

2. Digitalisierungskonferenz, 4.12.2021

Digitale Initiativen des Sorbischen Instituts

und Hinweise zur Weiterentwicklung des
„Digitalkonzepts“

Dr. Hauke Bartels

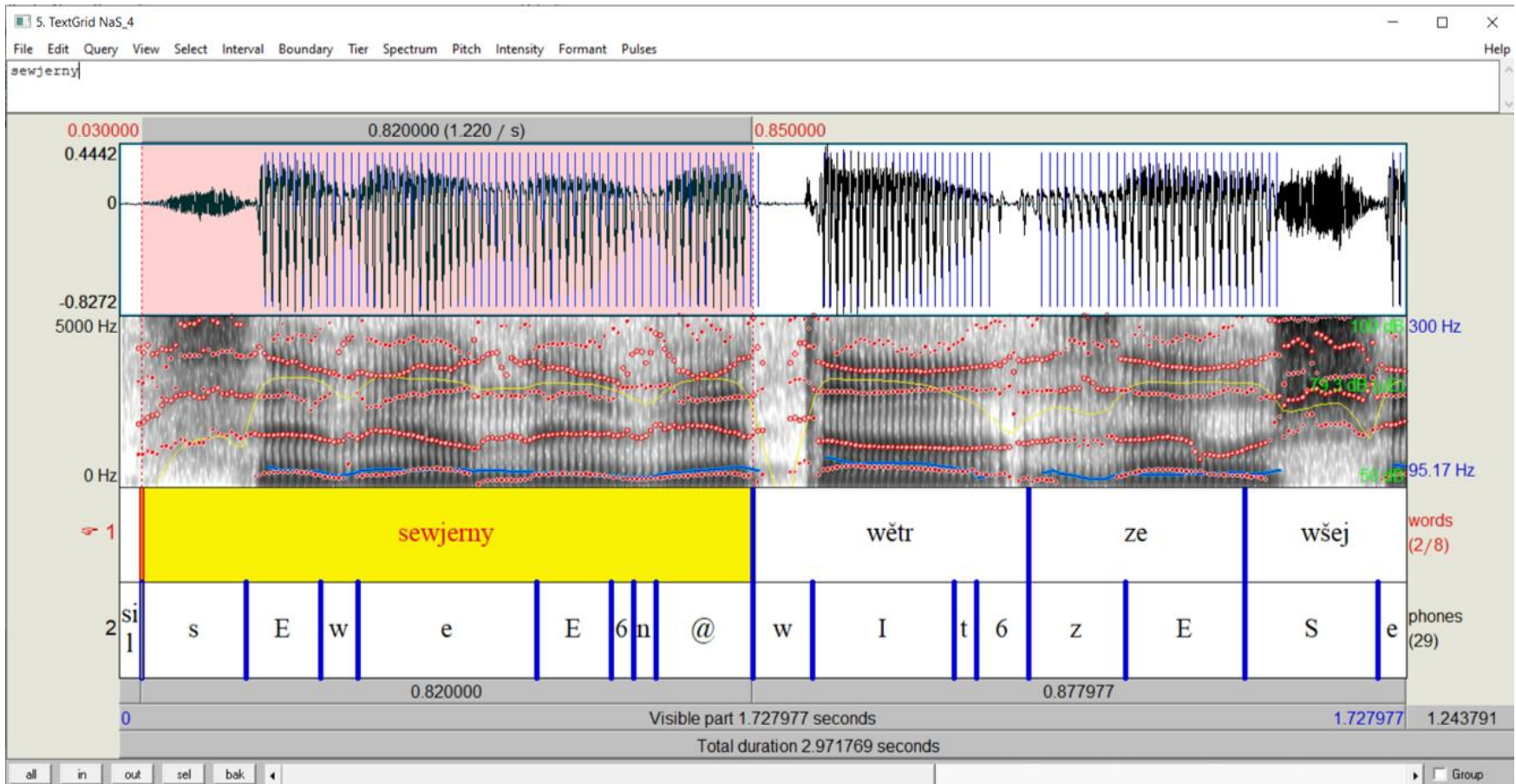


**Serbski Sorbisches
institut Institut**

Text-to-speech: Vorlesefunktion



Serbski Sorbisches
institut Institut



→ Präsentation auf dem Institutstag am 10.11.2021 in Hoyerswerda – zugänglich über Youtube (SI)



Breite der SI-Aktivitäten

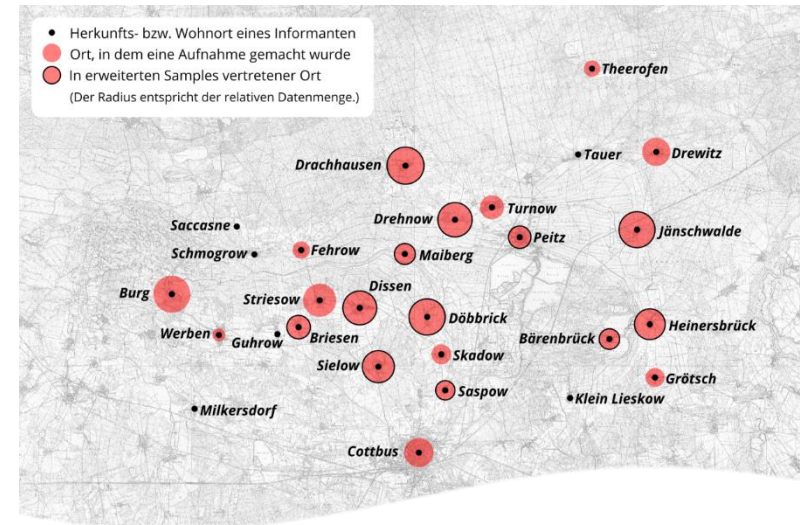
1. (Weiter)**Entwicklung von Technologien** oder spez. techn. Systeme zur Digitalisierung: TTS, OCR, ...
2. **Metadaten:** Dig. Sorb. Bibliographie, Katalog des SKA, Dig. Fotoarchiv, ... → GND
3. **Dig. Dokumentation gesprochener Sprache** (Tonaufnahmen, Metadaten, ...)
- <https://www.dolnoserbški.de/dobes/>
4. **Dig. Präsentation “traditioneller” Werke:** Wörterbücher, Sorb. Kulturlexikon, ...
5. **Entwicklung neuer Sprachressourcen** in dig. Form:
- Autom. Rechtschreibkontrolle, autom. Silbentrennung, Informationsservice zu Eigennamen oder zur Aussprache/Orthoepie, Neue Lexik (i. Vorb.), ...
6. **Textdigitalisierung:** Images und Volltexte → Korpustexte → Dig. Bibliothek
7. **Digitalisierung** (Inventarisierung und Bereitstellung) verschiedener Objekte des sorbischen **Kulturerbes** = sehr großes Handlungsfeld → ...

Digitalisierung = neue Qualität

100 Std. Aufnahmen/Transkription/Übersetzung



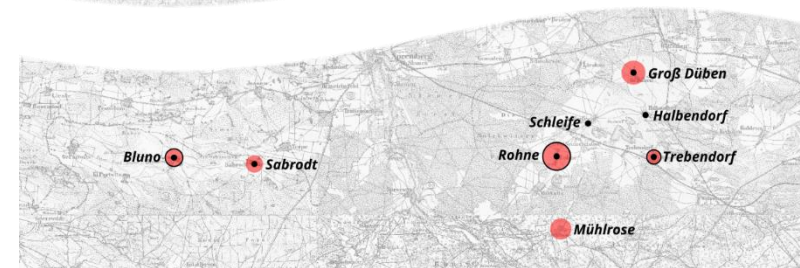
Kartenbasierter Zugang:



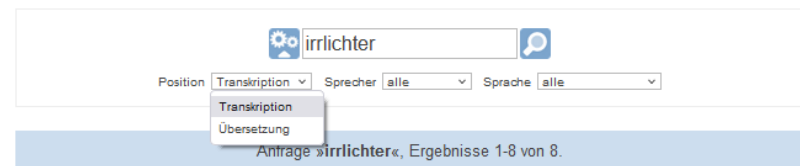
Burg

Thematischer Zugang:

- 2011-12-17:
 1. Hauptthemen: Kindheit, Familie, Sprache, Arbeit, Mythen um Kindern Angst einzujagen, Südamerika, Dorf Fehrow, Kontakt mit anderen Dörfern. (22:31)
 2. Hauptthemen: Familie, Zuhause, Altenheim, Fehrow, Tracht, Kirche, Handarbeit, Weihnachten, Sprache, Aberglauben, Räuchern von Fleisch, Feuer. (33:08)
 3. Gespräch über Bekannte, ein Lied, Aberglauben, Weihnachten, Kochen, Flachs, Backen. (23:20)
- 2011-12-18:
 1. Gespräch zu Familie, Sorbischsprechern, Landwirtschaft, ein Geschenk, Tod des Ehemanns, Wetter, altes Gebetsbuch, Lesen, Sprache. (12:02)
 2. Gespräch und Scherzen. (0:18)
 3. Hauptthemen: Tod, Jugendanedkten, Familie, Tiere, Essen, Alter, Dorf Saccasne, Landwirtschaft, DDR-Zeit, Arbeit, LPG, Schweine, Tabak, Rüben, persönliche Erinnerungen. (29:54)
 4. Hauptthemen: Anekdoten über Russen, Milch, Fahrrad, Essen, Fisch, Weihnachten, Gesellschaft, Tracht. (11:16)
 5. Hauptthemen: Hof, Familie, Jugend, Alkohol, Krankheit, Nachbarschaftshilfe, Umgang mit Vieh, Anekdote über Verhexung von Vieh. (21:16)



Zugang über Transkription/Übersetzung:



Digitalisierung = neue Qualität



Serbski Sorbisches
institut Institut

```
<textClass>
  <classCode scheme="http://www.deutschestextarchiv.de/doku/klassifikation#dws">Zeitung</classCode>
</textClass>
</profileDesc>
<revisionDesc status="structure-4">
  <change when="2019-11-14" who="MSI" status="structure-2">initial TEI-P5 encoding</change>
  <change when="2019-11-22" who="JSA" status="structure-4">TEI-P5 encoding</change>
</revisionDesc>
</teiHeader>
<text xml:lang="dsb">
  <front>
    <pb n="1"/>
    <titlePage>
      <docTitle>
        <titlePart type="main">Nowy CASNIK</titlePart>
        <titlePart type="desc">TYŽENIK ZA SERBSKI LUD</titlePart>
      </docTitle>
      <docEdition>33/Lětnik 70</docEdition>
      <docDate>15. awgust 2019</docDate>
    </titlePage>
    <p>0,65 €</p>
    <p>ISSN 0029-54 18</p>
    <note type="encoders-comment" subtype="missing-text" resp="JSA">The table of contents is ...
  </front>
  <body>
    <pb n="1"/>
    <div type="article">
      <head>Wjeliki kokot w Turjeju</head>
      <p><chi rend="bold">Z Turjeja/Tauer. </hi>Jaden z nejwětšych a wěšće teke nejrdnjějšyc
      <p>Dogromady 18 rejтары jo na jubilejnym kokoše před pisaćem towrynt pšiglědarjami w
      <byline>H. A.</byline>
      <figure>
        <p>Naš jubilej jo se wjelgin derje ražil - su gronili wěykne, ako su sobu gótowali
        <byline>Foto: Horat Adam</byline>
      </figure>
    </div>
    <div type="article">
      <head>Bórkowy: 15 zajmcow za rědownju notnych</head>
      <p><chi rend="bold">Z Pódstupima/Potadam. </hi> Na pósejženju Rady za nastupnoći Serbow
      <p>Nejprerweje pak jo šulake lěto se zachopilo a take něčo jo nanejžšnjeje k zachopjen
      <p>Serbski ceptař Uwe Gutšmýt jo weto optimistiski, aš jo móžno gromadu dostaš lichu w
      <p>K argumentam staršejšych, ako smy w Casniku wózwjawali, aš serbsku wucba wupadnje di
      <byline>I. N.</byline>
    </div>
    <div type="article">
      <head type="main">„Južo pšecaj knigly wjelgin lubujom“</head>
      <head type="sub">Bernd Pittkunings žěla něnto w Dolnoserbskej bibliotece</head>
      <p>W lěše 1991 su w Serbskem domje w Chóšebuzu založyli Dolnoserbsku biblioteku.W tych
```



<https://niedersorbisch.de/korpus/format/>

2020-08-14 TEI Format zur Annotation der sorbischen Korpora 1/27

TEI Format zur Annotation der sorbischen Korpora

Dieses Dokument beschreibt die Anwendung des TEI-P5-Formats als Modell und zur Kodierung der Inhalte der sorbischen Textkorpora.

Elemente der Korpusebene

teiCorpus

TEI-Modul	core – Elements Available in All TEI Documents
Unterelement von	—
Überelement für	teiHeader, TEI
Beschreibung	4 Default Text Structure • http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DS.html

Ein technisches Mittel, um mehrere Korpusdokumente zusammenzuführen (wenn in einer XML-Datei mehrere Dokumente enthalten sein sollen).
Hat keine semantische Funktion (wie z.B. ein Umschlag für eine Serie oder einen Jahrgang).

Elemente der Dokumentebene

(Die genaue Definition eines Korpusdokuments wird noch benötigt.)

Korpus text (XML-TEI) als Grundlage für verschiedene Anwendungen, zum Beispiel auch für eine **Digitale Bibliothek** →

Digitalisierung = neue Qualität

Katalog:











NIEDERSORBISCH DE
DIGITALE BIBLIOTHEK
INFORMATIONEN KATALOG SUCHE

DOLNOSERBSKI

Suche im katalog

Titel Autor 1899-1914

Anfrage »1899-1914«, Ergebnisse 1-10 von 40.

-  *Kužde to lět — nowy bśwít, Ale ta Boža gnada wolfoha niherrie* (Towarstwowe knihy, 29)
Jordan, Hajno (awtor)
Mašica Serbska, Wórzeje, 1900
-  *Lužizkizmo mjas zarnymi we Afrize* (Niederwendliche Missionsschriften, Nr. 2)
Merensky, Alexander (awtor), Jordan, Hajno (pšestajaf)
Barišiske misiońske towařstwo, Bariń, 1900
-  *Nowe a itare sorbske arje la dwa głoša*
Jordan, Hajno (redaktor)
Smoler, Budyšin, 1900
-  *Wětrny bñak Čhinaŕkego miliońtwa i młnin „Li-tihyung-jin“ abo „Li-Tichungin.“* (Niederwendliche Missionsschriften, Nr. 4)
Krušwica, Jan Djarnat (pšestajaf), Hubrig, ... (awtor)
Barišiske misiońske towařstwo, Bariń, 1901
-  *Troscht chorym a tužnym* (Towarstwowe knihy, 30)
Broniš, Matěj Bogumil (pšestajaf)
Lapsch, Wórzeje, 1901
-  *Mimo Boga na bśwěče* (Towarstwowe knihy, 32)
Jordan, Hajno (pšestajaf), Royová, Kristina (awtor)
Lapsch, Wórzeje, 1903
-  *Pichrowońc: silwo, kak 6e wot Jěšuša dajo hugojich, ale ľašej do bśwěta 6e ľašlapro a našćdku do nihernege 6aleńa pchizo* (Towarstwowe knihy, 33)
Šwjela, Kito (awtor)
Mašica Serbska, Chóšebuz, 1903
-  *Szym pjanowzy*
Šwjela, Bogumil (pšestajaf)
Lapsch, Wórzeje, 1903
-  *Kleiner Katechismus mit den vom Evangelischen Oberkirchenrat ausgewählten Sprüchen = Malý Katechismus le ľchpřucham*
Luther, Martin (awtor)
Lapsch, Wórzeje, 1904
-  *Drobenize 6e Bariškeje historje : k 25 lětnemu ľakóženikemu bśwěženju Malchize Sserbkeje we Dolnej Lužyzy* (Towarstwowe knihy, 45)
Jordan, Hajno (awtor)
Lapsch, Wórzeje, 1905

welero

Leseausgabe:

 *Bramborski Casnik, [lětnik 38] Probe-Nummer, stwórk, 26. mārca 1885*
Šwjela, Kito (redaktor), Jordan, Hendrich (redaktor)
Brandt + Wenzel, Chóšebuz + Wórzeje, 1885

Wopšimješe

- Roskłaženje
 - Nach kajžor 88 lět
- Worch Bismark
- Politka
 - Nimka
 - Awstria
 - Dta Braunichweigkeje
 - S Ruskojeje
 - Engelka
 - Nowinki
- S načheje hokolończi
 - S Dřewka
 - S Borkow
 - S Bóbowo
 - S Malych Dobryń
 - Se Skarbočča, 14. Merz
 - Se Stareje Darbuoje
 - S Rudneje pla Kalawy
 - Serlinke 6božowe wki 23. februara
 - Litnica

53

Bramborški Zašnik.

Serbke nowiny fa politiku a pohuzenje we hokrejfach Chočhebus, Kalawy a Grodk.

Probe-Nummer.

Stwórk ten 26. Merz. 1885.

„Zašnik“ pchizo na flwórk kuždeg tyženja a plałchi na bertyl ľeta 75 př., na polće 1 mk.

Wot fñatezyńerow 6e fa malu želonu 6mufchku 10 př. plałchi; pchi wězej raľach jo tuřliche.

Hudawaře a gorějbranje „Zašnika“ a tych fñatezyńerow jo w Chočhebulu 1) pla kn. J. Bojta bliko butrowych wikow, 2) pla kn. Poelcha, Grodkkej gašy No 115, a 3) pla kn. Rafalki pla Žandojkich rot, w Pizru pla kn. Zurabela pla zerke, we Wětoľchowe pla Schulza, w Gořenowe pla pichekuza kn. Schullera, w Dřewku pla kn. Petrika, w Borkowach pla kn. Reinh. Krügera atd.

53

Roskłaženje.

Nach nowy ľachopěrik 6e ľtani f Jolušom, tym kichizowanymi Zogdla to pchěrheńenje f načnym zašnikom? Tak budu načhe ľalowaje pchafchalch. Wotgrono: Juž dawno pchizochu na zašnikařa ľkaržby, až to ľopeno jo pomě nehužytnych fñatezyńerow, a až to pišmo řejo ľkoro wězej k ľalowawu. Pla tog donětejřchnege ľchifchčzařa řeběčho pomoz k namakařu, togdla 6om drugege ľchifchčzařa pýtal a f nim kontrakt ľgotowal, až ma won tak ľchifchčzalach, ak ja jo požadam. Smejoli weto hyřchčzi něcht ľkaržbu, abo niži dobru radu, ga kichěl mě to pišalch, ja a moej 6obuzašnikaf, kněs ľchulaf Jordan, bužomej 6e pilnowalch, aby ten Zašnik ľchym ľalowafam pchěze ľubřy horowal. Ten duch togo, zož pľchomy, wolfořa kaž doněta ten ľtary a niher pchawy: „F Bogom ľa kraľa a wolchzjoju řemju!“ Togdla pchóšymy, ab te ľtaro pľchjajchelo nam wěrne wolali a welo nowych ľalowafow pchiwabili.

Aspekte der Bedeutung von Daten am Beispiel - Eigennamen



Serbski Sorbisches
institut Institut

Namen als Wörter:

Brjazyna

Dorf in der Gemeinde Briesen

im Siedlungsgebiet der Sorben/Wenden in Brandenburg*

annähernde Koordinaten: [GEO 51,81554°N 14,2442°E](#)

veraltete Formen: **Brjazyn, Břazyn, Břazyń, Břazyna, Brásyna**

deutscher Name: **Briesen**

veraltete deutsche Form: **Briefen**

verbundene Lexik

- Einwohner: **Brjazynaŕ**
- Einwohnerin: **Brjazynaŕka**
- Adjektiv: **brjazyński**

	sg	du	pl
nom	Brjazynaŕ	Brjazynarja	Brjazynarje
gen	Brjazynarja	Brjazynarjowu	Brjazynarjow
dat	Brjazynarjeju	Brjazynarjoma	Brjazynarjam
acc	Brjazynarja	Brjazynarja Brjazynarjowu	Brjazynarje Brjazynarjow
instr	Brjazynarjom	Brjazynarjoma	Brjazynarjami
loc	Brjazynarju	Brjazynarjoma	Brjazynarjach

Aspekte der Bedeutung von Daten am Beispiel - Eigennamen



Serbski Sorbisches
institut Institut

Namen als Schlüssel zu Informationen/Wissen:

```
<body>
  <pb n="1"/>
  <div type="article">
    <head><w lemma="WJELIKI_aflex,WJELIK,WJELIKI_ad
    <p><hi rend="bold"><w lemma="Z">Z</w> <w lemma=
    <p><w lemma="DOGROMADY">Dogromady</w> <w xml:id
    <byline><w xml:id="dle807702">H.</w> <w xml:id=
    <figure>
      <p><w lemma="NAŠ">Naš</w> <w lemma="JUBILEJ
      <byline><w lemma="FOTO_dekl,FOTO_indekl">Fo
    </figure>
  </div>
  <div type="article">
    <head><w lemma="BÓRKOW,BÓRKOWY" type="name">Bór
    <p><hi rend="bold"><w lemma="Z">Z</w> <w lemma=
    <p><w lemma="NEJPJERWJEJ">Nejppjerwjej</w> <w le
    <p><w lemma="SERBSKI">Serbski</w> <w lemma="CEP
    <p><w lemma="K">K</w> <w lemma="ARGUMENT">argum
    <byline><w xml:id="dle808627">I.</w> <w xml:id=
  </div>
  <div type="article">
    <head type="main"><pc></pc><w lemma="JUŽO">Juž
    <head type="sub"><w lemma="BERND">Bernd</w> <w
    <p><w lemma="W">W</w> <w lemma="LÉŠE_aflex,LÉT,
    <p><w lemma="NC">NC</w><pc></pc> <w lemma="KAK
    <p><w lemma="BERND">Bernd</w> <w xml:id="dle809
    <p><w lemma="DLA">Dla</w> <w lemma="CO">cogo</w
    <p><w lemma="DOKULAŽ">Dokulaž</w> <w lemma="SOM
    <p><w lemma="A_bed2,A_bed1">A</w> <w lemma="JAD
    <p><w lemma="BYŠ,JO,WÓNO_flexl">Jo</w><pc></pc>
    <p><w lemma="BYŠ">Bužoš</w> <w lemma="W">W</w>
    <p><w lemma="TO,TEN">To</w> <w lemma="JA">ja</w
    <p><w lemma="KŠĚŠ,CO">Co</w> <w lemma="HĚŠ">maš
    <p><w lemma="GROMADA,GROMAŽE,GROMAŽIŠ">Gromažec
    <p><w lemma="A_bed2,A_bed1">A</w> <w lemma="TO,TEN">to</w> <w lemma="WAŽNY,NEJWAŽNJEJŠY">nejvažnjěše</w><pc></pc> <w lemma="BYŠ,BUŽO" type="name">bužo</w> <w
    <p><w lemma="JA">Ja</w> <w lemma="HOW">how</w> <w lemma="ŽĚLO,ŽĚL,ŽĚLAŠ">žělam</w> <w xml:id="dle809981">45</w> <w lemma="GÓŽINA">góžinow</w> <w lemma="WOB">wo
    <p><w lemma="A_bed2,A_bed1">A</w> <w lemma="NA">na</w> <w lemma="KÓŇC_inanim,KÓŇC_anim" type="name">kóńcu</w><pc></pc> <w lemma="BYŠ,BUŽO" type="name">bužo</w
    <p><w lemma="TO,TEN">To</w> <w lemma="SE_pron,SE_aflex">se</w> <w lemma="WĚŽEŠ_p,WĚŽEŠ_ip">wě</w><pc></pc> <w lemma="TO,TEN">to</w> <w lemma="WÓSTANUŠ,WÓSTAŠ"
    <p><w lemma="WUTŠOBNY">Wutšobny</w> <w lemma="ŽĚK">žěkc</w> <w lemma="A_bed2,A_bed1">a</w> <w lemma="WJELA_subst,WJELI,WJELE_aflex,WJELE_subst,WJELO" type="name
    <byline><pc></pc><w lemma="Z">z</w> <w xml:id="dle810323">B.</w> <w xml:id="dle810326" lemma="PITTKUNINGS_M2019" type="name">Pittkuningsom</w> <w lemma="BYŠ,J
    <figure>
      <p><w lemma="TO,TEN">To</w> <w lemma="BYŠ,JO,WÓNO_flexl">jo</w> <w lemma="NĚNT">něnt</w> <w lemma="NOWY_subst,NOWA_gross,NOWY_adj,NOWA_klein_dekl" type="na
      <byline><w lemma="FOTO_dekl,FOTO_indekl">Foto</w><pc></pc> <w xml:id="dle810398" lemma="MICHAEL_M2019" type="name">Michael</w> <w xml:id="dle810401" lemma
    </figure>
```


Aspekte der Bedeutung von Daten am Beispiel - Eigennamen



Serbski Sorbisches
institut Institut

Namen als Metadaten:

<docTitle>

<titlePart type=„main“>

<w lemma="BRAMBORSKI" cpc:norm="bramborski|Bramborski">**Bramborski**</w>

<w lemma="SERSKI" cpc:norm="serski|Serski">**ßerski**</w>

<w lemma="CASNIK" cpc:norm="casnik">**Zaffnik**</w><pc>.</pc>

</titlePart>

</docTitle>

[...]

<docAuthor xml:lang="deu">

<w>Herausgegeben</w>

<w>von</w>

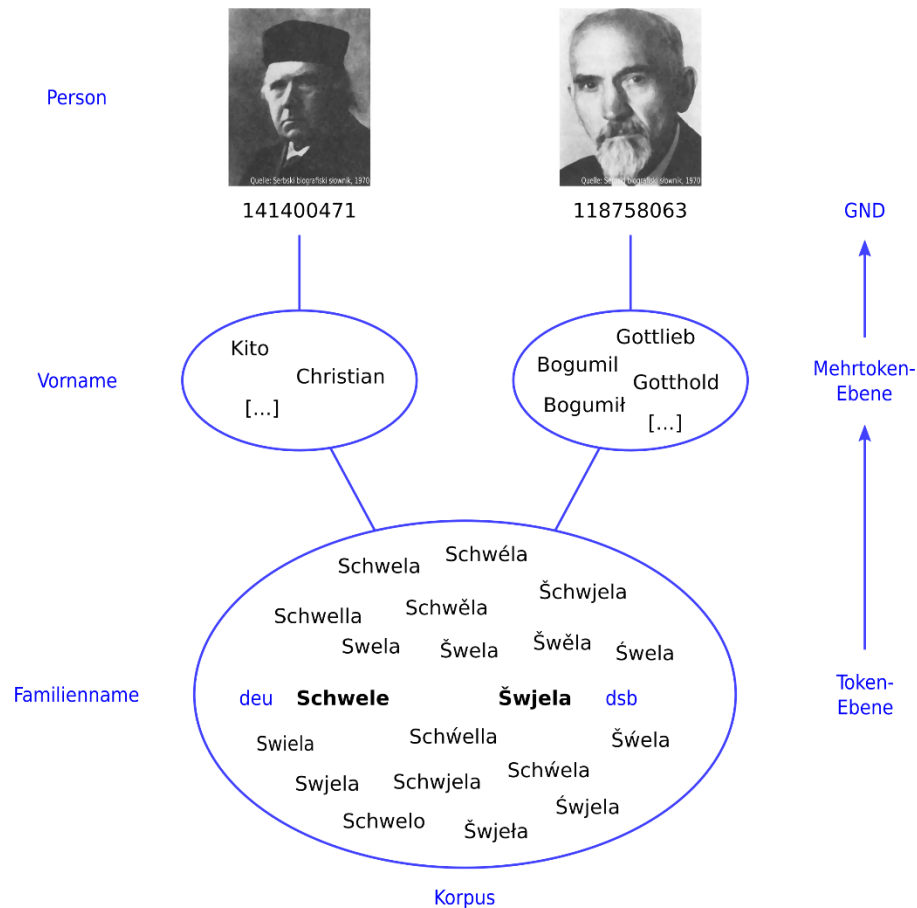
<w>**Nowka**</w><pc>.</pc>

</docAuthor>

Aspekte der Bedeutung von Daten am Beispiel - Eigennamen



Das Problem der Varianten:



Aspekte der Bedeutung von Daten am Beispiel - Eigennamen



Serbski Sorbisches
institut Institut

Namen als Teil des Kulturerbes:

DOLNOSERBSKI

NIEDERSORBISCH.DE

NIEDERSORBISCHE NAMEN

INFORMATIONEN NAMENSUCHE HILFE ZUR SUCHE ETYMOLOGIE

Namenerklärungen

Viele Menschen interessieren sich für den Ursprung der Namen. Deshalb wird dieses Portal schrittweise um Erklärungen zur Entstehung der einzelnen Namen erweitert. Zunächst wurde eine geeignete Gestaltung entwickelt und auf eine Auswahl von ungefähr 50 Orts- und Familiennamen angewendet. Dabei handelt es sich bislang um Prototypen.

Einleitung

- Hinweise zur Benutzung – unbedingt zuerst lesen
- Hinweise für Wissenschaftler
- Namenbedeutung
- Modulstruktur
- Prinzipien der Namenerklärung
- Sprachliche Rekonstruktion
- Schreibkonventionen

Namenverzeichnis

Nur die Namen, denen bereits eine Erklärung hinzugefügt worden ist, werden aufgelistet.

Ortsnamen

- Bóšojce / Bohsdorf
- Bronkow / Bronkow
- Carna / Zschorno
- Chóšebuz / Cottbus
- Drjenow / Drehnow
- Drjenow / Fürstlich Drehna
- Drjowk / Drebkau
- Gubin / Guben
- Kalawa / Calau
- Krosno / Krossen
- Kšušwica / Krausnick
- Libanojce / Laubsdorf
- Lubin / Lübben
- Lubnjow / Lübbenau
- Luboraz / Lieberose
- Mała Kšušwica / Kleinkrausnick
- Małe Libice / Klein Liebitz
- Popojce / Papitz
- Pynow / Pinnow
- Rakow / Raakow
- Rogow / Horno
- Rogow / Horno
- Rogow / Ragow
- Rogow / Ragow
- Tšupc / Straupitz
- Wjelika Kšušwica / Großkrausnick
- Wjelike Libice / Groß Liebitz
- Źarkow / Sarkow
- Źrakow / Schrakau

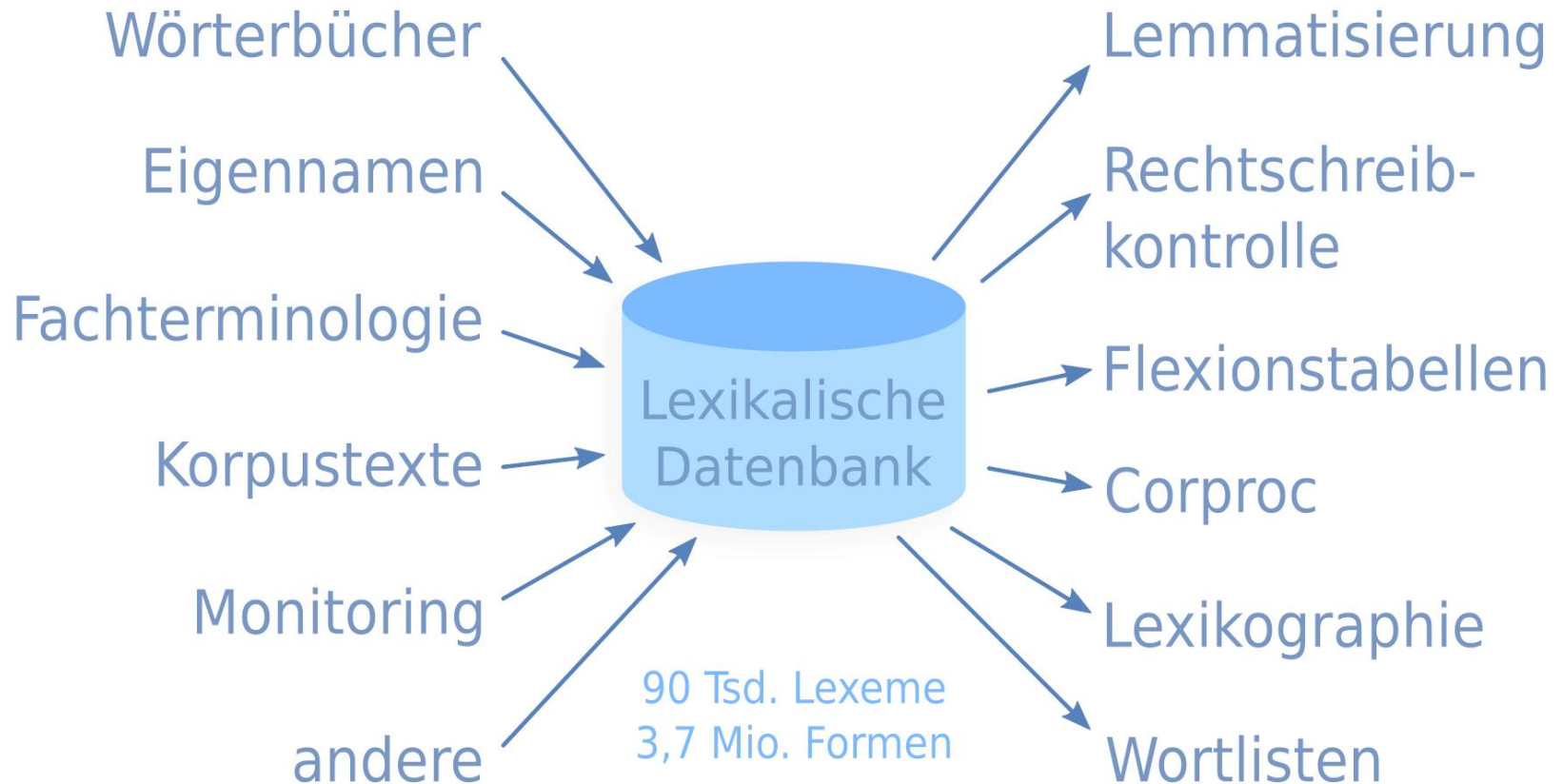
Personennamen

- Burik
- Buriš
- Domań
- Domanja
- Domańk
- Hupac
- Hupack
- Kócor
- Kócur
- Kokot
- Kóńcak
- Kóńcač
- Kowal
- Kral
- Libuš
- Libuša
- Lubiš
- Mudrik
- Mudrink
- Natus
- Natuš
- Pólak
- Wójtko
- Zeleń
- Zeleńk
- Zelezk
- Zelezko

Lexikalische Datenbanken als Teil der digitalen Infrastruktur



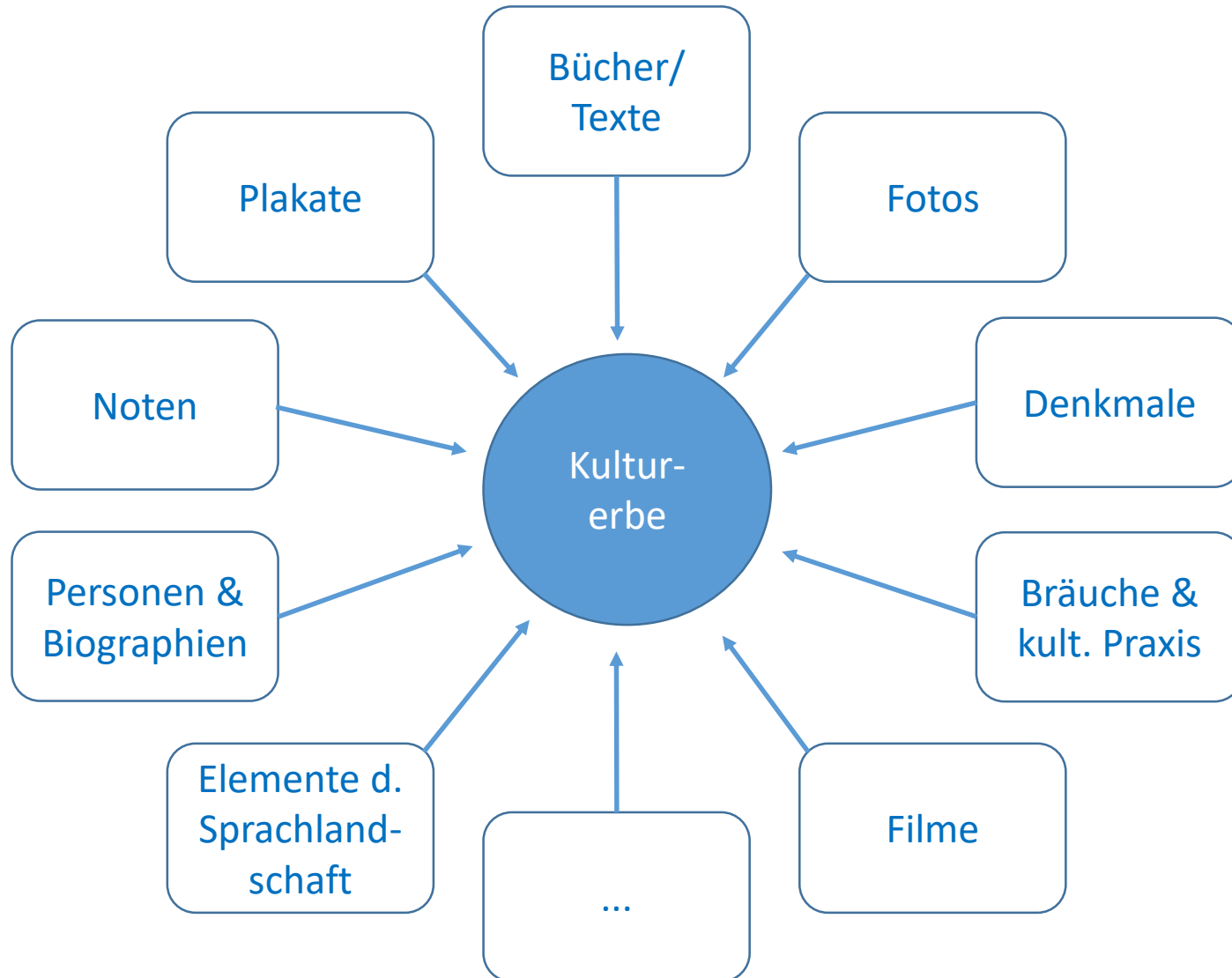
Serbski Sorbisches
institut Institut



Digitalisierung des Kulturerbes



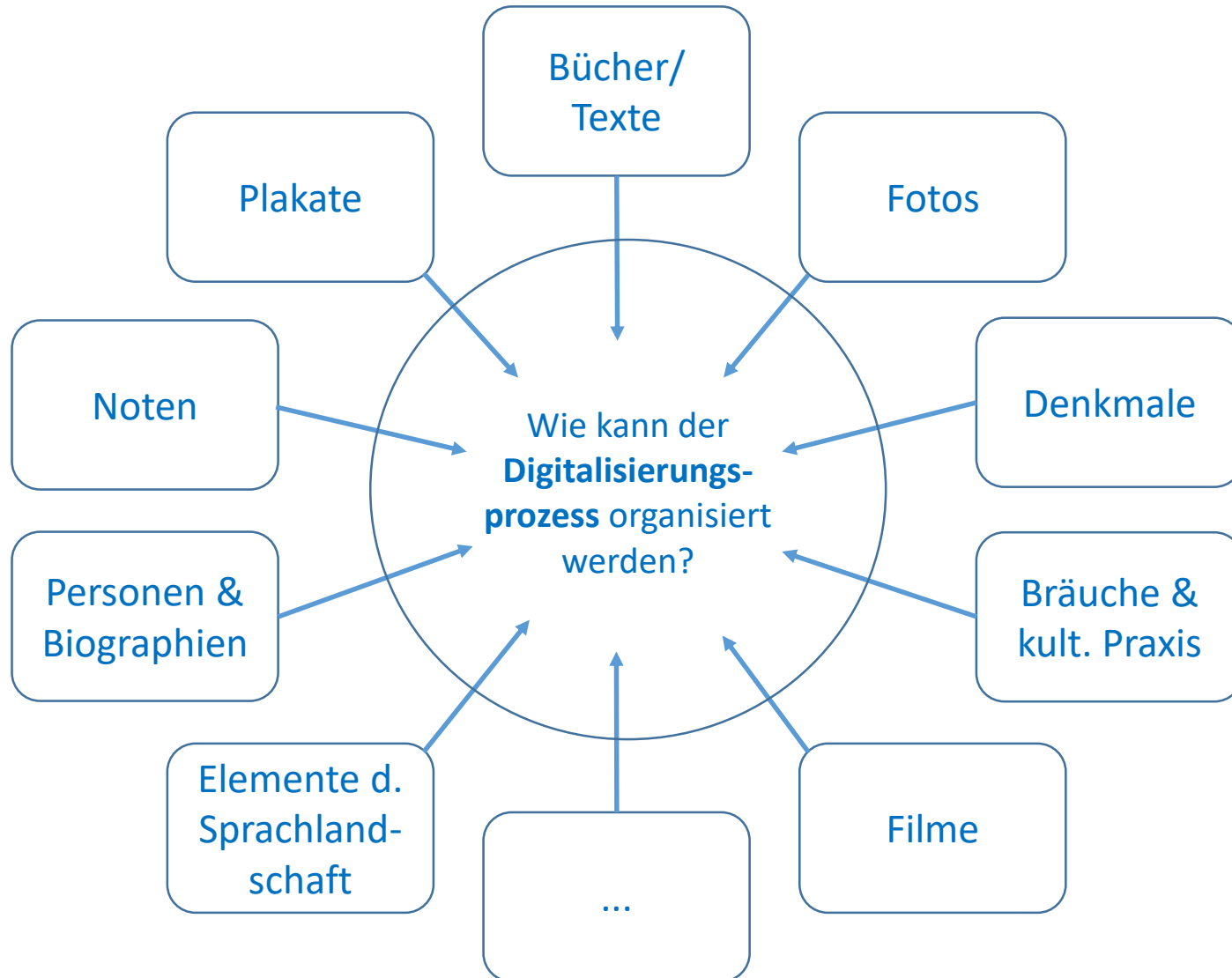
Serbski Sorbisches
institut Institut



Digitalisierung des Kulturerbes



Serbski Sorbisches
institut Institut



Infrastruktur zur KE-Digitalisierung



Serbski Sorbisches
institut Institut



Bedeutung des (dig.) Kulturerbes

Wichtig für die Weiterentwicklung des sorbischen Digitalkonzepts“ (mit Blick auf das Kulturerbe):

1. Die Bedeutung des Kulturerbes und seiner Digitalisierung ist größer als bisher (an)erkannt.
2. Der Digitalisierungsprozess lässt sich nicht durch trad. Projektförderung organisieren – notwendig: professionelle Organisation und Infrastruktur.
3. Wenn wir vorankommen wollen, brauchen wir
 - (a) (eine rechtzeitige) Entscheidung über die Einrichtung eines Zentralregisters Sorbisches Kulturerbe sowie
 - (b) eine Entscheidung für das **Digitalisierungszentrum**.

(Der Begriff “Digitalisierungszentrum” taucht im Digitalkonzept nicht auf.)
4. Priorisierung ist wichtig, sollte aber dynamisch sein!