. Maybe need for labels per issue to refer to reference guidelines
   b) Individual repositories for each recommendation effort consolidated as submodules in the main repository.

   Decision: in favour of separate repositories per guideline.

   ACTION (CHAIRS): will check whether git submodule or git tree will be used for the consolidation.

4) Discussion about the protocol for technical exchanges

   We suggest that we follow the W3C recommendation to have the technical discussions as Github issues. If you prefer to use another venue of communication and discussion, then we would like to ask that someone takes care of creating github issues that summarise main discussions that happened in the other venues regularly.

   JG: It make sense, maybe write a short procedural information, so people know how to proceed.

   CC: I think github issues are the right choice for specific todos. but certainly not for broader discussions. Pointing from the wiki to github issues.

   JG: I wouldn't expect too much technical discussion in this activity. This is more an effort to put guidelines on specific issues. Find examples, use cases, etc. My feeling that no much technical issues would arise

   AT: This to me feels like a technical issue. Keeping updates on git issues.

   JG: Let's start with github issues as recommendation and see how it goes.

   AT: We won't avoid having issues resolved in git issues, wiki or on the meeting.

   CC: what issues can be used for is streamlining workflows and coordination ("project tables"): https://github.com/features/issues:

ACTION (Chairs): Write a document explaining the main organisation idea + Added on wiki and send an email to inform everyone.

5) Please add your Github id:

   a) Cosimo Palma: https://github.com/Glottocrisio

b) Andon: https://github.com/twktheainur
c) Ranka Stanković: https://github.com/rankastankovic
d) Fahad: anasfkhan81
e) Milan: m1ci
f) Katerina Gkirtzou: https://github.com/gkirtzou
g) Thierry – Declerck (Github ID): https://github.com/Declerck
h) Penny Labropoulou: pennyl67
i) Ilan Kernerman: ilank9
j) Mike Rosner: rosnermike
k) Christian Chiarcos: chiarcos
l) Goedre Valunaite Oleskeviciene: Giedre-hub

ACTION (KG): Write an email, requesting the github ids of people interested in participating in the guidelines.
ACTION (KG): Doodle for the next coordination meeting.

NEXT STEP: Each group of guidelines should start organising its material and getting people involved.

From MEET CHAT:

Christian Chiarcos
12:16

NIF 2.1 is half-documented only, AFAIK
same question from my side, actually ;)
Ranka Stanković
12:17

https://nif-docs.readthedocs.io/_/downloads/en/latest/pdf/
nif:EntityOccurrence nif:confidenceProperty nif:entityOccurrenceConf ;
  nif:provenanceProperty nif:entityOccurrenceProv .
Christian Chiarcos
12:18

so everything except for provenance and con fidence is just inmherizted from NIF 2.0?
Ranka Stanković
nif:EntityOccurrence
Christian Chiarcos
12:20

wasn't there something line identRef, in NIF 2.0?
Christian Chiarcos
12:24

we could reach out to robert forkel on LLOD for underresourced languages
Ranka Stanković
12:33

https://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core/nif-core.html has no identRef, it is in 2.1
https://nif.readthedocs.io/en/latest/prov-and-conf.html#spot-class-scenario
Christian Chiarcos
12:35

@Ranka: https://www.w3.org/TR/its/
its:taIdentRefPointer

Ranka Stanković
12:36

@Christian, thanks
Christian Chiarcos
12:36

its is linked from https://persistence.uni-leipzig.org/nlp2rdf/
unlike NIF, ITS is a W3C standard. so that's "more normative"
Christian Chiarcos
12:39

pro 3.b)
(but not git submodule, but git subtree)
Christian Chiarcos
12:42

https://www.atlassian.com/git/tutorials/git-subtree
Christian Chiarcos
12:46

I think github issues are the right choice for specific todos. but certainly not for broader discussions.
but those can be part of the accompanying document
Christian Chiarcos
12:50

sometimes, discussions occur in github issues, too. problem is to arrive at a concensus:
https://github.com/TEIC/TEI/issues/1860
Christian Chiarcos
12:52

so, I guess there will be issues + meetings and minutes ;)
Christian Chiarcos
12:53

what issues can be used for is streamlining workflows and coordination ("project tables"):
https://github.com/features/issues

Meeting 15/02/2023

Agenda and connection details

Participants:

Jorge Gracia (Univ. of Zaragoza)
Katerina Gkirtzou (Athena Research Center)
Ranka Stanković, University of Belgrade
Thierry Declerck (DFKI)
Ineke Schuurman (KU Leuven (BE))
Ilan Kernerman (Lexicala)
Milan Dojchinovski (CTU Prague/INFAI)
Andon Tchechmedjiev (IMT Mines Alès)
Patricia Martín-Chozas (UPM)
Mike Rosner (UM)
Bev Corwin
Penny Labropoulou (ILSP/ARC)
Fahad Khan (CNR-ILC)

MINUTES

1. Voting on the renewal of the panel of chairs (see previous [email])

From previous email:

> [...] we received two last minute nominations for the panel of chairs,
> in addition to the proposed four initial ones (see the minutes for more
> details). Thus, we decided to postpone the voting until having informed the
> rest of the group about this.
>
> Having considered that six chairs is an ineffective number, we (current
> chair and candidates) have met after the telco and arranged a consensus
> list that adds gender balance to the initial list, as well as balance
> between BPMLOD and other W3C groups in coordination tasks. The proposed
> list is:
>
> * Katerina Gkirtzou (Athena Research Center, Greece)
> * Milan Dojchinovski (Czech Technical University in Prague, Czechia;
> representing NexusLinguarum WG1)
> * Andon Tchechmedjiev (IMT Mines Alès, France; representing NexusLinguarum WG2)

Jorge Gracia (University of Zaragoza, Spain; current BPMLOD chair)

I myself [Jorge] plan to stay a bit longer as chair to allow for a smooth transition with the new chairs, and step down at some point, thus recovering the original number of three chairs for the group.

DECISION: Approved

2. Stabilise the list of proposed guidelines and best practices. Next steps.

See https://www.w3.org/community/bpmlod/wiki/Guidelines_and_best_practices_for_LLOD

Check status of proposed guidelines:
- Terminology generation (TBX) - discussion will start taking current paper and termlex module as starting point and update existing guidelines
- Bilingual dictionaries - to start, idea is updating Aperium example to ontolex, maybe with inputs with DBdary and other examples
- Crosslingual linking - discussing different links and CL resources that might be built, prons and cons of every kind of links (e.g., at sense level, at entry level, …)
- Corpora annotation (NIF, Web annotation) - novelties of last NIF version (2.1). For Web Annotation Christian (maybe Max) could help
- Guidelines for Developing NIF-based NLP Services - collect more recent use cases, update with latest upgrade (2.1). Ranka can contribute with NIF annotated data.
- Wordnets. Fahad starts compiling guidelines based on poster session of last globalWordnet conference
- Guidelines for LLOD aware services- link with survey efforts in WG2, which is closely related topic-wise
- LLOD for Under-resourced languages - related with Sina's Virtual Mobility Grant
- Neuro-symbolic LLOD - need clarify objectives. Can FrAC ideas be reused? Maybe relation with T3.1 (current position paper in preparation)
- Licensing and metadata - relation with data spaces; ODRL; GPDR; Metashare ontology
- Multimodal data - very related with the sign lang. one, merge at some point?
- Sign languages

Ilan> see overlap between three guidelines: bilingual, cross-lingual, lexicography
ADD definitions (what we mean by bilingual, cross-lingual, etc.) in the guidelines
Mike> we did that in previous paper (let's check)
Ineke> consider sign lang. dictionaries (like to join Wordnet and Sign Languages)
Thierry> Here the link to documents already available for the Sign Language Best Practices
https://drive.google.com/drive/folders/113rtGuUZ3IDOvSYcvNFn4JZNNV9fw4Vu
I would be interested in joining the Wordnets group (and I have to leave the one on Under-resourced language, sorry!)

3. Where to store technical discussions? as Github issues (as suggested by R. Ishida)? See [thread](#)

From R. Ishida email:

"I should mention, btw, that this is not a simple matter of open preference – the group should have a strong argument for using emails for technical discussions that takes into account the people outside the group who will interact with it. (Email is typically used for admin related topics, like this one, though.) The W3C actually discourages use of emails for technical discussions by Working Groups, and use of GH issues is the standard mode of operation. [...]. The difference in effectiveness of reviews since we moved away from email threads is like night and day. Use of GH issues make threads findable and much more easy to read, it also avoids splintering (in the way this thread has), and can be labelled for automatic processing, crossreferencing, etc."

From C. Chiarcos response:

"I think it's in fact a difference whether a discussion is more strategic or whether it is specific to, say, a vocabulary under discussion or an emerging standard. This discussion here is strategic, indeed. But for the latter case, if it is maintained in GitHub (and most vocabularies from the wider context of this mailing list are, e.g., OntoLex, lexinfo, NIF or the Metashare ontology), *and* the contributors actually engage with it via GitHub, GH issues make a lot of sense. Talking about OntoLex: [...] at least as far as OntoLex-Morph and OntoLex-FrAC are concerned, the primary channel of communication are minute documents and the specification itself. [...] Enforcing such a shift has obvious benefits, but it may also lead to scaring away potential contributors. [...]
Given the cons and pros, that should really be openly discussed, either in the group as a whole or among the chairs, at least. There is a benefit in transparency, but also the risk of losing parts of the user/contributor base -- and developing best practices without having their potential users involved is clearly a bad idea. We cannot really do multilingual linked open data without having the language people on board."

From R.Ishida response:

"For what it's worth, we've also in the past had experiences such as the following with groups that contained people who were previously unaware of GitHub:

a. one email list that was moving slowly suddenly came to life once we switched to GitHub, and has been very active since.

b. in another group there were indeed people who resisted the move away from email because GH was something new and unknown, but eventually they realised the benefit and that (GH issues especially) is really easy to use, and now the whole group has all technical discussions in GH.

Btw, if useful, i created some simple, picture-based introductions to using GitHub issues that i can point you to (for people who were previously inexperienced and fearful).

Just mentioning it as an alternative possibility. I guess we'll see what the chairs decide."

Penny> used GH issues in previous projects and it's useful, in favour of that
Andon, Keteina, Milan, … +1

4. AOB

Next meeting in one month, Katerina to circulate a doodle poll

Github IDs:

Cosimo Palma: https://github.com/Glottocrisio
Andon: https://github.com/twktheainur
Ranka Stanković: https://github.com/rankastankovic
Fahad: anasfkhan81
Milan: m1ci
Katerina Gkirtzou: https://github.com/gkirtzou
Thierry – Declerck (Github ID): https://github.com/Declerck
Penny Labropoulou: pennyl67
Ilan Kernerman: ilank9

Meeting 17/01/2023

Agenda and connection details

https://www.w3.org/community/bpmlod/wiki/Teleconference,_17/01/2023

Participants:

Fahad Khan (ILC)
Katerina Gkirtzou (Athena Research Center)
Jorge Gracia (University of Zaragoza)
Penny Labropoulou (ILSP/ARC)
Mike Rosner (UMalta)
Andon Tchechmedjiev (IMT Mines Alès)
Patricia Martín-Chozas (UPM)
Hugo Gonçalo Oliveira (CISUC, University of Coimbra)
Ranka Stanković (University of Belgrade, Serbia)
Dagmar Gromann (University of Vienna)
Gilles Sérasset (Université de Grenoble)
Ilan Kernerman (Lexicala)
Giedre Valunaite Oleskeviciene (Mykolas Romeris University)

MINUTES

1. Summary of future plans for guidelines and best practices for multilingual linked open data. From previous email (5/12/22):

> It has been a long *while since this group became "dormant" after the*
>
> *successful release of a good number of WC3 community group reports*
>
> *<https://www.w3.org/community/bpmlod/>. Now, in the context of the*
>
> *NexusLinguarum <https://nexuslinguarum.eu/> COST Action, we would like to*
>
> *expand and update the guidelines and best practices for multilingual linked*
>
> *open data that we produced, in two directions:*
>
> *- Generation, interlinking, publication, and validation of LLOD.*
>
> *- Linguistic Linked Open Data-aware Natural Language Processing.*
>
> *Since BPMLOD is the natural forum for this, we would like to move here such*
>
> *discussion and involve more interested partners, to produce a new round of*
>
> *W3C reports by the end of next year.*
>
> *This recently published paper, containing an analysis of how things are and*

what is missing, could serve as starting point:

Khan, A. F., Chiarcos, C., Declerck, T., di Buono, M. P., Dojchinovski, M., Gracia, J., Oleskeviciene, G. V., & Gifu, D. (2022). A Survey of Guidelines and Best Practices for the Generation, Interlinking, Publication, and Validation of Linguistic Linked Data <https://aclanthology.org/2022.ldl-1.9.pdf>. Proc. of the 8th Workshop on Linked Data in Linguistics (LDL 2022) at LREC 2022, 69–77.

2. Draft an initial list of possible topics

(added my name (ineke), Febr 15th)

Expressions of interest

topic	proposer	interested people	comments / relevant links
Terminology generation (TBX)	Patricia	Andon, Maria Pia, Fahad	https://www.w3.org/2015/09/bpmlod-reports/multilingual-terminologies/
Bilingual dictionaries	Jorge	Ilan, Fahad, Gilles	http://www.w3.org/2015/09/bpmlod-reports/bilingual-dictionaries/
Crosslingual linking	Mike	Jorge, Katerina, Ilan, Gilles, Thierry	related to Nexus T1.3 activities and position paper
Corpora annotation (NIF, Web annotation)	Milan	Ranka	see reference card guidelines document
Guidelines for Developing NIF-based NLP Services	Milan	Ranka Andon Katerina Gkirtzou	
Wordnets	Fahad	Mike , Ineke	
Guidelines for LLOD aware services	Andon	Katerina, Patricia	https://bpmlod.github.io/report/LLOD-aware-services/index.html
LLOD for Under-resourced languages		Giedre, Jorge, Ranka, Fahad, Thierry	see policy brief and LREC paper
Neuro-symbolic LLOD		(Andon), Hugo, Dagmar, Katerina,	

		Cosimo	
licensing, metadata	Penny	(Ilan), Fahad	Lider reference cards
multimodal data		Andon, Fahad, Thierry	
sign languages	Thierry	Fahad, Thierry, Ineke	Task 3.4 of NexusLinguarum
lexicographical data (lexicog)		Jorge, Ilan	

Fahad: go back to the position paper and check which gaps we need to address. Translate into other languages.

To publish in Zenodo afterwards (to have a DOI)

3. Voting on the renewal of the panel of chairs. Proposal:

* Dr Jorge GRACIA (University of Zaragoza) [already CHAIR]

* Dr Andon TCHECHMEDJIEV (IMT Mines Alès, France)

* Dr Milan DOJCHINOVSKI (Czech Technical University in Prague, Czech Republic)

* Dr Anas Fahad KHAN (Istituto della Linguistica Computazionale, Italy)

can we add women chairs too

Today we add two volunteers: Katerina Gkirtzou, Patricia Martin-Chozas
Voting POSTPONED until the rest of the group is informed