

European Summer School 2017

Text Mining with Canonical Text Services

Theory Session 5 – Post Processing, Advanced Requests, CTS Cloning,
Structure based Text Alignment



Federal Ministry
of Education
and Research

Text Passage Post Processing

Text Passage can be formatted differently using different “views”

```
<passage>
    O Christmas tree, O Christmas tree ! How are thy leaves so verdant ! Not only in the summertime, But even in winter is thy prime. O Christmas tree, O Christmas tree, How are thy leaves so verdant ! O Christn
    tree, O Christmas tree, Much pleasure doth thou bring me ! For every year the Christmas tree, Brings to us all both joy and glee. O Christmas tree, O Christmas tree, Much pleasure doth thou bring me ! O
    Christmas tree, O Christmas tree, Thy candles shine out brightly ! Each bough doth hold its tiny light, That makes each toy to sparkle bright. O Christmas tree, O Christmas tree, Thy candles shine out brightly !
<passage>
- <tei:TEI>
- <tei:text>
- <tei:body>
- <tei:div n="1" type="song">
- <tei:div n="1" type="strophe">
    <l n="1">O Tannenbaum, O Tannenbaum, </l>
    <l n="2">Wie treu sind deine Blätter. </l>
    <l n="3">Du grünst nicht nur zur Sommerzeit, </l>
    <l n="4">Nein auch im Winter wenn es schneit. </l>
    <l n="5">O Tannenbaum, O Tannenbaum, </l>
    <l n="6">Wie grün sind deine Blätter! </l>
</tei:div>
- <tei:div n="2" type="strophe">
    <l n="1">O Tannenbaum, O Tannenbaum, </l>
    <l n="2">Du kannst mir sehr gefallen! </l>
    <l n="3">Wie oft hat schon zur Winterszeit </l>
    <l n="4">Ein Baum von dir mich hoch erfreut! </l>
    <l n="5">O Tannenbaum, O Tannenbaum, </l>
    <l n="6">Du kannst mir sehr gefallen! </l>
</tei:div>
- <tei:div n="3" type="strophe">
    <l n="1">O Tannenbaum, O Tannenbaum, </l>
    <l n="2">Dein Kleid will mich was lehren: </l>
    <l n="3">Die Hoffnung und Beständigkeit </l>
    <l n="4">Gibt Mut und Kraft zu jeder Zeit! </l>
    <l n="5">O Tannenbaum, O Tannenbaum, </l>
    <l n="6">Dein Kleid will mich was lehren. </l>
</tei:div>
</tei:div>
<tei:body>
</tei:text>
</tei:TEI>
</passage>

- <passage>
- <div1 n="1" type="song">
- <div2 n="1" type="strophe">
    <div3 n="1" type="line">O Christmas tree, O Christmas tree ! </div3>
    <div3 n="2" type="line">How are thy leaves so verdant ! </div3>
    <div3 n="3" type="line">Not only in the summertime, </div3>
    <div3 n="4" type="line">But even in winter is thy prime. </div3>
    <div3 n="5" type="line">O Christmas tree, O Christmas tree, </div3>
    <div3 n="6" type="line">How are thy leaves so verdant ! </div3>
</div2>
- <div2 n="2" type="strophe">
    <div3 n="1" type="line">O Christmas tree, O Christmas tree, </div3>
    <div3 n="2" type="line">Much pleasure doth thou bring me ! </div3>
    <div3 n="3" type="line">For every year the Christmas tree, </div3>
    <div3 n="4" type="line">Brings to us all both joy and glee. </div3>
    <div3 n="5" type="line">O Christmas tree, O Christmas tree, </div3>
    <div3 n="6" type="line">Much pleasure doth thou bring me ! </div3>
</div2>
- <div2 n="3" type="strophe">
    <div3 n="1" type="line">O Christmas tree, O Christmas tree, </div3>
    <div3 n="2" type="line">Thy candles shine out brightly ! </div3>
    <div3 n="3" type="line">Each bough doth hold its tiny light, </div3>
    <div3 n="4" type="line">That makes each toy to sparkle bright. </div3>
    <div3 n="5" type="line">O Christmas tree, O Christmas tree, </div3>
    <div3 n="6" type="line">Thy candles shine out brightly ! </div3>
</div2>
</div1>
</passage>
```

Configuration Parameter

Dynamically provide parameters for post processing

Can be combine using the underscore “_”

Parameter	Description
divs	Structures the document with numbered divs (div1 type=chapter, div2 type=line,...)(false if @-notation is used, div=true overwrites epidoc=true)
epidoc	Structures the document in EpiDoc-Format (false if @-notation is used, div=true overwrites epidoc=true)
escapePassage (plus epidoc=true)	XML-Escape the passage (true if @-notation is used)
smallinventory	GetCapabilities without metainformation (more efficient for crawling)
xmlformatting	Should the XML be formatted to be easier to read or not to be more efficient? Your browser formats it anyway, so you might only see a difference in the website's sourcecode.
deletexml	Should the xml tags inside the text (e.g. meta information) be deleted to make it more readable?

Configuration Div View

```
<passage>
O Christmas tree, O Christmas tree ! How are thy leaves so verdant ! Not only in the summertime, But even in winter is thy prime. O Christmas tree, O Christmas tree, How are thy leaves so verdant ! O Christm
tree, O Christmas tree, Much pleasure doth thou bring me ! For every year the Christmas tree, Brings to us all both joy and glee. O Christmas tree, O Christmas tree, Much pleasure doth thou bring me ! O
Christmas tree, O Christmas tree, Thy candles shine out brightly ! Each bough doth hold its tiny light, That makes each toy to sparkle bright. O Christmas tree, O Christmas tree, Thy candles shine out brightly !
</passage>
```

configuration=divs=true

```
-<passage>
- <div1 n="1" type="song">
- <div2 n="1" type="strophe">
  <div3 n="1" type="line">O Christmas tree, O Christmas tree ! </div3>
  <div3 n="2" type="line">How are thy leaves so verdant ! </div3>
  <div3 n="3" type="line">Not only in the summertime, </div3>
  <div3 n="4" type="line">But even in winter is thy prime. </div3>
  <div3 n="5" type="line">O Christmas tree, O Christmas tree, </div3>
  <div3 n="6" type="line">How are thy leaves so verdant ! </div3>
</div2>
- <div2 n="2" type="strophe">
  <div3 n="1" type="line">O Christmas tree, O Christmas tree, </div3>
  <div3 n="2" type="line">Much pleasure doth thou bring me ! </div3>
  <div3 n="3" type="line">For every year the Christmas tree, </div3>
  <div3 n="4" type="line">Brings to us all both joy and glee. </div3>
  <div3 n="5" type="line">O Christmas tree, O Christmas tree, </div3>
  <div3 n="6" type="line">Much pleasure doth thou bring me ! </div3>
</div2>
- <div2 n="3" type="strophe">
  <div3 n="1" type="line">O Christmas tree, O Christmas tree, </div3>
  <div3 n="2" type="line">Thy candles shine out brightly ! </div3>
  <div3 n="3" type="line">Each bough doth hold its tiny light, </div3>
  <div3 n="4" type="line">That makes each toy to sparkle bright. </div3>
  <div3 n="5" type="line">O Christmas tree, O Christmas tree, </div3>
  <div3 n="6" type="line">Thy candles shine out brightly ! </div3>
</div2>
</div1>
</passage>
```

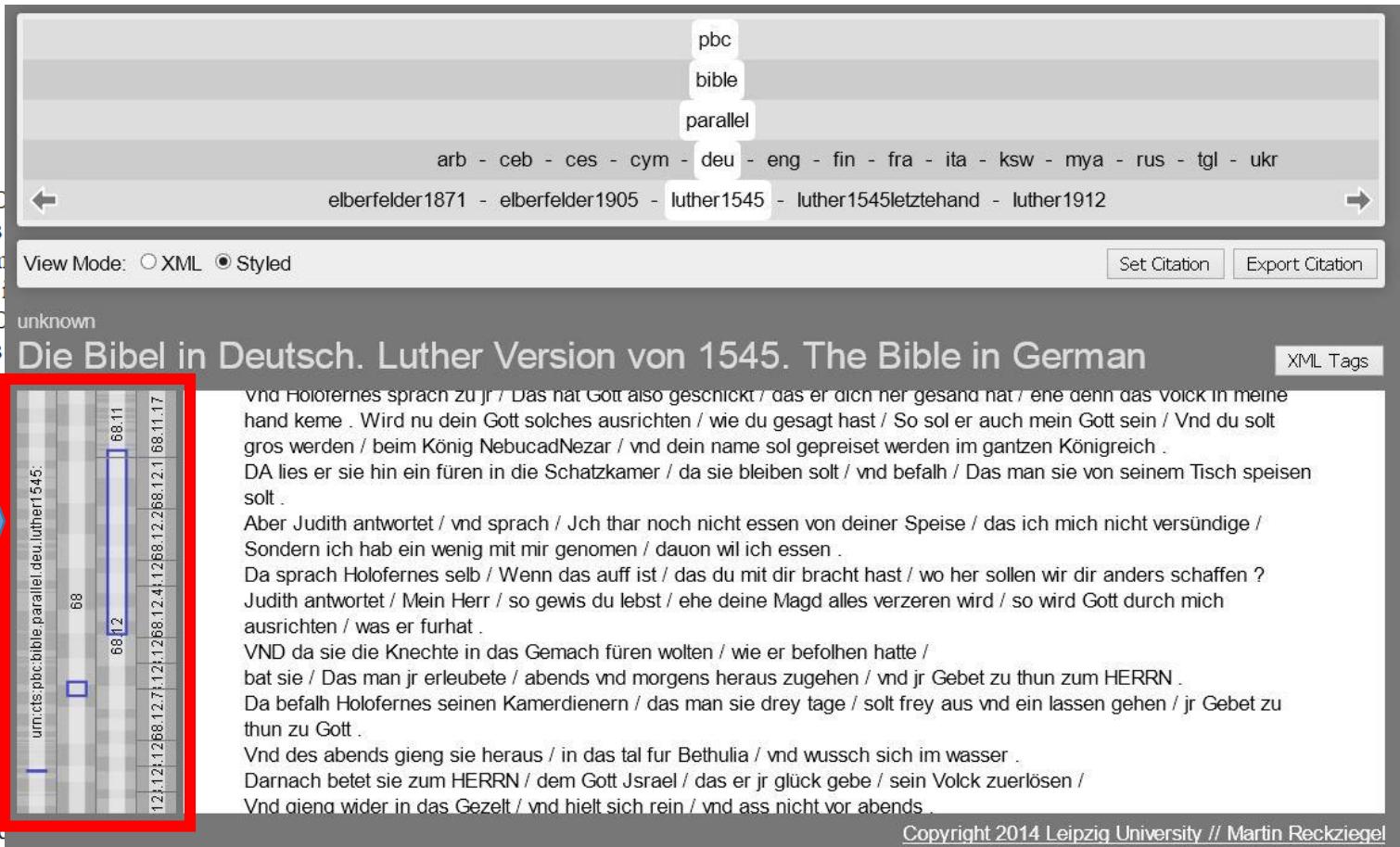
Configuration Parameter Example Request

`http://ctstest.informatik.uni-leipzig.de/perseus/cts/?request=GetPassage&urn=urn:cts:greekLit:tlg0007.tlg001.perseus_eng2:4.1&configuration=deletexml=true_escapePassage=false`

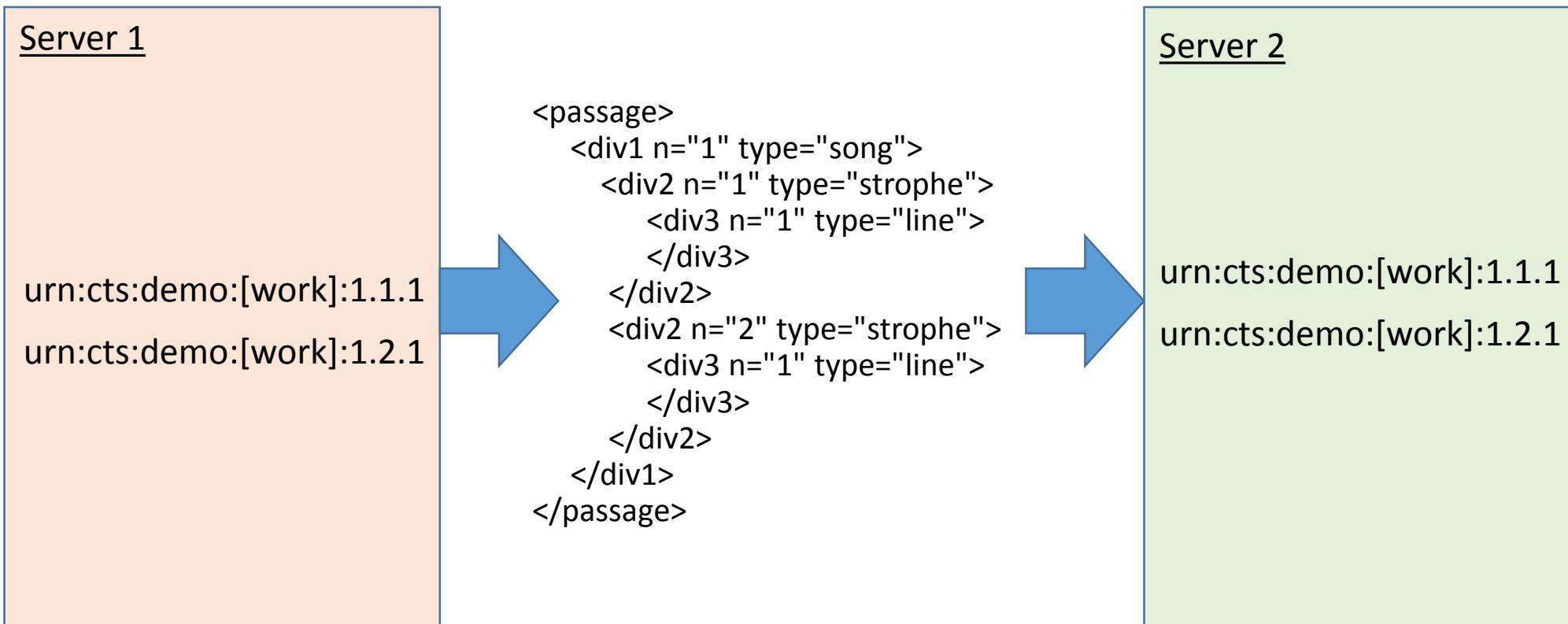
Configuration Parameter that combines deletexml with escapePassage

Configuration Div View

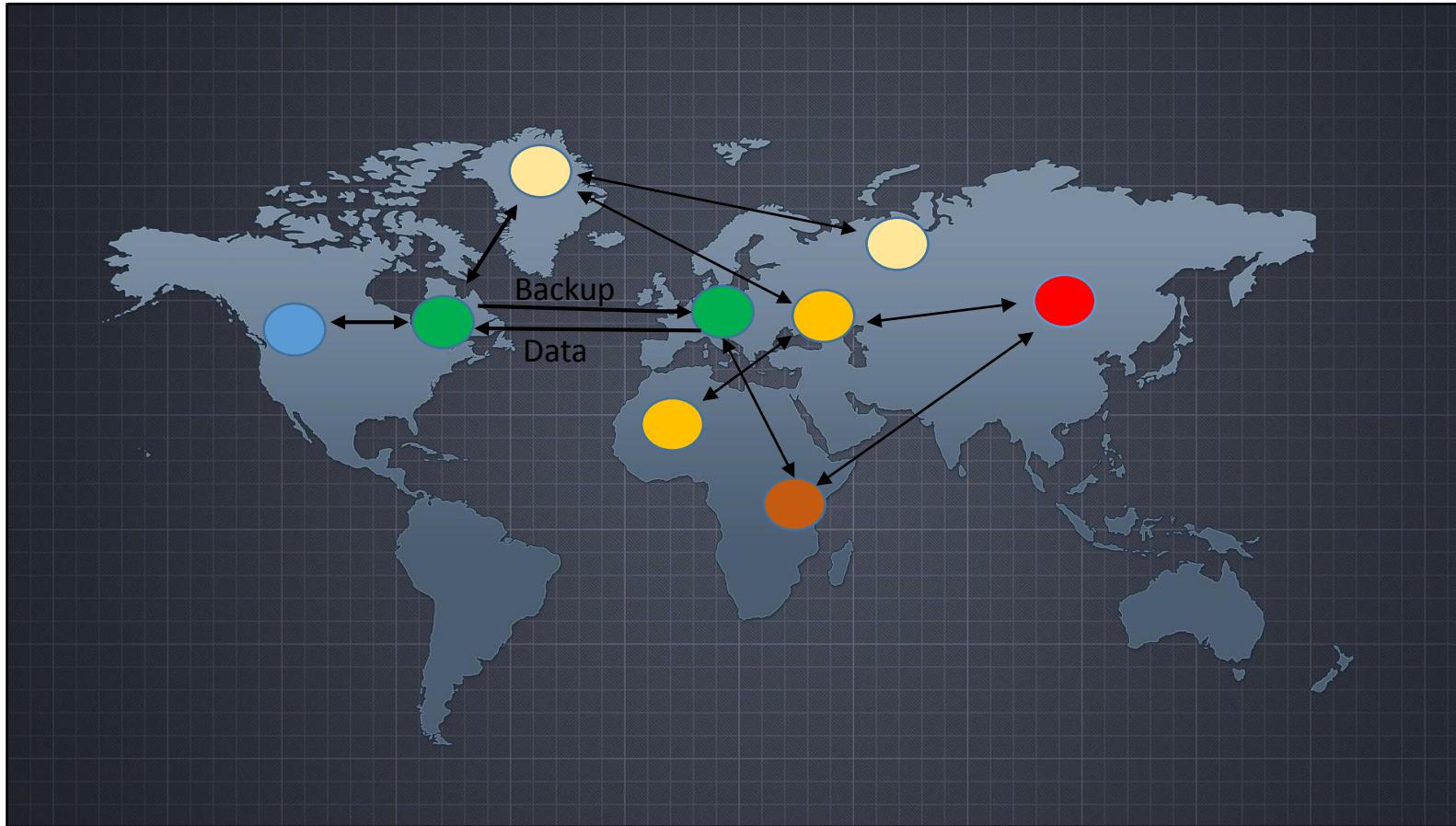
```
- <passage>
  - <div1 n="1" type="song">
    - <div2 n="1" type="strophe">
      <div3 n="1" type="line">O Christmas tree, O
      <div3 n="2" type="line">How are thy leaves
      <div3 n="3" type="line">Not only in the sun
      <div3 n="4" type="line">But even in winter
      <div3 n="5" type="line">O Christmas tree, O
      <div3 n="6" type="line">How are thy leaves
    </div2>
  - <div2 n="2" type="strophe">
    <div3 n="1" type="line">O
    <div3 n="2" type="line">M
    <div3 n="3" type="line">I
    <div3 n="4" type="line">R
    <div3 n="5" type="line">C
    <div3 n="6" type="line">O
  </div2>
  - <div2 n="3" type="strophe">
    <div3 n="1" type="line">O Christmas tree, O
    <div3 n="2" type="line">Thy candle doth shine
    <div3 n="3" type="line">Each bough doth hold
    <div3 n="4" type="line">That makes each to
    <div3 n="5" type="line">O Christmas tree, O
    <div3 n="6" type="line">Thy candles shine
  </div2>
</div1>
</passage>
```



CTS Cloning



CTS Cloning



<http://hdw.eweb4.com/out/1369880.html>

Advanced Request Functions

Plain text based output instead of XML

Goal: More performance & tool developer friendly requests

Use different URL branch to assure compatibility

`http://cts.informatik.uni-leipzig.de/pbc/cts/...`

`http://cts.informatik.uni-leipzig.de/pbc/plain/...`

Not covered by CTS specifications

Advanced Request Functions

- Meta Information

/urns

/urncount

/urncount?urn=[urn]

/doccount

/authors

/titles

/titles?pagesize=3&page=2

/titlesandurns

/titlesandurns?pagesize=3&page=2

/editions

/editions?pagesize=3&page=2

Advanced Request Functions

- Document Structure

/urnpartsonlevel?urn=[urn]&level=2

/urnstypes?urn=[urn]

/depthtypes?urn=[urn]

/compactdepthtypes?urn=[urn]

/texts?urn=[urn]

/urnstexts?urn=[urn]

/urnstypestext?urn=[urn]

/urnstypestextlength?urn=[urn]

/depthstypestextlength?urn=[urn]

Advanced Request Functions

/urnstypestext?urn=[urn]

```
urn:cts:pbc:bible.parallel.eng.kingjames: edwrap null
urn:cts:pbc:bible.parallel.eng.kingjames:1 book null
urn:cts:pbc:bible.parallel.eng.kingjames:1.1 chapter null
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.1 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.2 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.3 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.4 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.5 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.6 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.7 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.8 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.9 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.10 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.11 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.12 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.13 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.14 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.15 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.16 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.17 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.18 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.19 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.20 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.21 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.22 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.1.23 sentence
urn:cts:pbc:bible.parallel.eng.kingjames:1.21 sentence
```

/depthstypestextlength?urn=[urn]

0	edwrap	-
1	book	-
2	chapter	-
3	sentence	56
3	sentence	147
3	sentence	58
3	sentence	89
3	sentence	119
3	sentence	116
3	sentence	148
3	sentence	92
3	sentence	131
3	sentence	125
3	sentence	184
3	sentence	181
3	sentence	53
3	sentence	174
3	sentence	105
3	sentence	134
3	sentence	79
3	sentence	121

Advanced Request Functions

- <http://cts.informatik.uni-leipzig.de/pbc/plain/depthtypestextlength?urn=urn:cts:pbc:bible.parallel.eng.kingjames>:

Alignment

Find comparable text parts in different documents

-> Comparison of documents/editions

Structural Alignment

-> Input for translation memory tools

:1.4.21	नामानेव भाईयोर नाम दिल खूब। नामा बींगा ओ बांचि बाजार खूब तालार सकालेर पुल्कुलुव.	Et le nom de son frère était Jubal ; il fut père de tous ceux qui jouent de la harpe et du chalumeau .	And his brother's name was Jubal ; he was the father of those who handle the harp and pipe .	Und sein Bruder hieß Jubal ; von dem sind hergekommen die Geiger und Pfeifer .
:1.4.22	सिलार गर्ते ठुबल-कमिनेव जम्म हल। (ज्ञात आर नामार सब तारमार मस्तुपांचि तेवी कम्मा दिल तार काज।) ठुबल-कमिनेव बोलेव नाम दिल नामा।	Et Tsilla , elle aussi , enfanta Tubal-Cain , qui forgeait toute sorte d'instruments tranchants d'airain et de fer , et la sœur de Tubal-Cain fut Naama .	And Zillah , she also bore Tubal-Cain , the forger of every kind of tool of brass and iron . And the sister of Tubal-Cain was Naamah .	Die Zilla aber gebaer auch , nämlich den Thubalkain , den Meister in allerlei Erz-und Eisenwerk . Und die Schwester des Thubalkain war Naema .
:1.4.23	एक्लिच लेमेक ताजल दुई श्वेते वाल , "आदा आर निला , डोमार आमार कना प्रिंगा , मेमाकेर श्वास , अमार उधार काल नाठ।" (ये श्वेत कामार जामा कराए , अदा , ये श्वेत कामार जामा राहत कुलेक , अदा ताके खूब कराए)	Et Lémech dit à ses femmes : Ada et Tsilla , écoutez ma voix ; femmes de Lémech , prétez l'oreille à ma parole : Oui ! J'ai tué un homme pour ma blessure , et un jeune homme pour ma meurtrissure .	And Lemech said to his wives : Adah and Zillah , hear my voice , Ye wives of Lemech , listen to my speech . For I have slain a man for my wound , and a youth for my bruise	Und Lamech sprach zu seinen Weibern Ada und Zilla : Ihr Weiber Lamechs , hört meine Rede und merkt , was ich sage : Ich habe einen Mann erschlagen für meine Wunde und einen Jungling für meine Beule ;
:1.4.24	करिल्क खूब करवार प्रहिलाख थानि साकेतण हल , तावे लेमेकक खूब करवार प्रहिलाख हले साकात्तर ऊर ! "	Car si Cain est vengé sept fois , Lémech le sera soixante-dix-sept fois .	If Cain shall be avenged seven-fold ; Lemech seventy and seven fold .	Kain soll siebenmal gerächt werden , aber Lamech siebenundsiebzigmal .
:1.4.25	गर्ते आदम आवार तार त्रित कल्ले देले एवं तार श्वीते फ्रेटो देले हल। तार श्वी देल नाम झार्लेन लेख। इया वर्गान , "करिन रेलेक-खूब कराए वाल तिल्लर बोलेव आकाय आमाके आर एकाटि सहाल दिलेन।"	Et Adam connaît encore sa femme , et elle enfanta un fils et l'appela Seth (remplaçant) ; car Dieu , dit-elle , m'a donné un autre fils au lieu d' Abel , que Cain a tué .	And Adam knew his wife again , and she bore a son , and called his name Seth : ... For God has appointed me another seed instead of Abel , because Cain has slain him	Adam erkannte abermals sein Weib , und sie gebaer einen Sohn , den hieß sie Seth ; denn Gott hat mir , sprach sie , einen andern Samen gesetzt für Abel , den Kain erwürgt hat .
:1.4.26	गर्ते श्वेते एकाटि घटे हल। तिलि तार नाम झार्लेन इनालौ। (सेहि समय थेके लोकेवा नमाप्राप्तुके तार नोरा नमान लिले शुक कराए)	Et un fils naquit aussi à Seth , et il l'appela Énoch . Alors on commence à invoquer le nom de l'Éternel .	And to Seth , to him also was born a son ; and he called his name Enosh . Then people began to call on the name of Jehovah .	Und Seth zeugte auch einen Sohn und hieß ihn Enos . Zu der Zeit fing man an , zu predigen von des HERRN Namen .
:1.5				Dies ist das Buch von des Menschen Geschlecht . Da Gott den Menschen schuf , machte er ihn nach dem Bilde Gottes ;
:1.5.1	ऐह इह आदमेव व्यश्वेव कथा। मानुस शुष्टि करवार समय ईश्वर ताके तार निलेव माते शुष्टि कराएँ;	Voici le livre de la postérité d' Adam . Au jour où Dieu crée l'Homme , il le fit à la ressemblance de Dieu .	This is the book of Adam's generations . In the day that God created man , in the likeness of God made he him .	
:1.5.2	शुष्टि करालेन मुख्य एवं श्वेतोक कर्त्त्र एवं तालेव अलीर्वाप करालेन। शुष्टिन समय तिलि ताराक नाम दिलेन शुष्टुः।	Il les créa mâle et femelle , et il les bénit , et leur donna le nom d' Homme , au jour qu'ils furent créés .	Male and female created he them ; and blessed them , and called their name Adam , in the day when they were created .	
:1.5.3	एक्लो तिल अवर व्याम आदमेव एकाटि घटेव जम्म हल। (घोलेटि वालिने एवं तिलेव ताराइ माते व्यामिले।) आदम तार नाम दिलेन (सेहि)	Or , Adam vécut cent trente ans , et engendra un fils à sa ressemblance , selon son image , et il lui donna le nom de Seth .	And Adam lived a hundred and thirty years , and begot [a son] in his likeness , after his image , and called his name Seth .	Und Adam war hundertunddreißig Jahre alt und zeugte einen Sohn , der seinem Bild ähnlich war und hieß ihn Seth
:1.5.4	प्रोटेव आवार पर आदम आर्कृ आदमेव बदर बोटे दिलेन। एवं माध्ये तार आर्कृ बोलेमेव आवार व्यामिले।	Et les jours d' Adam , après qu'il eut engendré Seth , furent de huit cents ans , et il engendra des fils et des filles .	And the days of Adam after he had begotten Seth were eight hundred years , and he begot sons and daughters .	und lebte damach achthundert Jahre und zeugte Söhne und Töchter ;
:1.5.5	मोटौ न 'लो त्रिन अवर पर आदम आर्कृ आदमेव बदर बोटे दिलेन।	Tout le temps qu' Adam vécut , fut donc de neuf cent trente ans ; puis il mourut .	And all the days of Adam that he lived were nine hundred and thirty years ; and he died .	daß sein ganzes Alter ward neunhundertunddreißig Jahre , und starb .
:1.5.6	प्रोटेव एक्लो गाँठ बधार व्याम तार घेले इनालौ जम्म हल।	Seth vécut cent cinq ans , et engendra Énoch .	And Seth lived a hundred and five years , and begot Enos	Seth war hundertundfünf Jahre alt und zeugte Enos

Alignment

Naive: Sentence Tokenization

- > Sentences may be merged or splitted

Statistical Methods

- Named Entity Recognition

- Precalculation, Cross Comparison

- > DTA: 3.5 GB * 5136 documents > 17 TB

Error probability accumulates

Alignment via CTS

Alignment of CTS URNs

urn:cts:demo:shakespeare.sonnets.en.1:3.2

urn:cts:demo:shakespeare.sonnets.de.1:3.2

urn:cts:demo:shakespeare.sonnets.fr.1:3.2

-> Error probability goes back to default when
text unit is „switched“

-> Fast, Realtime

1.4.21	नामलेप भौतिकर नाम चिन तुलन। यामा शीले ३ वर्षी वाजाय बुलन डासर सहित दुर्लभम्।	Et le nom de son frère était Jubal ; il fut père de tous ceux qui jouent de la harpe et du chalumeau.	And his brother's name was Jubal : he was the father of those who handle the harp and pipe.	Und sein Bruder hieß Jubal ; von dem sind hergekommen die Geiger und Pfeifer.
1.4.22	शिलार गद्द तुलन-करिनार जाव हा। त्रोड अत गोपाल सव राहन्यत द्वृष्टपूर्वि तेली काळ पिण ताव तावा। तुलन-करिनार वोलान नाम चिन वरमा।	Et Tsilla , elle aussi , enfanta Tubal-Cain , qui forgeait toute sorte d'instruments tranchants d'alain et de fer , et la sœur de Tubal-Cain fut Naama .	And Zillah , she also bore Tubal-Cain , the forger of every kind of tool , of brass and iron . And the sister of Tubal-Cain was Naamah .	Die Zilla aber gab auch , nämlich den Thubalkain , den Meister in allerlei Erz-und Eisenwerk . Und die Schwester des Thubalkain war Naema .
1.4.23	एकान्तिन गामक भालू द्वृष्ट श्रीके नामन। - आदा अत आदा। (तामार आमार कामा प्रोल , लेमेच कृत्तिला , आमार कमार कामा)। वे लेमेच आमारक काम करावः , अर्थात् , दे युक्त आमार पात्र वाह कुलकान , आमि उक्ते धूम कामावः।	Et Lémech dit à ses femmes : Ada et Tsilla , écoutez ma voix ; femmes de Lémech , prétez l'oreille à ma parole : Oui ! j' tue un homme pour ma blessure , et un jeune homme pour ma meurtrissure	And Lemech said to his wives . Adah and Zillah , hear my voice . Ye wives of Lemech , listen to my speech . For I have slain a man for my wound , and a youth for my bruise meurtrissure.	Und Lamech sprach zu seinen Weibern Ada und Zilla : Ihr Weiber Lamechs , hört meine Rede und merkt , was ich sage : Ich habe einen Mann erschlagen für meine Wunde und einen Jungen für meine Beule .
1.4.24	कृष्णाकृष्ण धूम करावर आदिशाख वानि वाक्तव्य रुप , भलू लेमेचक धूम करावर प्राक्तिकृष्ण धूम करावर वानि वाक्तव्य रुप।	Car si Cain est venge sept fois , Lémec sera soixante-dix-sept fois.	If Cain shall be avenged seven-fold , Lemech seventy and seven fold.	Kain soll siebenmal gerächt werden , aber Lamech siebenundsiebzigmal.
1.4.25	पत्र आदम अवार भालू श्रीके कालू (गोलन एवं भालू श्रीके एकान्ति श्रीके धूम।) जाव श्री उत्र आदा वालान्माने (लेम)। वहा वरमान , “अर्थात् श्रेष्ठकृष्ण करावर वलू श्रीकृष्ण वेलावर आमार आमार आमार एकान्ति श्रावन लिपिना।”	Et Adam connaît encore sa femme , et elle enfanta un fils et l'appela Seth (remplaçant) , car Dieu , dit-elle , m'a donné un autre fils au lieu d' Abel , que Cain a tué .	And Adam knew his wife again , and she bore a son , and called him name Seth : “For God has appointed me another seed instead of Abel , because Cain has slain him in his place .”	Adam erkannte abermals sein Weib , und sie gebar einen Sohn , den hiess sie Seth ; denn Gott hat mir , sprach sie , einen andern Samen gesetzt für Abel , den Kain erungert hat .
1.4.26	पत्र आदम एकान्ति धूम है। उपरि ताव माम तालन वेला। (मैंहि सम्म धूमेर लेमेच)। सम्मानकृतूक तालू। वानि वाक्तव्य विणे धूम करावर।	Et un fils naquit aussi à Seth , et il l'appela Enosh . Alors on commence à invoquer le nom de l'Étemel .	And to Seth , to him also was born a son ; and he called his name Enosh . Then people began to call on the name of Jehovah .	Und Seth zeugte auch einen Sohn und hieß ihn Enos . Zu der Zeit fing man an , zu predigen von des HERRN Namen .
1.5				
1.5.1	ऐ हम आदमक लालून कथा। मध्ये शुटि करावर सम्म देवर भालू शिलोर मृत करू शुटि करावर।	Voici le livre de la postérité d' Adam . Au jour où Dieu créa l'homme , il le fit à la ressemblance de Dieu	This is the book of Adam's generations . In the day that God created man , in the likeness of God made he him .	Dies ist das Buch von des Menschen Geschlecht . Da Gott den Menschen schuf , machte er ihn nach dem Bilde Gottes .
1.5.2	सूषि करावर सम्म एवं श्रीलेच करू एवं तामा आमार्यान करावर। शुष्टि सम्म तिलू तामा नाम लिपिना शाम्भव।	Il les crée mâle et femelle , et il les bénit , et leur donne le nom d' Homme , au jour qu'ils furent créés .	Male and female created he them ; and blessed them , and called their name Adam , in the day when they were created .	Und Adam war hundertdreißig Jahre alt und zeugte einen Sohn , der seinem Bild ähnlich war und hieß ihn Seth
1.5.3	प्राक्त्य॒ त्रिश वर्षम् वस्यते अस्मिन् एकान्ति वेलावर अमूर रुप। अर्थात् वास्तवि त्रिश विभूति तालूर धूम वरमानि। आदा भालू नाम लिपिना (लेम)।	Or , Adam vécut cent trente ans , et engendra un fils à sa ressemblance , selon son image , et il lui donna le nom de Seth .	And Adam lived a hundred and thirty years , and begot [a son] in his likeness , after his image , and called his name Seth .	Und Adam war hundertdreißig Jahre alt und zeugte einen Sohn , der seinem Bild ähnlich war und hieß ihn Seth
1.5.4	प्राप्त्य॒ आदमक पर आमार एकान्ति धूम करावर वानि वाक्तव्य लिपिन। एवं मैंहि तालू तालू औलालयम् लालू धूम वरमानि।	Et les jours d' Adam , après qu'il eut engendré Seth , furent de huit cents ans , et il engendra des fils et des filles .	And the days of Adam after he had begotten Seth were eight hundred years ; and he begot sons and daughters .	und lebte damach achthundert Jahre und zeugte Söhne und Töchter ;
1.5.5	मोटौ न॑ (ग॑ त्रिश वर्षम् वेला एकान्ति धूम करावर पर आमार मारा (गोलन)।	Tout le temps qu' Adam vécut , fut donc de neuf cent trente ans ; puis il mourut .	And all the days of Adam that he lived were nine hundred and thirty years , and he died .	daß sein ganzes Alter ward neunhundertunddreißig Jahre , und starb .
1.5.6	प्राप्त्य॒ एकान्ति धूम करावर वस्यते भालू वेला इवान्शर लालू धूम।	Seth vécut cent cinq ans , et engendra Enosh .	And Seth lived a hundred and five years , and begot Enosh .	Seth war hundertfünf Jahre alt und zeugte Enos

Realtime Alignment Tools for CTS

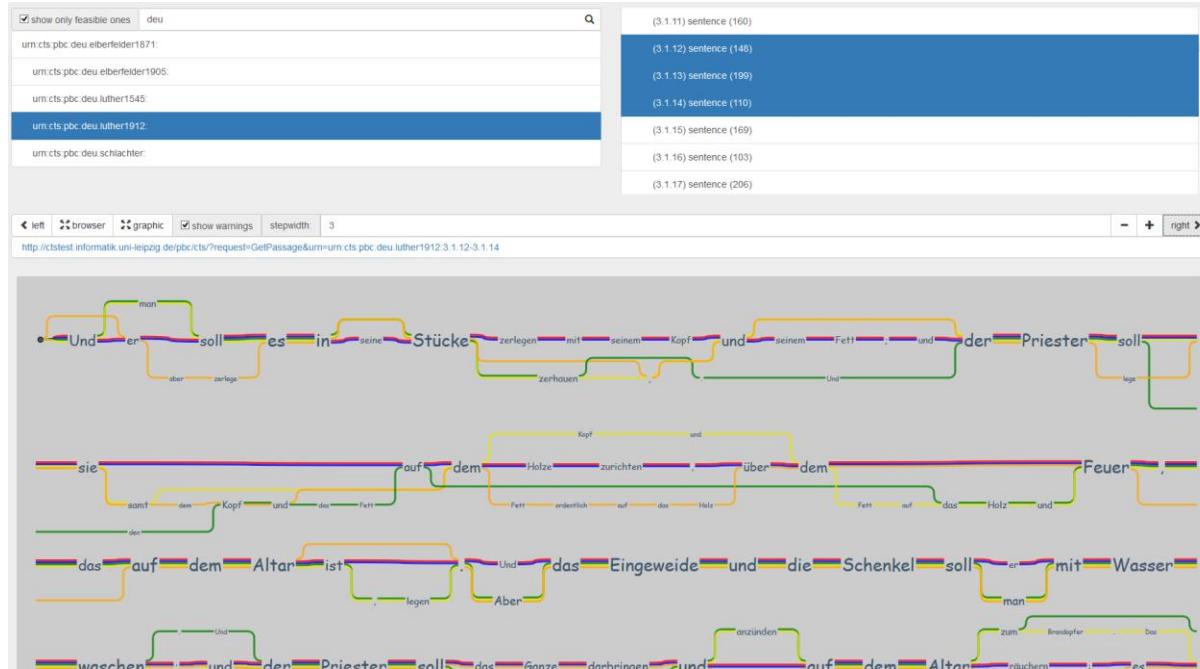
The screenshot shows a web-based application for aligning text passages from different CTS documents. At the top left is a search bar with the query "fr". To its right is a sidebar with a list of URNs and navigation controls (+ add, - rm, up, down). The main area displays a table with five columns, each representing a different language or document version. The columns are labeled: urn:cts:pbc:dan.frederik:, urn:cts:pbc:deu.elberfelder1871:, urn:cts:pbc:eng.darby:, urn:cts:pbc:fra.darby:, and urn:cts:pbc:gur.frafra:. The table rows represent specific text alignments, with each row containing five cells corresponding to the language columns. The first cell contains the section identifier (e.g., :40) and the subsequent cells contain the aligned text passages.

section	urn:cts:pbc:dan.frederik:40	urn:cts:pbc:deu.elberfelder1871:	urn:cts:pbc:eng.darby:	urn:cts:pbc:fra.darby:	urn:cts:pbc:gur.frafra:
:40					
:40.1	Jesu Christi , Davids Søns , Abrahams Søns , Slægtbog .	Buch des Geschlechts Jesu Christi , des Sohnes Davids , des Sohnes Abrahams .	Book of the generation of Jesus Christ , Son of David , Son of Abraham .	Livre de la généalogie de Jésus Christ , fils de David , fils d'Abraham .	Gorp wa dela Yesu Krista tuunhom . A dela Abraham yaanja la David yaanja .
:40.1.2	Abraham avlede Isak ; og Isak avlede Jakob ; og Jakob avlede Juda og hans Brødre .	Abraham zeugte Isaak , Isaak aber zeugte Jakob , Jakob aber zeugte Juda und seine Brüder ;	Abraham begat Isaac ; and Isaac begat Jacob , and Jacob begat Juda and his brethren ;	Abraham engendra Isaac ; et Isaac engendra Jacob ; et Jacob engendra Juda et ses frères ;	Abraham dayoña n daa de Isaaq , te Isaak daya dina Yakob . te Yakob dayoña dina Yuda la a keendoma la a yebene
:40.1.3	og Juda avlede Phares og Zara med Thamar ; og Phares avlede Esrom ; og Esrom avlede Abram ;	Juda aber zeugte Phares und Zara von der Thamar ; Phares aber zeugte Esrom , Esrom aber zeugte Aram .	and Juda begat Phares and Zara of Thamar ; and Phares begat Esrom , and Esrom begat Aram ,	et Juda engendra Pharès et Zara , de Thamar ; et Pharès engendra Esrom ; et Esrom engendra Aram ;	Yuda dayochee daa dela Perez la Zera . Ba ma daa dela Tamar . Perez daya dela Hezron , te Hezron daya dina Aram ,
:40.1.4	og Abram avlede Aminadab ; og Aminadab avlede Naasson ; og Naasson avlede Salmon ;	Aram aber zeugte Aminadab , Aminadab aber zeugte Nahasson , Nahasson aber zeugte Salmon ,	and Aram begat Aminadab , and Aminadab begat Naasson , and Naasson begat Salmon .	et Aram engendra Aminadab ; et Aminadab engendra Naasson ; et Naasson engendra Salmon ;	te Aram daya dina Aminadab , te Aminadab daya dina Nason , te Nason daya dina Salmon ,
:40.1.5	Salmon avlede Booz med Rachab ; og Booz avlede Obed med Ruth ; og Obed avlede Jesse ;	Salmon aber zeugte Boas von der Rahab ; Boas aber zeugte Obed von der Ruth ; Obed aber zeugte Jesse .	and Salmon begat Booz of Rachab ; and Booz begat Obed of Ruth , and Obed begat Jesse ,	et Salmon engendra Booz , de Rachab ; et Booz engendra Obed , de Ruth ,	te Salmon daya dina Boaz . A ma dela Arahab . Boaz daya n de Obed . Te a ma dina Arut . Obed daya n de Yesse ,
:40.1.6	og Jessai avlede Kong David ; og Kong David avlede Salomon med Urias' Hustru ;	Jesse aber zeugte David , den König . David aber zeugte Salomon von der , die Urias Weib gewesen ;	and Jesse begat David the king . And David begat Solomon , of her [that had been the wife] of Urias ;	et Obed engendra Jessé ; et Jessé engendra David le roi ; et David le roi engendra Salomon , de celle qui avait été femme d'Urié ;	te Yesse dayoña dina Naba David la . David dayoña n de Solomon se'lem te a ma dela Urià poga .
:40.1.7	og Salomon avlede Roboam ; og Roboam avlede Abia ; og Abia avlede Asa ;	Salomon aber zeugte Roboam , Roboam aber zeugte Abia , Abia aber zeugte Asa ,	and Solomon begat Roboam , and Roboam begat Abia , and Abia begat Asa ,	et Salomon engendra Roboam ; et Roboam engendra Abia ; et Abia engendra Asa ;	Solomon daya n daa de Arehoboam , te Arehoboam daya dina Abiya , te Abiya daya dina Asa ,
:40.1.8	og Asa avlede Josaphat ; og Josaphat avlede	Asa aber zeugte Josaphat , Josaphat aber zeugte Joram .	and Asa begat Josaphat , and Josaphat begat	et Asa engendra Josaphat ; et Josaphat	te Asa daya dina Yehosaphat te Yehosaphat daya

http://cts.informatik.uni-leipzig.de/cts_admin_tools/browser/?urn=urn:cts:pbc:bible.parallel.deu.luther1545:&ctsURL=../../pbc/cts/

Align text passages across different languages

Realtime Alignment Tools for CTS



http://cts.informatik.uni-leipzig.de/cts_admin_tools/alignbrowser/?ctsURL=../../pbc&urn=urn:cts:pbc:bible.parallel.deu.luther1545:3.2.11&stepwidth=1

Uses TRAViz
S. Jaenicke, M. Buechler and G. Scheuermann (2014). Improving the Layout for Text Variant Graphs.
VisLR workshop at LREC conference 2014, ReykjavÃk, Iceland.

Align text passage variations in one language

Contact

Jochen Tiepmar

E-Mail: jtiepmar@informatik.uni-leipzig.de

Scalable Data Solutions (ScaDS) Leipzig
Universität Leipzig
Ritterstraße 9-13
04109 Leipzig



Automatische Sprachverarbeitung



UNIVERSITÄT LEIPZIG