

Označevanje zbirke zapisnikov sej slovenskega parlamenta s smernicami TEI

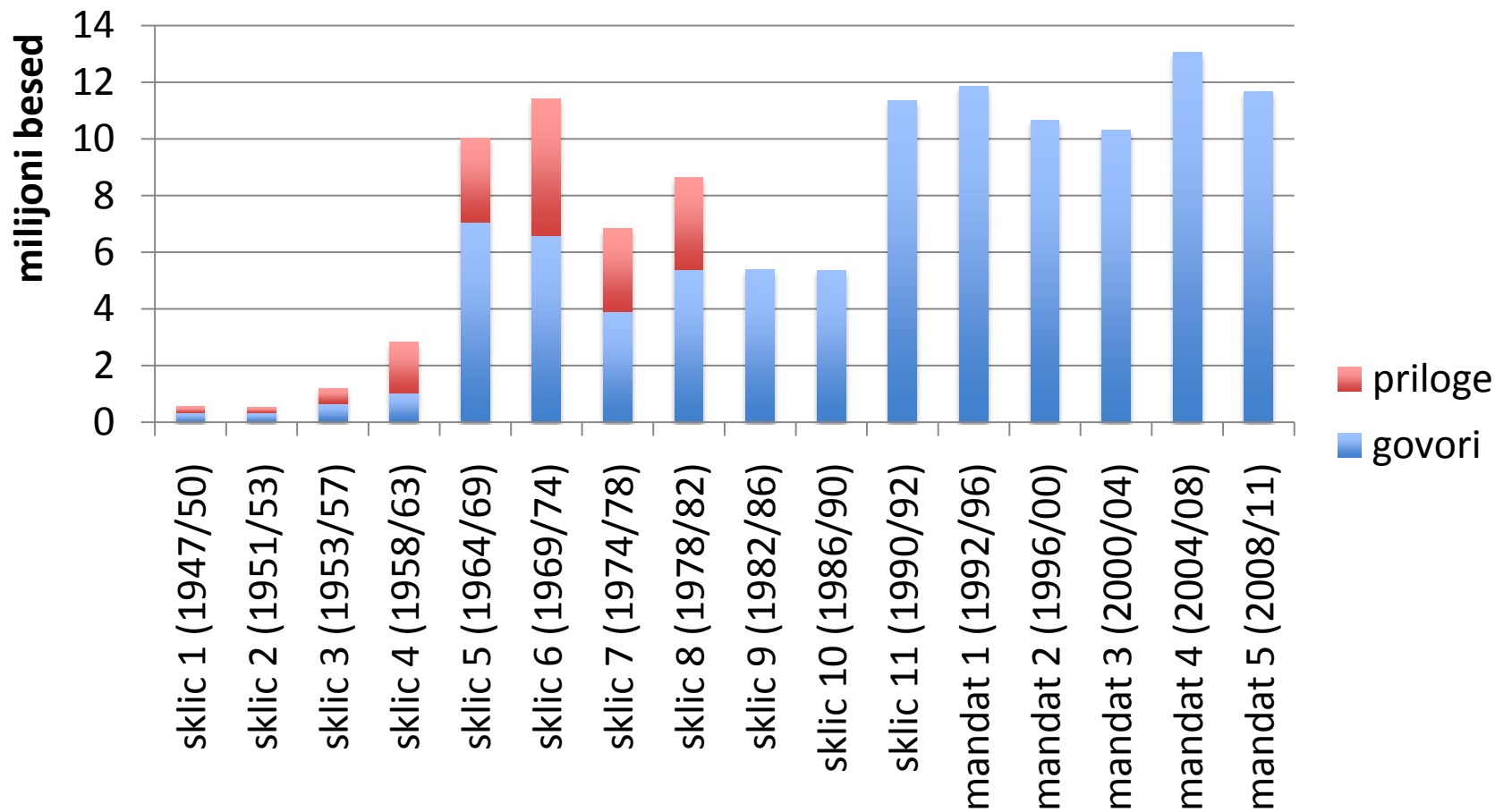
Andrej Pančur (Inštitut za novejšo zgodovino)

Mednarodna konferenca Jezikoslovne tehnologije in
digitalna humanistika

Ljubljana, 29. 9. — 1. 10. 2016

Masivni podatki (big data) več kot 115 milijonov besed

Parlamentarni govori v sejnih zapiskih, objavljeni na portalu Sistory (1947—1990) in na spletni strani državnega zbora (1990—)



Zgodovina Slovenije — Sistory: PDF format

Viri ▾ Literatura ▾ Prireditve ▾ Zgodovina in šola ▾ Digitalna humanistika ▾ Podatki ▾

sl ▾



PUBLIKACIJE

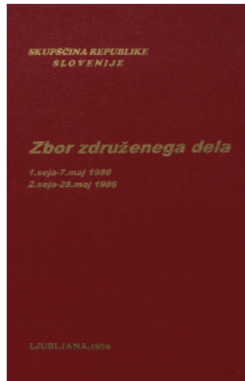
ŽRTVE

POPISI

ZIC



🏠 / SKUPŠČINA SOCIALISTIČNE REPUBLIKE SLOVENIJE (1963-1990) / 10. SKLIC (1986-1990)



Sejni zapisi Skupščine Socialistične republike Slovenije za 10. sklic (1986-1990), št. zvezka 1

Zbor združenega dela, 1. seja (7. maj 1986)

Zbor združenega dela, 2. seja (28. maj 1986)

Jezik: slovenski

Vrsta gradiva: Besedilo

Leto: 1986

Vsebina: predstaviška telesa, stenografski zapisniki, Socialistična republika Slovenija, družbeno samoupravljanje, representative bodies, stenographical minutes, Socialist Republic of Slovenia, social self-management, session, term, resumption of session, joint session

Založnik(i): Skupščina Socialistične republike Slovenije

Zbirka: Sejni zapisi Skupščine Socialistične republike Slovenije

Avtorske pravice:



This work (Sejni zapisi Skupščine Socialistične republike Slovenije za 10. sklic (1986-1990), št. 1, by Skupščina Socialistične republike Slovenije), identified by [Zgodovina Slovenije - Sistory](#), is free of known copyright restrictions.

To delo avtorskopravno ni varovano.

Trajna povezava: <http://hdl.handle.net/11686/16027>

Posamezni prispevki:


- Zbor združenega dela, 1. seja (7. maj 1986), Sejni zapisi Skupščine Socialistične republike Slovenije za 10. sklic (1986-1990), št. zvezka 1
- Zbor združenega dela, 2. seja (28. maj 1986), Sejni zapisi Skupščine Socialistične republike Slovenije za 10. sklic (1986-1990), št. zvezka 1

▲ Multimedia, povezave in prenosi



[Velikost: 48,5 Mb]

Državni zbor Republike Slovenije: HTML format



REPUBLIKA SLOVENIJA
DRŽAVNI ZBOR

English

[Domov](#) > [Delo Državnega zbora](#) > [Seje](#)

[Prijava](#)

[Kopiraj URL](#)

[NAZAJ](#) [NATISNI](#)

< • > September 2016

Pon	Tor	Sre	Čet	Pet	Sob	Ned
29	30	31	1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	1	2
3	4	5	6	7	8	9

Evidenca zapisa seje

Polni naziv telesa - Številka in vrsta seje: 50. Seja Zbora združenega dela

Datum zasedanja:

Vsebina zapisa seje

ZBOR ZDRUŽENEGA DELA

SKUPŠČINE REPUBLIKE SLOVENIJE
50. seja
(9., 10. september in 7. oktober 1992)

Sejo sta vodila: Jože Zupančič, predsednik Zbora združenega dela in Bogo Rogina, podpredsednik zbora
Seja se je pričela ob 10.10 uri.

PRESEDNIK JOŽE ZUPANČIČ: Dovolite, da pričnem po avgustovskih počitnicah z našo 50. sejo Zbora združenega dela. Slavimo majhni jubilej, petdeseto sejo in bi vas samo obvestil o naslednjem. Sedaj nas je prisotnih 64, 4 so opravičeni, upam, da bo še nekaj poslancev prišlo, tako da bo zbor še bolj številčno zastopan. Sedaj ne moremo sprejemati dnevnega reda kljub temu, da ste lahko vsi prebrali - zaradi tega, ker so dnevni redi razmnoženi na vaših mizah - kaj imamo danes in jutri pred seboj še za postoriti.

Nekaj na novo, nekaj pa je seveda tudi še starih zadev, ki jih moramo tudi zaključiti in uspešno opraviti. Mi bi imeli v tem dopoldanskem času naslednji dnevni red: najprej bi imeli skupno sejo, ki jo bo vodil gospod predsednik dr. France Bučar, in bi opravili še tista imenovanja in volitve, s katerimi smo ostali na dolgu še v mesecu juliju. In potem bi po krajšem odmoru nadaljevali s skupno delovno sejo, na kateri bi samo odločali o teh zadevah, ki so še odprte in potem skušali sprejeti volilno zakonodajo. Na podlagi te volilne zakonodaje bo pa čisto normalno lahko sprejet, če bo sprejeta ostala volilna zakonodaja, ki bi jo morali mi obravnavati v nadaljevanju te naše 50. seje.

Poročevalec

[Gradiva v obravnavi](#)

[Arhiv od 14.4.2006 do 15.7.2010](#)

Obiščite nas

Dan odprtih vrat je drugo soboto v mesecu ob 9., 10. in 11. uri.

[▶ več o tem](#)

Od masivnih podatkov (big data) do pametnih podatkov (smart data): masivni pametni podatki (big smart data). Označevanje v XML: Smernice TEI (Text Encoding Initiative)

modul za dramska besedila



P5: Guidelines for Electronic Text Encoding and Interchange

Version 3.0.0. Last updated on 29th March 2016, revision 89ba24e

Table of contents

- 7.1 Front and Back Matter
- 7.2 The Body of a Performance Text
- 7.3 Other Types of Performance Text
- 7.4 Module for Performance Texts

« 6 Verse

» 8
Transcriptions
of Speech

Home

7 Performance Texts

This module is intended for use when encoding printed dramatic texts, screen plays or radio scripts, and written transcriptions of any other form of performance.

Section [7.1 Front and Back Matter](#) discusses elements such as cast lists, which can appear only in the front or back matter of printed dramatic texts. Section [7.2 The Body of a Performance Text](#) discusses the structural components of performance texts: these include major structural divisions such as acts and scenes (section [7.2.1 Major Structural Divisions](#)); individual speeches (section [7.2.2 Speeches and Speakers](#)); groups of speeches (section [7.2.3 Grouped Speeches](#)); stage directions (section [7.2.4 Stage Directions](#)); and the elements making up individual speeches (section [7.2.5](#)

[Speech Contents](#)). Section [7.2.6 Embedded Structures](#) discusses ways of encoding units which cross the simple hierarchic structure so far defined, such as embedded songs or masques. Finally, section [7.3 Other Types of Performance Text](#) discusses a small number of additional elements characteristic of screen plays and radio or television scripts, as well as some elements for representing technical stage directions such as lighting

modul za transkribiranje govorov



P5: Guidelines for Electronic Text Encoding and Interchange

Version 3.0.0. Last updated on 29th March 2016, revision 89ba24e

Table of contents

- 8.1 General Considerations and Overview
- 8.2 Documenting the Source of Transcribed Speech
- 8.3 Elements Unique to Spoken Texts
- 8.4 Elements Defined Elsewhere
- 8.5 Module for Transcribed Speech

« 7
Performance
Texts

» 9
Dictionaries

Home

8 Transcriptions of Speech

The module described in this chapter is intended for use with a wide variety of transcribed spoken material. It should be stressed, however, that the present proposals are not intended to support unmodified every variety of research undertaken upon spoken material now or in the future; some discourse analysts, some phonologists, and doubtless others may wish to extend the scheme presented here to express more precisely the set of distinctions they wish to draw in their transcriptions. Speech regarded as a purely acoustic phenomenon may well require different methods from those outlined here, as may speech regarded solely as a process of social interaction.

This chapter begins with a discussion of some of the problems commonly encountered in transcribing spoken language (section [8.1 General Considerations and Overview](#)). Section [8.2 Documenting the Source of Transcribed Speech](#) documents some additional TEI header elements which may be used to document the recording or other source from which transcribed text is taken. Section [8.3 Elements Unique to Spoken Texts](#) describes the basic structural elements provided by this module. Finally, section [8.4](#)

Pretvorbe:

- PDF → DOCX → TEI
- HTML → TEI

M a r t a K a d u n c : Poročilo Komisije za verifikacijo pooblastil in nunitetna vprašanja Zbora združenega dela Skupščine SR Slovenije za 35. sejo ne 20.6.1984;

Zbor združenega dela flteje 130 delegatov, Za današnja sejo je predložilo Blegatska pooblastila 145 delegatov, in sicer: z gospodarskega področja 91 delegatov, s prosvetno-kulturnega področja 13 delegatov, s socialno-zdravstvenega odročja 13 delegatov, s kmetijskega področja 17 delegatov, z obrtnega področja 5 delegatov, s področja državnih organov, družbenopolitičnih organizacij in ruštev 5 delegatov ter s področja oboroženih sil en delegat.

Print Layout View

Sec 7

Pages: 6 of 112

Words: 496 of 56468



156%

M a r t a K a d u n c : Poročilo Komisije za verifikacijo pooblastil in nunitetna vprašanja Zbora združenega dela Skupščine SR Slovenije za 35. sejo ne 20.6.1984:

Zbor združenega dela šteje 130 delegatov. Za današnja sejo je predložilo Blegatska pooblastila 145 delegatov, in sicer: z gospodarskega področja 91 delegatov, s prosvetno-kulturnega področja 13 delegatov, s socialno-zdravstvenega odročja 13 delegatov, s kmetijskega področja 17 delegatov, z obrtnega področja 5 delegatov, s področja državnih organov, družbenopolitičnih organizacij in ruštev 5 delegatov ter s področja oboroženih sil en delegat.

Struktura vsebine zapisnikov sej

- dokumenti
 - kazalo vsebine
 - seznam govornikov
 - priloge
 - seje
 - didaskalije (stage-direction)
 - vsebinski sklopi (točke dnevnega reda ipd.)
 - didaskalije
 - govori
 - » didaskalije
 - » odstavki
 - didaskalija

Osnovna struktura zapisnikov sej kodiranih v TEI

```
<text>
  <front>
    <!-- možen titlePage, docImprint -->
    <div type="contents">
      <!-- kazalo vsebine; lahko tudi v back -->
    </div>
    <div>
      <!-- seznam govornikov; lahko tudi v back -->
    </div>
  </front>
  <body>
    <!-- govori -->
  </body>
  <back>
    <div type="appendix">
      <!-- priloge -->
    </div>
    <div>
      <!-- seznam govornikov; lahko tudi v front -->
    </div>
    <div type="contents">
      <!-- kazalo vsebine; lahko tudi v front -->
    </div>
    <!-- možen div[@type='colophon'] -->
  </back>
</text>
```

```
<body><!-- govori -->
  <div><!-- vrsta zbora oz. skupščine -->
    <div><!-- seja oz. zborovanje -->
      <docDate><date when="1990-07-02">dan zasedanja</date>,
        <date>možen drugi dan zasedanja</date></docDate>
      <timeline>
        <!-- navezava na tei:stage[@type='time'] -->
      </timeline>
      <castList>
        <!-- predsedujoči seje -->
      </castList>
      <stage type="time">
        <!-- čas začetka govorov -->
      </stage>
      <div><!-- pred dnevnim redom -->
        <div>
          <!-- vsebinski sklopi -->
        </div>
      </div>
      <div><!-- dnevni red -->
        <div>
          <!-- vsebinski sklopi oz. točke -->
        </div>
      </div>
    </div>
  </div>
</body>
```


Seznam govornikov in seznam predsedujočih

```
<front>
  <!-- ... -->
  <div>
    <!-- seznam govornikov -->
    <castList>
      <castItem>
        <actor xml:id="sp.DolinšekDrago">Dolinšek
          Drago</actor> 26, 46</castItem>
        <!-- ... -->
        <castItem>
          <actor sameAs="#sp.ZupančičJože">Jože
            Zupančič</actor> 56, 59</castItem>
        <!-- ... -->
      </castList>
    </div>
  </front>
  <body><!-- govori -->
    <div><!-- vrsta zbora oz. skupščine -->
      <div><!-- seja oz. zborovanje -->
        <!-- ... -->
        <castList>
          <!-- predsedujoči -->
          <castItem>
            <roleDesc>Sejo je vodil</roleDesc>
            <actor xml:id="sp.ZupančičJože">Jože Zupančič</actor>,
            <role>predsednik Zbora združenega dela</role>
          </castItem>
        </castList>
      </div>
    </div>
  </body>
```

Kodiranje govorov

```
<div><!-- vsebinski sklop -->
  <sp who="#sp.PriimekIme">
    <speaker>Ime in priimek govornika in
      morebitne oznake govornika</speaker>
    <p>Besedilo govora, besedilo <title>naslov
      sklopa</title> besedilo govora,
      <stage>komentar zapisnikarja magnetograma
        razprave</stage> besedilo.</p>
  </sp>
  <stage type="time">
    <!-- čas prekinitve in ponovnega začetka razprave -->
  </stage>
  <sp who="#sp.PriimekIme">
    <speaker>Ime in priimek</speaker>
    <ab>Govor iz klopi, ki ga je komentator zapisal v sklopu
      govora govornika iz govorniškega odra.</ab>
  </sp>
  <stage type="time">
    <!-- čas konca razprave -->
  </stage>
</div>
```

Dodatni TEI dokument: govorniki (TEI modul: Imena, datumi, osebe in kraji)



P5 Guidelines — English

P5: Guidelines for Electronic Text Encoding and Interchange

Version 3.0.0. Last updated on 29th March 2016, revision 89ba24e

Table of contents

- 13.1 Attribute Classes Defined by This Module
- 13.2 Names
- 13.3 Biographical and Prosopographical Data
- 13.4 Module for Names and Dates

« 12 Critical Apparatus

» 14 Tables, Formulae, Graphics and Notated Music

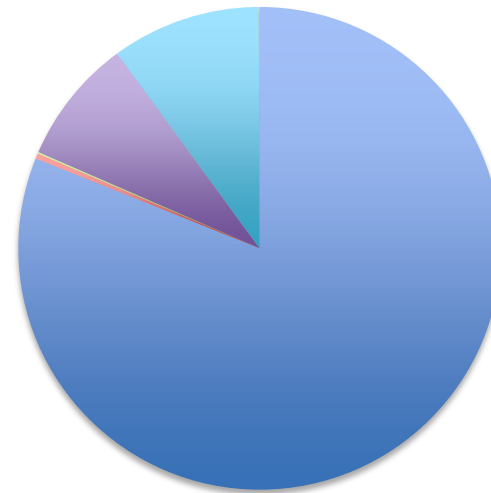
Home

13 Names, Dates, People, and Places

This chapter describes a module which may be used for the encoding of names and other phrases descriptive of persons, places, or organizations, in a manner more detailed than that possible using the elements already provided for these purposes in the Core module. In section [3.5 Names, Numbers, Dates, Abbreviations, and Addresses](#) it was noted that the elements provided in the core module allow an encoder to specify that a given text segment is a proper noun, or a *referring string*, and to specify the kind of object named or referred to only by supplying a value for the @type attribute. The elements provided by the present module allow the encoder to supply a detailed sub-structure for such referring strings, and to distinguish explicitly between names of persons, places, and organizations.

This module also provides elements for the representation of information about the person, place, or organization to which a given name is understood to refer and to represent the name itself, independently of its application. In simple terms, where the core module allows one simply to represent that a given piece of text is a *name*, this module allows one further to represent a *personal name*, to represent the *person* being named, and to represent the *canonical name* being used. A similar range is provided for names of places and organizations. The main intended applications for this module are in biographical, historical, or geographical data systems such as gazetteers and biographical databases, where these are to be integrated with encoded texts.

št. besed



- ZbZdruDel
- DruzPolZb
- ZbObc
- vlada
- poročevalci
- neznano

Dodatni TEI dokument: tematsko kazalo vsebine govorov (shema Poslovnika državnega zbora in Pravni red Republike Slovenije): TEI taxonomy

št. besed

2.3.7 The Classification Declaration

The `classDecl` element is used to group together definitions or sources for any descriptive classification schemes used by other parts of the header. Each such scheme is represented by a `taxonomy` element, which may contain either a simple bibliographic citation, or a definition of the descriptive typology concerned; the following elements are used in defining a descriptive classification scheme:

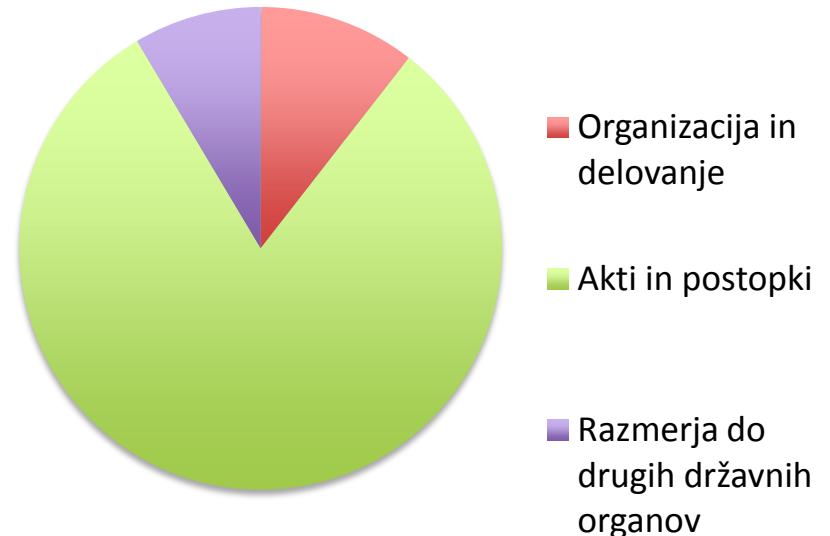
`<classDecl>` (classification declarations) contains one or more taxonomies defining any classificatory codes used elsewhere in the text.

`<taxonomy>` defines a typology either implicitly, by means of a bibliographic citation, or explicitly by a structured taxonomy.

`<category>` contains an individual descriptive category, possibly nested within a superordinate category, within a user-defined taxonomy.

`<catDesc>` (category description) describes some category within a taxonomy or text typology, either in the form of a brief prose description or in terms of the situational parameters used by the TEI formal textDesc.

« 2.3.6
The
Reference
System
Declaration
» 2.3.8
The
Geographic
Coordinates
Declaration
Home



Dostopnost:

- GitHub:
 - https://github.com/Slstory/Sejni_zapiski
 - https://github.com/Slstory/Seje_DZ
 - <https://github.com/Slstory/SlovParl>
- TEI modul za transkribiranje govorov
 - <https://github.com/DARIAH-SI/CLARIN.SI/tree/master/SlovParl>
 - Tomaž Erjavec (IJS): tokenizacija, oblikoskladenjsko označevanje in lematizacija.
 - Možnost dodajanja imenskih entitet, npr. Stanford NER za slovenščino.
 - CLARIN.SI repozitorij
 - NoSketch Engine

Corpus *SlovParl (parlament RS 1990-1992)* - statistics and info

Counts	
Tokens	3,167,912
Words	2,658,568
Sentences	162,881
Paragraphs	56,740
Documents	133

General info	
Language	Slovenian
Encoding	UTF-8
Compiled	08/28/2016 19:29:54
Tagset doc	Description
Infolink	Not specified

Lexicon sizes	
word	76,616
lc	68,439
lempos	28,856
lemma	27,899
lemma_lc	27,677
tag_en	998
tag	998

Structures and attributes

text	133
p	56,740
s	162,881
g	462,666