



STAATSBIBLIOTHEK ZU BERLIN – PK

(NEUE) INFORMATIONSMANGEBOTE IM SLAVISTIK-PORTAL :

**MultiSlavDict & Co**

Vladimir Neumann

- Die Idee entstand beim Treffen des Fachbeirats im Jahre 2017
- 2018 begann die Vorbereitung für die Antragsphase 2019-2021
- Konzept (2019 auf Aufforderung der DFG nachgeschärft), hochwertige Daten durch Doble-Keying zu bekommen und diese in Datenbanken zu konvertieren.
- Bewilligung 2019, Freigabe der Mittel und Aufnahme des Konversionsverfahren 2020
- Arbeiten:
  - Proofreading
  - Datenstrukturierung / -konvertierung
  - Indizierung / Präsentation / Ausgabe

## 1. Phase (2019-2021): 3 Wörterbücher

	Kürzel	WB	Spezifika
1	LexiconPGL	Lexicon palaeoslovenico-graeco-latinum / Franz Miklosich. Vindobonae : Braumueller, 1862-1865. XXII, 1171 S.	Kirchenslawisch, Griechisch, Latein
2	KratkijSlovar	Kratkij slovar' šesti slavjanskich jazykov (ruskago s cerkovnoslavjanskim, bolgarskago, serbskago, češkago i pol'skago) a takže francuzskij i německij Franz Miklosich. - S. Peterburg : Vol'f, 1885, 955 S.	Kirchenslawisch, Russisch, Bulgarisch, Serbisch, Tschechisch, Polnisch, Deutsch, Französisch
3	SlownikPRN	Słownik polsko-rossyjsko-niemiecki / Johann Adolf Erdmann Schmidt. - w Wrocławiu : W. B. Korn, 1834. 917 S.	Polnisch, Deutsch, Russisch, Fraktur

## 2. Phase (2022-2024):

	Kürzel	WB	Spezifika
4	Pawlowsky	Russisch-deutsches Wörterbuch / Iwan Pawlowsky. 3. Aufl. 2. Abdruck. Riga 1911. XI, 1774 S.	Russisch, Deutsch, Kirchenslawisch, Fraktur
5	LexiconSGL	Srpski rječnik istumačen nemačkim i latinskijem riječima / Karadžić, Vuk Stefanović. Belgrad 1898. XLII S., 1 Bl., 880 S.	Serbisch, Deutsch, Latein, Fraktur, Diakritika
6	Želechowski1886	Ruthenisch-deutsches Wörterbuch / Eugen Zelechowski. 2 Bände. Leipzig 1886 (2 Bde: 1117 S.)	Ukrainisch, Deutsch, Fraktur
7	Miladinov1927	Bulgaro-nemski pulen recnik / Miladinov, Ivan An. († 1942). Sofia 1927, ~960 S.	Bulgarisch, Deutsch

<https://slavistik-portal.de/msd/>

The screenshot shows the homepage of MultiSlavDict. At the top, there is a navigation bar with links for 'Suche', 'Aktuelles', 'Neuerwerbungen', 'Zeitschriften', 'Datenbanken', and 'Über uns'. Below this is a banner with the 'slavistik-portal' logo and the text 'Portal slavistiki'. The main heading is 'MultiSlavDict Beta'. A descriptive paragraph states: 'Das multilinguale Wörterbuch der slawischen Sprachen (MultiSlavDict) ist eine Datenbank mit mehreren konvertierten historischen Wörterbüchern und enthält ca. 153.900 lexikographische Einträge (über 3.8 Mio. Wortformen) in zehn Sprachen.' Below the text are search filters for 'Alle Sprachen' and 'Info & Tools', along with 'Übersetze!' and 'Suchtipps' buttons. At the bottom, four book covers are displayed: 'LexiconPGL', 'Kratkij Slovar', 'Slovník PRN', and 'Pawlowsky'.

<https://slavistik-portal.de/dicthub/dict-pgl.html>

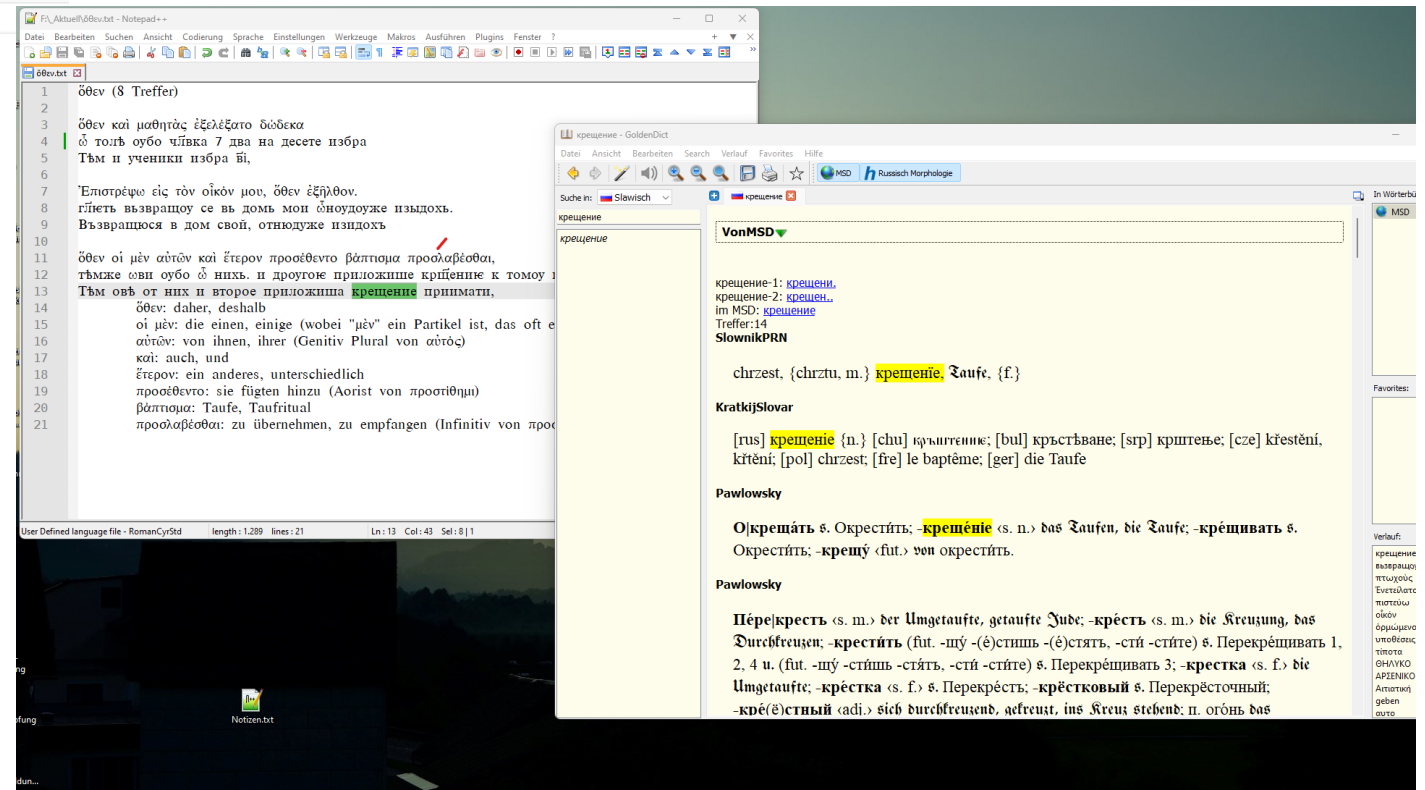
The screenshot shows the entry page for 'Lexicon palaeoslovenico-graeco-latinum' von Franz Miklosich (LexiconPGL) in the 'dicthub' section. The page title is '"Lexicon palaeoslovenico-graeco-latinum" von Franz Miklosich (LexiconPGL)'. A short description follows: 'Das "Lexicon palaeoslovenico-graeco-latinum" von Franz Miklosich, gedruckt in Vindobonae bei Braumueller, 1862-1865 wurde im Rahmen des Projektes MultiSlavDict verschriftlicht (über ein mehrstufiges Proofreadingverfahren) und in eine Datenbank konvertiert.' Below this is a search bar with a dropdown menu set to 'Alles', a search button, and a question mark icon. A pagination bar shows 'Paginierung: 1 2 3 4 5 6 »' and an 'Übersicht' button. The main content area displays two entries for the letter 'A':  
 1. **A**  
 prima alphabeti palaeo-slovenici littera, азъ (gen. аза, dat. азѡу *chrab.* 89.) dicta. A cyrillicum unitatem, „A numerum millenarium designat. [CIT]  
 2. **A** <adv.>  
 1) ἀλλά, δέ sed, vero *claz.* 1. 564. *sup.* 5. 16; 5. 20; 6. 16 etc. 2) καί et A τῶν γράδων κ' αὐτῆ καὶ τῶ ἐρχῆ πρός με *math.* 3. 14. - *ostrom.* земля вѣсѣи на водахъ, а тежка вещь соуци *hom.:mlh.* pro da positum putat *vost.* *nest.* 146. 18, ubi nos da scripsimus 3) a cum condicionali iunctum aequiparat аште: а бы не моанаъ са *sup.* 222. 17. а би былаъ съде *sup.* 226. 4. а бы аннакаъ смокве *sup.* 258. 9. а бышаъ не извѣганъ *sup.* 16. 29. а быша не знаанъ *sup.* 246. 14. а быша не извѣзанъ *sup.* 299. 4. радъ бы, а бы са доуша отъ тѣла не разоучнаа *men.* - *leop.* *nest.* 40.3. *rec.* 4.4; 61.13; 70.22. 4) ūna ut *chrab.* 1. 163. 12. cum dat. et inf. *chrab.* 1. 169. 6. а - не мѣ не заповѣдалаъ, а бы не повѣдала *esth.* 2. 10. *op.* 1. 55. *chrab.* 1. 145. 12; 200. 27; 224. 42. 5) а ли in interrogatione: а ли продаешин; *hom.:mlh.* 14. а ли надѣшеш се съдолѣти; *men.:mlh.* *men.:serb.* 145. 149. 151. 6) а ли δέ sed, vero; καί et а ли богъ послѡухъ мой *nest.* ингы съпасе, а ли себе не можеть съпастн *marc.* 15. 31. - *kyll.* ubi *ostrom.* interrogationem habet: а себе ли etc. а ли себе не можеть съпастн *math.* 27. 42. - *op.* 2. 2. 263. а ли тѣ градешн къ мѣркъ *math.* 3. 14. - *kyll.* *chrab.* 1. 168. 18. а ли in interrogatione habes *sup.* 165. 27; 226. 2; 313. 10; 315. 25; 385. 28. а ли богъ не испѣтаетъ срьдѣца *op.* 2. 2. 55. а ли дѣва не можеть родити *op.* 2. 1. 27. 7) ū vel *chrab.* 1. 163. 11, ubi alius cod. legit а лѡво. in *nest.* 52.35; 150.3; 158.18. scribendum est ѡли, unde ѡли, а ли orta sunt. а ли же ни *hor.* 18. 8) а лѡво pro а ли *chrab.* 1. 163. 1. 9) а не нъ мѣ ѡти multo minus *ezech.* 15. 5. in cod. saec. xv. а не нъ нама не прѣва *greg.:naz.* не

## API-Service (JSON):

[https://slavistik-portal.de/api/msd/?q=slav\\*](https://slavistik-portal.de/api/msd/?q=slav*)

## API-Webservice, Integration in lokale Anwendungen (hier: GoldenDict)

```
JSON Rohdaten Kopfzeilen
Speichern Kopieren Alle einklappen Alle ausklappen JSON durchsuchen
▼ responseHeader:
  status: 0
  QTime: 2
  ▼ params:
    q: "govoriti*"
    rows: "20"
    wt: "json"
▼ response:
  numFound: 17
  start: 0
  ▼ docs:
    ▼ 0:
      id: "ks_1403"
      ▼ rus:
        0: "вѣтѣствовати {v.}"
        chu: "вѣтовати"
        bul: "сладкодумен съм"
        srp: "беседити, говорити беседу"
        cze: "uměle, ozdobně mluvíti, býti řečníkem, krasomluvcem"
        pol: "być wymownym, krasomówcą"
        fre: "parler en orateur"
      ▼ ger:
        0: "zierlich reden"
        charo: "в"
      ▼ charo_f:
        0: "в"
        chart: "v"
      ▼ chart_f:
        0: "v"
        page: "72"
```



The screenshot shows the GoldenDict application interface. The search bar contains 'крещение'. The results are displayed in a list view. The first result is 'крещение 1: крещение', followed by 'крещение 2: крещен...', and 'im MSD: крещение'. Below the results, there are sections for 'Kratkij Slovar' and 'Pawlowsky'. The 'Kratkij Slovar' section shows the word 'крещение' in various languages: [rus] крещение (n.), [chu] крѣщение; [bul] кръстъване; [srp] крштење; [cze] křestění, křtění; [pol] chrzest; [fre] le baptême; [ger] die Taufe. The 'Pawlowsky' section shows the word 'Окрещать в. Окрестить: -крещение <с. п.> das Laufen, die Taufe; -крещивать в. Окрестить: -крещу {fut.} von окрестить.

## Liste der Kandidaten für die Antragsphase 2025-

- Weitere historische Wörterbücher mit ergänzender Funktion
- Welche weiteren Wörterbücher sind als Ergänzung sinnvoll?
  - Z.B. Leksikon von Pamvo Berynda (1627), Lexicon von Jan Mączyński (1564) oder Leksikon von Fedor Polikarpov (1704)
- Welche weiteren Services sind gewünscht / notwendig?

	Kürzel	WB	Spezifika
8	Vostokov1861	Vostokov, A. Ch.: Slovar' cerkovno-slavjanskago jazyka. Sanktpeterburg 1858-1861 (2 Bde: 255+299 S.)	Kirchenslawisch, Russisch, Griechisch, (hochwertige Digitalisate durch Slovanska knihovna angefertigt)
9	Danicic1863	Daničić, Đuro: Rječnik iz književnih starina srpskih. Beograd 1863 (3 Bde: XI, 521 + 519 + 596 S.)	Kirchenslawisch, Latein, Serbisch (mit Varianten aus Bosnien, Dubrovnik, mit altserb. Nomina proprium und Toponymen)
10	Šulek1860	Šulek, Bogoslav: Deutsch-Kroatisches Wörterbuch. Njemačko-hrvatski rječnik. Agram 1854-1860 (2 Bde: 1712 S.)	Kroatisch, Deutsch, Latein, Fraktur
11	SreznevskijMDS	Sreznevskij, Izmail Ivanovič: Materialy dlja slovarja drevne-russkago jazyka po pis'mennym pamjatnikam. Sanktpeterburg 1893-1912 (3 Bde: IX, 49 S., 1420 Sp. + 15 S., 1802 Sp + 1684, 272 Sp., 13 S. = ~2660 S.)	Kirchenslawisch, Altrussisch, Russisch, Griechisch, Multilingual (Hochwertige Digitalisate in der SBB)
12	Palkovic1821	Palkovič, Juraj: Böhmisches-deutsch-lateinisches Wörterbuch. mit Beyfügung der den Slowaken und Mähren eigenen Ausdrücke und Redensarten. Prag 1820-1821 (2 Bde: XIV, 1182 Sp.; Sp. 1184 - 2913, CXXXVI S. = notwenig ~1500 Seiten)	Tschechisch, Deutsch, Lateinisch (mit mährischen und slowakischen Äquivalenten), Fraktur
13	MiklosichEW	Miklosich, Franz. Etymologisches Wörterbuch der slavischen Sprachen. Wien 1886 (VIII, 547 S.)	Multilingual, Slawisch, Etymologien
14	Cnapius1754	Knapiusz, Grzegorz: Thesaurus R. P. Gregorii Cnapii e Societate Jesu Latino-Polono-Germanicus. Posnaniae 1754. (920 S., [49] Bl.)	Polnisch, Latein, Deutsch, Fraktur



<https://slavistik-portal.de/corphub.html>

[root][abs]

## (Slav)CorpHub

Textkorpus der kirchenslawischen Drucke

Das Korpus besteht aus mehreren Komponenten, die Übersichten und duchsuchbare Widgets beinhalten.

<a href="#">KSL-OldPrints</a>	Die Datenbank der kirchenslawischen Drucke der SBB-PK bietet eine Übersicht über die Drucke (momentan 144 Datensätze) und die ca. 1000 mit Transkribus gewonnenen und geprüften Seiten in Ground Truth-Qualität. Unter Detailsichten sind die Seiten mit den geprüften Volltexten zu finden, die auch in der Parallelansicht "Bild-Text" angezeigt werden können.
<a href="#">Suche in GT-Texten der KSL-OldPrints-digital</a>	Die Volltextsuche über ca. 1.000 Seiten der "Kirchenslavica Digital" der SBB-PK, die mit Transkribus mit verschiedenen (HTR+)-Modellen bearbeitet und anschließend manuell (Zeile für Zeile) bis zur Ground Truth-Qualität überprüft wurden. Die Volltextsuche erfolgt über ein Widget, das auf die SOLR-indizierten Textdaten zugreift. Die angezeigten Ergebnisse werden über einen PoS-Wrapper von "UDPipe" mit den grammatischen Informationen angereichert. Für die Suche im Index können "Reguläre Ausdrücke" (RegEx) verwendet werden.
<a href="#">Suche in RAW-Texten der KSL-OldPrints-digital</a>	Die Volltextsuche über ca. 90.000 Seiten "Kirchenslavica Digital" der SBB-PK, die mit Transkribus mit verschiedenen (HTR+)-Modellen bearbeitet und in so genannten RAW-Format automatisch (und ohne manuelle Prüfung) durch Skripting bereinigt worden sind. Die Volltextsuche erfolgt über ein Widget, das auf die SOLR-indizierte Textdaten zugreift. Die angezeigten Ergebnisse sind über ein PoS-Wrapper von "Mystem" mit den grammatischen Informationen angereichert. Die Suche kann trunkiert oder in N-Gram-Modus ausgeführt werden.
<a href="#">Kirchenslavica Digital</a>	Die kirchenslawischen Drucke der SBB-PK in digitalisierter Form können über die Digitale Bibliothek der SBB in hochauflösender Qualität angezeigt und heruntergeladen werden.

### Slawistische Korpora im Netz

Liste der slawischen Korpora, die im Netz zu finden sind.

<a href="#">CorpSlavColl</a>	Eine Sammlung der slawistischen Korpora im Netz (als Auszug aus dem Slavistik-Guide).
------------------------------	---

## KSL-OldPrints

Die Datenbank der kirchenslawischen Drucke der SBB-PK bietet eine Übersicht über die Drucke (momentan 145 Datensätze) und die ca. 1500 mit **Transkribus** gewonnenen und geprüften Seiten in Ground Truth-Qualität.

## Suche in GT-Texten der KSL-OldPrints-digital

Die Volltextsuche über ca. 1.500 Seiten der "Kirchenslavica Digital" der SBB-PK, die mit **Transkribus** mit verschiedenen (HTR+)-Modellen bearbeitet und anschließend manuell (Zeile für Zeile) bis zur Ground Truth-Qualität überprüft wurden. Die Volltextsuche erfolgt über ein Widget, das auf die SOLR-indizierten Textdaten zugreift.

## Suche in RAW-Texten der KSL-OldPrints-digital

Die Volltextsuche über ca. 90.000 Seiten "Kirchenslavica Digital" der SBB-PK, die mit **Transkribus** mit verschiedenen (HTR+)-Modellen bearbeitet und in so genannten RAW-Format automatisch (und ohne manuelle Prüfung) durch Skripting bereinigt worden sind.

- Welche Ausbaufunktionalitäten sind sinnvoll und notwendig? Ist die Erhöhung der Anzahl der „Grounth Truth“-Datensätze im KSL-Textkorpus nützlich?

The image shows two overlapping screenshots of the SLAW-DE website. The top screenshot shows the main page with a search bar and a list of search results. The bottom screenshot shows a detailed search result for 'mickiewicz'.

**Übersetzungsdatenbank (SLAW-DE) Beta**  
Die Übersetzungsdatenbank (SLAW-DE) ist eine Sammeldatenbank und enthält die bibliographischen Daten zu den Übersetzungen aus einer slawischen Sprache ins Deutsche.  
Die Liste der Übersetzer\*innen befindet sich hier.

Suche: alle -> ger | mickiewicz | Suche | ? | Sortierung: relevance

Paginierung: 1 | 2

Ihre Suche nach **mickiewicz** lieferte **18** Treffer

1	Die Krnischen Sonette des Adam Mickiewicz = [Sonety krymskie (dt.]] Mickiewicz, Adam - Posen, Decker, 1871
2	Adam Mickiewicz als Dramatiker - Dichtung und Bühnengeschichte "Dziady = Totenfeier", deutsche Übersetzung Mickiewicz, Adam - Bern, Lang, 1999
3	Todtenfeier [Dziady (dt.]] Mickiewicz, Adam - Leipzig, Breitkopf und Härtel, 1887
4	Konrad Wallenrod Mickiewicz, Adam - Berlin, Faudel, 1855
5	Herr Thaddäus oder der letzte Einritt in Löhauen Mickiewicz, Adam - Leipzig, Breitkopf und Härtel, 1882
6	Herr Thaddäus oder der letzte Einritt in Löhauen Mickiewicz, Adam - Leipzig, Breitkopf und Härtel, 1898
7	Gedichte Mickiewicz, Adam - Berlin, Nauck, 1836
8	Herr Thaddäus oder der letzte Sajassid in Löhauen [Pan Tadeusz (dt.]] Mickiewicz, Adam - Leipzig, Weber, 1836
9	Puppenmuseum Mickiewicz, Iwona - Berlin, Oberbaum, 1992
10	Lyrik Mickiewicz, Adam - Leipzig, Reclam, 1978

- Grundstock: 20.000 Datensätze aus dem Index Translationum
- Ergänzung durch Konversion der Metadaten aus dem Katalog der SBB (ca. 2000 DS)
- Sind die weiteren Ergänzungen gewünscht? (Katalog der BSB?, aus den Bibliographien?)
- Ist die Integration von Volltexttexten / Volltextsuche gewünscht / sinnvoll?



Vielen Dank für Ihr Aufmerksamkeit!

## Haus Unter den Linden

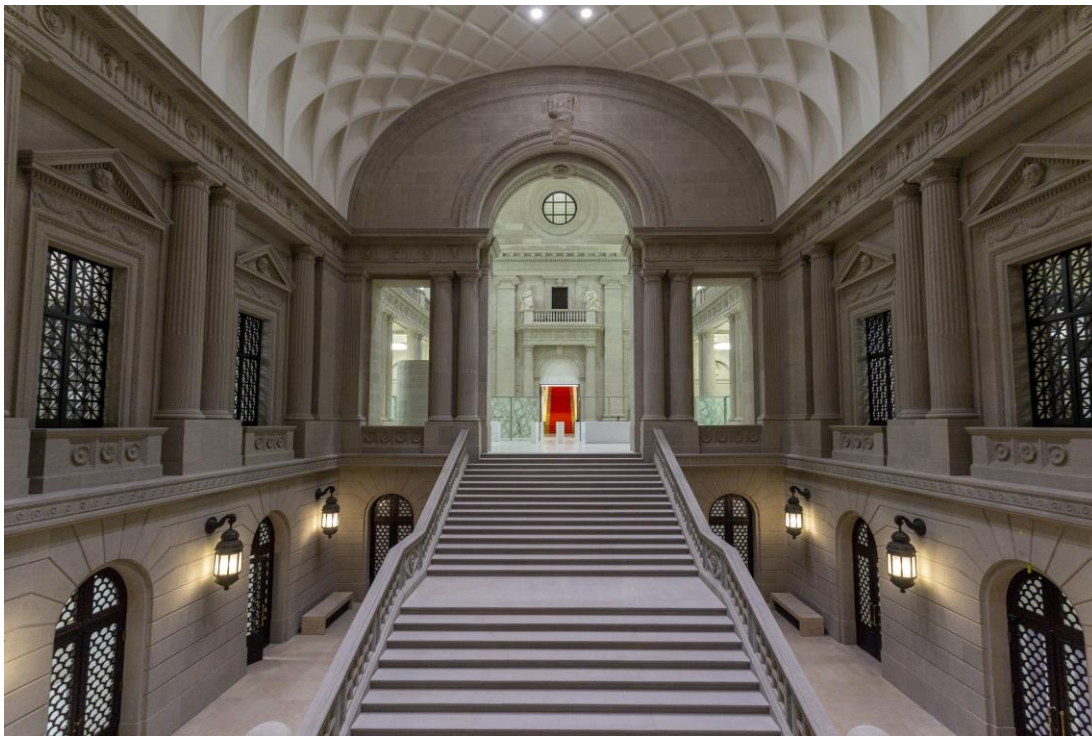


Foto: SBB-PK / Ralf Stockmann

## Haus Potsdamer Straße



Foto: Karl Johaentges