



OLIF2 Review Meeting

OLIF2 Consortium
Walldorf, Germany
November 14, 2000



Participants

Hartmut Bohn, *SAP*

Andrew Bredenkamp, *DFKI*

Mike Dillinger, *Logos*

Daniel Grasmick, *SAP*

Anke Holler, *IBM*

Hubert Lehmann, *Linguattec*

Christian Lieske, *SAP*

Susan McCormick, *SAP*

Carlo Mergen, *EC*

Yves Mahé, *Xerox*

Christel Nahkor, *IBM*

Markus Pilzecker, *DFKI*

Peter Quartier, *Lotus*

Johannes Ritzke, *Sail Labs*

Laurent Romary, *SALT*

Jean Senellart, *Systran*

Christine Thielen, *SAP*

Gregor Thurmair, *Sail Labs*

Michael Wetzel, *Trados*

Purpose of Meeting

- ★ To review, amend and approve proposals for the structure and content of version 1 of the OLIF2 exchange format
- ★ To plan for the final development, implementation and testing of OLIF2

Agenda

9.00 – 9.30	Welcome and introductory remarks	<i>Daniel Grasmick</i>
	Review and approve agenda	<i>Susan McCormick</i>
9.30 – 10.30	Discuss proposal for OLIF2 file header	<i>Christian Lieske</i>
10.30 – 10.45	Coffee break	
10.45 – 11.15	How to represent user-based data	<i>Christian Lieske</i>
11.15 – 13.00	Discuss proposal for structure and content of OLIF2 file body	<i>Susan McCormick</i>
13.00 – 14.00	Lunch	
14.00 – 14.30	Continue discussion of proposal for structure and content of OLIF2 file body	<i>Susan McCormick</i>
14.30 – 15.30	Review of efforts to include inflection classes	<i>Gregor Thurmair</i>
15.30 – 15.45	Coffee Break	
15.45 – 16.00	Discuss concertation with SALT	<i>all participants</i>
16.00 – 16.15	Discuss Asian languages, esp. Japanese	<i>all participants</i>
16.15 – Close	Plan for final phase of OLIF2 development; final specifications and DTD; testing and implementation	<i>all participants</i>



Technical Issues: File Header

✱ Reviewed by Christian Lieske, *SAP*



User-based Extensions

- ✱ How open should OLIF2 be?
- ✱ User-defined elements?
- ✱ User-defined attributes?
- ✱ Mechanisms
 - ✱ attached URL's
 - ✱ namespaces
 - ✱ restrictive vs. nonrestrictive content models
 - ✱ etc.



Structure of the File Body

List of entries that contain data that is grouped according to the linguistic/lexical/terminological character of the information being represented.

Main groups are:

- **monolingual:** defines monolingual data.
- **cross-reference:** defines cross-reference relations between the given entry and other entries in the lexicon in the same language.
- **transfer:** defines bilingual transfer relations between the given entry and other entries in the lexicon in different languages.

Identifying an OLIF2 Entry

Semantic unit that is identified uniquely by a set of five basic monolingual features:

- **canonical form:** the entry string, represented in canonical form in accordance with OLIF2 guidelines (to be published in conjunction with SALT)
- **language:** the language represented by the entry string
- **part of speech:** the part of speech represented by the entry string
- **subject field:** the knowledge domain to which the lexical/terminological entry is assigned.
- **reading number:** the number identifier used to distinguish readings for entries with identical values for *canonical form*, *language*, *part of speech*, and *subject field*.

Sample OLIF2 Entry

```
<entry>  
  <mono>  
    <canForm>Briefkurs</canForm>  
    <language>de</language>  
    <ptOfSpeech>noun</ptOfSpeech>  
    <subjField>gac-fi</subjField>  
    <readingNo>1</readingNo>  
  </mono>  
  <transfer>  
    <canForm>bank selling rate</canForm>  
    <language>en</language>  
    <ptOfSpeech>noun</ptOfSpeech>  
    <subjField>gac-fi</subjField>  
    <equival>full</equival>  
  </transfer>  
</entry>
```



Content of Entries

☀ Monolingual features

- ☀ basic
- ☀ administrative
- ☀ morphological
- ☀ syntactic
- ☀ semantic

☀ Cross-reference features

☀ Transfer features



Inflection Classes

✦ Reviewed by Gregor Thurmair, *Sail Labs*

Concertation with SALT

- ✱ Currently, OLIF2 is integral part of DXLT standard
- ✱ Will review OLIF2 features/values to coalesce OLIF2 with MARTIF as much as possible (feature names, values)
- ✱ SALT will provide domain hierarchy model
- ✱ OLIF2 will provide guidelines for canonical forms



Status of Japanese

- ✦ At present, no special consideration of Japanese requirements reflected in the proposal
- ✦ Manpower/vendors currently lacking in consortium

Final Phase of OLIF2 Work

- ✱ Final format description and specifications - latest target date March 2001
- ✱ Testing and implementation plans to be decided