

GWDG NACHRICHTEN

03|16 Inhalt

.....

**4 Einheitliche Passwortrichtlinie im Bereich der
Universität Göttingen 6 Persistent Identifiers
for Research Data Management 9 Metadata
Quality in Europeana 13 Kurz & knapp
15 Dr. SOAP – oder wie ich lernte, den Domino
Web-Datenbankzugriff zu lieben
19 Personalia 21 Kurse**

Impressum

.....
Zeitschrift für die Kunden der GWDG

ISSN 0940-4686
39. Jahrgang
Ausgabe 3/2016

Erscheinungsweise:
monatlich

www.gwdg.de/gwdg-nr

Auflage:
550

Fotos:
© XtravaganT - Fotolia.com (1)
© Brian Jackson - Fotolia.com (4)
© intheskies - Fotolia.com (6)
© chagin - Fotolia.com (14)
© pterwort - Fotolia.com (18)
© Sashkin - Fotolia.com (20)
© MPLbpc-Medienservice (3, 19)
© GWDG (2, 21)

Herausgeber:
Gesellschaft für wissenschaftliche
Datenverarbeitung mbH Göttingen
Am Faßberg 11
37077 Göttingen
Tel.: 0551 201-1510
Fax: 0551 201-2150

Redaktion:
Dr. Thomas Otto
E-Mail: thomas.otto@gwdg.de

Herstellung:
Maria Geraci
E-Mail: maria.geraci@gwdg.de

Druck:
Kreationszeit GmbH, Rosdorf



Prof. Dr. Ramin Yahyapour
ramin.yahyapour@gwdg.de
0551 201-1545

Liebe Kunden und Freunde der GWWDG,

die Bewahrung von wissenschaftlichem und kulturellem Wissen ist eine wichtige Aufgabe, die insbesondere von Bibliotheken wahrgenommen wird. Die Erschließung dieser Informationen basiert dabei zusätzlich zu Veröffentlichungen in klassischen Medien wie Büchern und Journalen zunehmend auf digitalen Objekten. In dieser Ausgabe finden Sie einen Artikel, der sich mit dem Metadaten-Management in Europeana beschäftigt. Bei Europeana handelt sich um die virtuelle europäische Bibliothek, die das kulturelle Erbe Europas in digitaler Form für jedermann zugänglich macht. Sie ist ein Verbund von europäischen National- und Universitätsbibliotheken, Archiven und Museen, deren Bestände digital zugreifbar gemacht werden. Die digitalen Objekte werden nicht in einem zentralen Archiv gehalten, sondern verbleiben bei der ursprünglichen Institution. Die Kontext- bzw. Metadaten werden von Europeana gesammelt, um die Auffindbarkeit und Erschließung der digitalen Objekte zu gewährleisten. Die automatische Bewertung der Qualität der Metadaten ist dabei ein wichtiger Aspekt. Es freut mich, dass die GWWDG dieses wichtige Unterfangen unterstützen kann.

Wie Ihnen bekannt sein dürfte, ist Forschungsdaten-Management ein zentrales Thema der GWWDG, über das wir auch in den GWWDG-Nachrichten schon mehrmals berichtet haben. Wir hoffen, dass Ergebnisse aus diesem und anderen Projekten langfristig auch für allgemeine Forschungsdaten genutzt werden können, um die Arbeit mit Metadaten zu verbessern. In dem zugehörigen Artikel finden Sie einige Anhaltspunkte, welche Aspekte im Umgang mit Kontext- und Metadaten zu beachten sind.

Ich wünsche Ihnen viel Spaß beim Lesen dieser Ausgabe.

Ramin Yahyapour

GWWDG – IT in der Wissenschaft



Einheitliche Passwortrichtlinie im Bereich der Universität Göttingen

Text und Kontakt:
Andreas Ißleiber
andreas.issleiber@gwdg.de
0551 201-1815

Der Benutzung von Passwörtern gehört zum IT-Alltag. Damit Passwörter auch ihren Zweck erfüllen, sollten sie bestimmten Qualitätsanforderungen genügen. Für den Bereich der Universität Göttingen gibt es dazu eine einheitliche Passwortrichtlinie. Neben den üblichen Komplexitätskriterien bei der Passwortgestaltung soll auch die regelmäßige Änderung des Passwortes für eine möglichst hohe Sicherheit des Passwortes sorgen.

PASSWORTGESTALTUNG

Die Benutzung von Passwörtern ist seit Langem die gängigste Methode, um auf sichere, schützenswerte Daten zuzugreifen. Umso wichtiger ist ein verantwortungsvoller Umgang mit Passwörtern, damit Daten nicht in falsche Hände geraten. Hierbei ist immer wieder der Spagat zwischen „Merkbarkeit“ und Komplexität des Passwortes zu schaffen. Schon lange schützen einfache Passwörter nicht mehr vor dem Zugriff Dritter, weil sie leicht automatisiert erraten werden können. Auch viele Betriebssysteme sowie passwortgeschützte Zugänge erfordern mittlerweile ein Mindestmaß an Komplexität des Passwortes.

Zum Schutz vor Zugriffen durch unautorisierte Personen müssen Passwörter gewissen Qualitätsanforderungen (Komplexitätskriterien) genügen.

Kriterien für die Sicherheit eines Passwortes sind

- Mindestlänge,
- Groß-/Kleinschreibung,
- Verwendung von Sonderzeichen wie z. B. „!“ oder „-“
und
- Verwendung von Zahlen.

Vor diesem Hintergrund wurden auch die Anforderungen an die Komplexität des Passwortes, welches innerhalb der Universität Göttingen (einschließlich UMG) verwendet wird, entsprechend erhöht.

LAUFZEIT EINES PASSWORTES ALS ZUSÄTZLICHES KRITERIUM

Zusätzlich zu den o. g. Kriterien ist vor einiger Zeit von der Universität Göttingen ein weiteres Passwortkriterium verbindlich für alle ihre Benutzer vorgeschlagen worden: Es ist die Laufzeit eines Passwortes. Hierbei muss der Benutzer sein Passwort in bestimmten Abständen ändern. Die maximale Laufzeit eines Passwortes für die Universität Göttingen wurde auf ein Jahr festgelegt. Vor Ablauf dieses Zeitpunktes muss der Benutzer sein Passwort neu setzen. Dieses Kriterium ist Anfang des Jahres 2015 technisch umgesetzt worden.

Alle Benutzer bekommen dazu bereits drei Wochen vor Ablauf ihres Passwortes eine entsprechende E-Mail mit dem Hinweis, das Passwort am Kundenportal der GWDG zu ändern. Dieser Hinweis wird bei Nichtreaktion des Benutzers zunächst wöchentlich wiederholt und in der letzten Woche des Ablaufdatums dann sogar täglich. Um solch eine E-Mail von den vielen Phishing-E-Mails zu unterscheiden, die oft eine ähnliche Aufforderung enthalten, ist die E-Mail der GWDG selbstverständlich entsprechend elektronisch signiert.

Werden alle Warnhinweise zur Passwortänderung ignoriert, so wird der Account gesperrt. Eine Entsperrung kann dann nur noch durch den zuständigen Administrator des Instituts oder durch die GWDG erfolgen.

Hier ein Beispiel einer Erinnerungs-E-Mail der GWDG zur Passwortänderung:


Sehr geehrte(r) Lisa Mustermann

Ihr Passwort läuft am 26.02.2016 10:45 Uhr ab.
Bitte ändern Sie rechtzeitig Ihr Passwort, da mit Ablauf dieses Termins Ihr Account gesperrt wird.

Mit freundlichen Grüßen

Ihre GWDG

[Customer Portal](#) | [Support](#)



Gesellschaft für wissenschaftliche
Datenverarbeitung mbH Göttingen

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen
Am Fassberg 11 - 37077 Göttingen
Fon: +49 551 201-1523
E-Mail: support@gwdg.de

Geschäftsführer: Prof. Dr. Ramin Yahyapour
Aufsichtsratsvorsitzender: Prof. Dr. Christian Griesinger
Sitz der Gesellschaft: Göttingen
Registergericht: Göttingen
Handelsregister-Nr. B 598

PASSWORTRICHTLINIE DER UNIVERSITÄT GÖTTINGEN

Die Universität Göttingen hat folgende Richtlinie für die Verwendung eines Passwortes für alle ihre Benutzer verbindlich festgelegt:

1. Die Länge des Passwortes muss mindestens zehn Zeichen betragen.
2. Von den vier Zeichengruppen Großbuchstaben, Kleinbuchstaben, Sonderzeichen und Zahlen muss das Passwort mindestens drei abdecken.
3. Die Laufzeit eines Passwortes beträgt maximal ein Jahr.

WO KANN ICH MEIN PASSWORT ÄNDERN?

Die Änderung des Passwortes für Benutzer muss ausschließlich am Kundenportal der GWDG <https://www.gwdg.de> unter dem Button „ANMELDEN“ bzw. „SIGN IN“ erfolgen. In diesem Portal wird auch die „Korrektheit“ des Passwortes überprüft und damit auch die Einhaltung der o. g. Richtlinie.

PROJEKT „EINHEITLICHER MITARBEITER-ACCOUNT“ UND ZUGANG ZU SAP-SYSTEMEN

Abseits der hier definierten Passwort-Policy unterliegen Personen, die Zugang zu besonders sicherheitsrelevanten Diensten bekommen, in aller Regel einer noch strengeren Richtlinie hinsichtlich der Laufzeit ihres Passwortes sowie der Passwort-History. Bei der Passwort-History „merkt“ sich das System die letzten fünf

verwendeten Passwörter. Ein neu definiertes Passwort darf dann keinem der bereits verwendeten fünf Passwörter entsprechen.

Im Projekt „Einheitlicher Mitarbeiter-Account“ am Standort Göttingen, welches in Kürze in die Produktivphase übergeht und über das wir in einer späteren Ausgabe der GWDG-Nachrichten berichten werden, wird eine Vereinheitlichung der Benutzernamen sowie der Passwortrichtlinien für den gesamten Bereich der Universität Göttingen sowie der UMG und der GWDG angestrebt. In diesem Zusammenhang ist die hier beschriebene Erhöhung der Komplexitätsanforderung bei der Passwortgestaltung zwingende Voraussetzung.

FAZIT

Der verantwortungsvolle Umgang mit Passwörtern wird immer wichtiger, da auch die Methoden und Möglichkeiten Dritter, illegal an Passwörter zu gelangen, immer vielfältiger und raffinierter werden. Die Erhöhung der Passwortkomplexität ist eine wichtige Reaktion auf diesen Umstand und sollte überdies durch einen gewissenhaften Umgang des Benutzers mit dem eigenen Passwort flankiert werden.

WEITERE HINWEISE

Hinweise des Bundesamtes zur Sicherheit in der Informationstechnik (BSI) zur Gestaltung eines sicheren Passwortes:

https://www.bsi.bund.de/DE/Themen/ITGrundschutz/ITGrundschutzKataloge/Inhalt/_content/m/m02/m02011.html

Improvements for the password policy of the University of Göttingen

The use of passwords have been around since a long time. A password is the favorite authentication criteria at all for many systems. It should protect you and your data against unauthorized access. For the entire network of the University of Göttingen, we have a global and unique password policy with different complexity rules. Since January 2007, the passwords must meet at least three of four minimum requirements.

To improve the quality of passwords, we had to have at least one more criteria. Since January 2015, the users passwords must have a minimum length of ten characters. Aside from that, the password must be periodically changed by the user himself. The defined time frame for the periodic password changes is maximum one year. With that condition, the users can make sure that they use a strong password which is sufficient for all local services.

TECHNICAL REALIZATION

The fundamental function of a PID system is to resolve a research data resource's immutable and globally unique identifier into its current Internet address. Therefore, the core components of any PID system are a naming scheme and a resolution system. The naming scheme is a syntax definition for the identifiers to be resolved. Usually the name is composed of a prefix and a suffix, which is comparable with a telephone number. An elaborate resolution system is required in order to make a PID actionable in the sense that it functions in the current Internet. In addition, a resolution system has to guarantee sustainability, trustworthiness, high performance and high availability.

Various PID systems have been developed, whereas each is favoured in specific areas. Among them the DOI System [1] is the most prominent one, which is particularly established in the

publication area. Nevertheless, the Handle System [2] is the most important PID system, because it provides the fundamental basis for several other PID systems. It stands out because of its compactness, versatility and especially its resolution system.

HANDLE SYSTEM ARCHITECTURE

PIDs resolved by the Handle System are composed of two parts: its prefix and a unique local name under the prefix known as its suffix. The prefix and suffix are separated by the ASCII character „/“. The prefixes are controlled by the CNRI (Corporation for National Research Initiatives) with an annual registration fee of \$50 per prefix, which additionally covers the usage of the Handle System Resolution Infrastructure as depicted in figure 1. The Handle System architecture is composed of three entities: Global Handle Record, Local Handle Service and the Global Proxy.

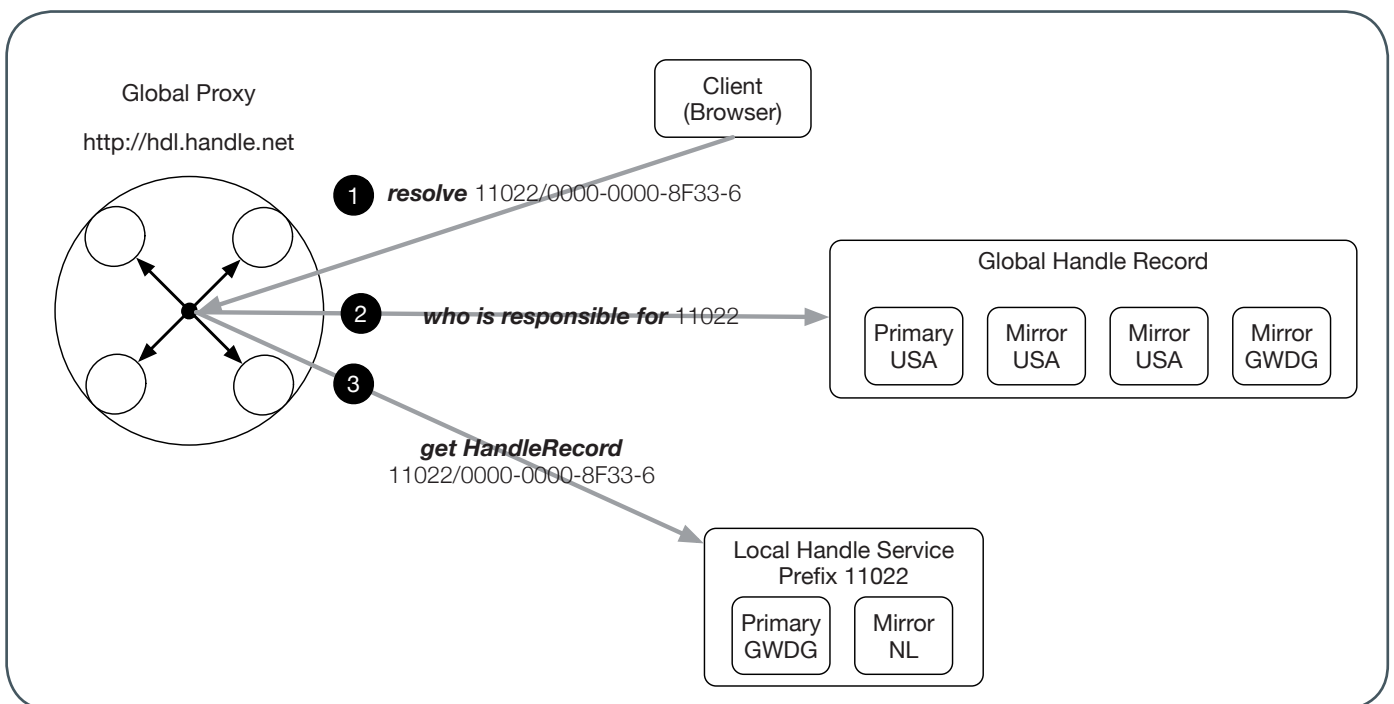


Figure 1: The Handle System Architecture

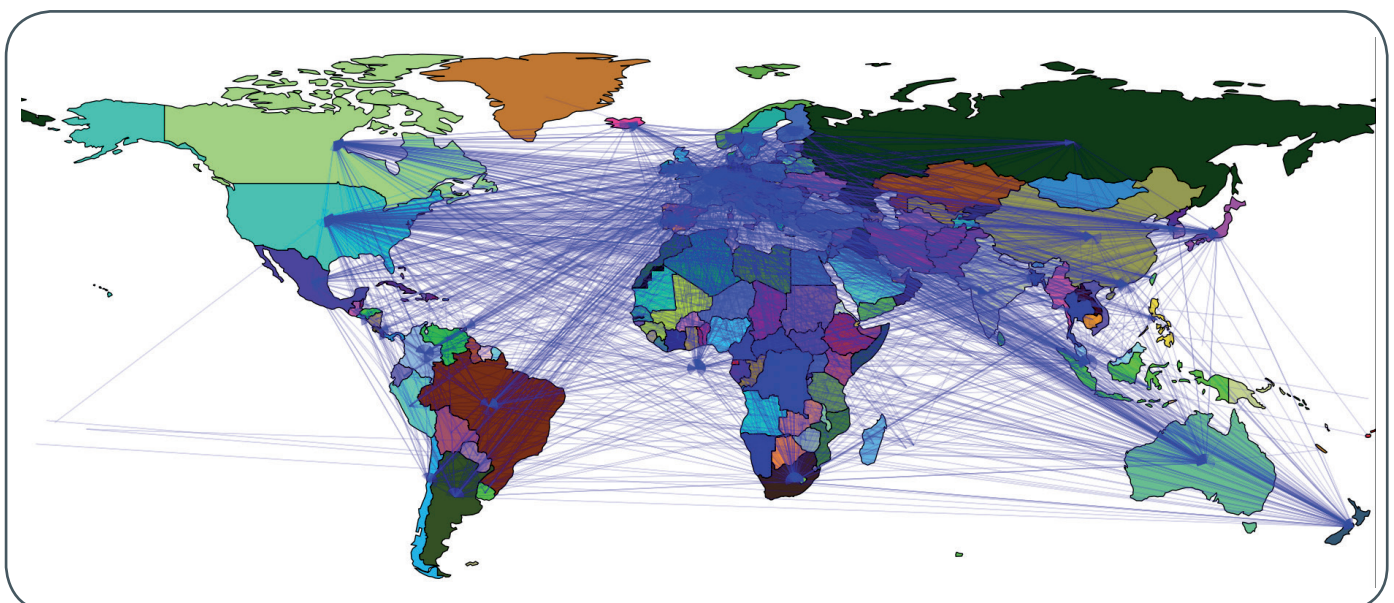


Figure 2: Resolution Path: from User Location to Resource

The Global Handle Record (GHR) is a distributed system consisting of one primary server located in the US and three mirror servers, whereas one of them is hosted at GWDG. The GHR records the information about each prefix' responsible Local Handle Service (LHS). A prefix' record is basically a list of IP addresses of servers, which compose the Local Handle Service of the corresponding prefix. Each Local Handle Service is composed of at least a primary server and optionally one more mirror servers.

To retrieve a resource, which is referenced by a PID, its PID has to be resolved by the Global Proxy to its current Internet address.

The Global Proxy, also composed as a distributed system, consists of four servers, whereas one of them is again hosted by GWDG. The resolution process is depicted in figure 1. As we can see, a client, who wants to resolve the particular PID „11022/0000-0000-8F33-6“ sends the resolution request to the Global Proxy. The Global Proxy in turn, internally queries the GHR about the responsible LHS. In the concrete case of prefix „11022“, which belongs to GWDG, it receives two IP addresses: the primary server hosted at GWDG and the mirror server hosted at a partner datacentre in Amsterdam. We will discuss this partnership later in detail. Having the IP addresses of the responsible LHS, the Global Proxy sends the resolution request to that LHS. The LHS responds with the corresponding Internet address. Finally, the Global Proxy redirects the client to that address. At the GWDG mirror proxy server there are up to 9 Million requests per month issued from clients all over the world. Figure 2 shows the resolution path recorded at the GWDG mirror proxy server in summer 2015.

As we could see, the Handle System is pretty much an imitation of the DNS-System. However the DNS-System is much more hierarchical than the Handle System, but in principle both systems are resolution systems. A query in the DNS-System results in a DNS-Record whereas in the Handle System it results in a Handle-Record.

PERSISTENT IDENTIFIERS AT GWDG

Persistent Identifiers play a important role at GWDG. It is a founder member of the European Persistent Identifier Consortium (ePIC) [3], a consortium consisting of European datacentres with the objective to provide Persistent Identifiers for the European Research Community. The ePIC PID system can be considered as a high level service on top of the Handle System, comparable with the DOI System. The key features of ePIC are:

- **RESTful API:** Since recent research data repositories mostly rely on RESTful interfaces, the ePIC-API has been developed to act as a translator from a REST interface into the native Handle Protocol interface. Until recently, the Handle System did not provide a native RESTful Interface in order to manage Handle-PIDs. The ePIC software stack is depicted in figure 3.

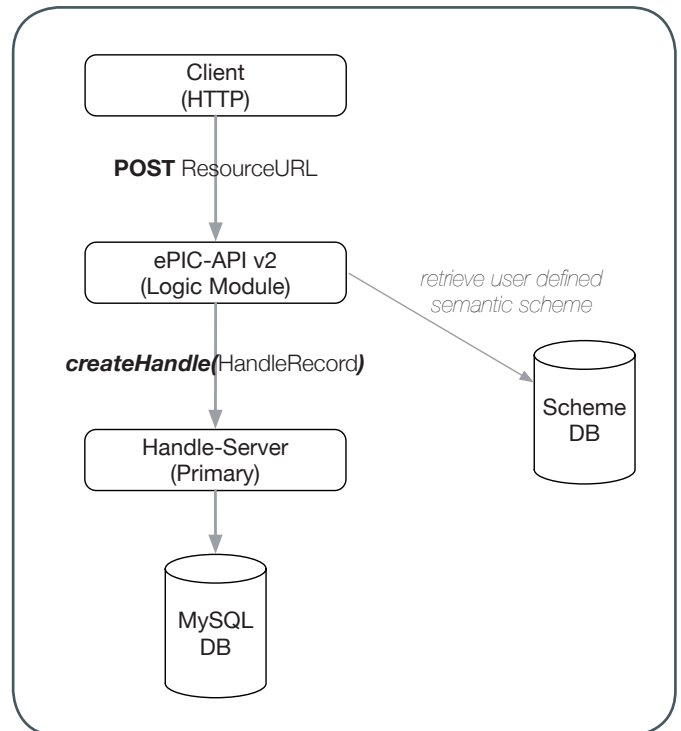


Figure 3: The ePIC Software Stack

- **Semantic Data Model:** As mentioned before, the real challenge with research data is to enable the extraction of as much information as possible from a research data resource. Therefore, an important feature of the ePIC-API is the ability of ingestion of semantic information into the PID-Record. This is managed by a logic module as can be seen in figure 3.
- **High-Availability:** Since the ePIC members are datacentres, all the ePIC PID-Records are replicated within the members. This is highly important in order to provide high available resolution and long-term sustainability of PID-Records. Currently, there are more than 20 Million PID-Records registered in the entire consortium.

In addition to its role as current chair of ePIC and its mirror role in the Global Handle Infrastructure, the role of GWDG has recently even more increased. GWDG is now also able to register and control prefixes, which will consolidate GWDG's position in the research data management area.

REFERENCES

- [1] <https://www.doi.org>
- [2] <http://www.handle.net>
- [3] <http://www.pidconsortium.eu>
- [4] <https://rd-alliance.org>
- [5] <http://www.heise.de/developer/artikel/Microservices-im-Zusammenspiel-mit-Continuous-Delivery-Teil-1-die-Theorie-2376386.html>

Metadata Quality in Europeana

Text und Kontakt:

Peter Király
peter.kiraly@gwdg.de
0551 39-20557

In several large digital collections the only way a user could access a digital object (such as a digitized cultural heritage material or a born digital research data) is via its metadata. If metadata is not precise, contains inappropriate or less information users miss the object, and the data creators' energy they put in the creation and maintenance of it will not lead to results. Europeana, the digital library of Europe launched a working group to define metrics and make proposals for improving the quality. GWDG supports this effort in a research project.

INTRODUCTION

Europeana [1] is the European digital library. It collects the metadata of digitized objects from more than 3,500 European libraries, archives, museums and similar cultural heritage institutions. At the time of writing the collection consists of more than 48 million objects. The objects stored in Europeana are not full text books, high quality images or videos, only their metadata which describe the more important properties of the digitized objects, such as title, creator and subjects.

The metadata elements are usually organized in metadata schemas, which are similar than database schemas, but instead of technical details they concentrate more on the semantic level. Europeana developed its own schema, called Europeana Data Model (EDM), and it covers almost 200 different data elements. Most of the partner organizations use some other metadata schemas which fit for the domain of their objects and for their existing workflow and technology stack. Some of the most important ones are MARC (Machine Readable Catalogue) for books, Dublin Core for digital library objects, EAD (Encoded Archival Description) for archival materials, LIDO (Lightweight Information Describing Objects) for objects in museums.

When these institutions provide records for Europeana they transform them to EDM. The expressive power of all those schemas are different because they were developed for specific use cases. There is however a common important feature: each one are loose standards, they allow quite a space for the cultural heritage institutions for their own cataloging habits and traditions (don't forget that most of the institutions have much longer history than that of these standards, and the specific way they create their catalogue is part of this history). There are differences in the roles of the catalogues even inside a single institution type and there can be big differences among records inside one catalogue. For example records created for an exhibition usually contain lots

of details, while there are records having only the minimal set of fields. And since we are talking about an international organization, there are many differences because of the variations in terms of languages and cultural traditions.

This various origin of Europeana's collection resulted a heterogeneity in quality of the metadata. The main functionality of Europeana is to provide the main catalogue of European cultural heritage. The user first finds a record and if it is interesting for him he follows the links to the data provider's web site where he can access the digitized object. Low metadata quality means in this context that the record does not have enough information content to fulfill its purpose and thus the user is not able to access the digitized object. In the past two years the topic of data quality became more and more important. The Europeana Network [2] created a Metadata Quality Task Force [3] to start investigations in this topic. In May, 2015 they published a detailed report [4] highlighting the overall issues and analyzing some typical data quality problems,

Metadatenqualität in Europeana

In mehreren großen digitalen Sammlungen kann ein Nutzer nur über die Metadaten eines digitalen Objektes (z. B. digitalisiertes Kulturerbematerial oder ursprüngliche digitale Forschungsdaten) auf dieses zugreifen. Wenn die Metadaten nicht präzise genug sind, ungeeignete oder zu wenige Informationen enthalten, dann kann der Nutzer das Objekt nicht finden und die Energie, die die Ersteller der Daten für dessen Erzeugung und Pflege eingesetzt haben, führt zu keinen Ergebnissen. Europeana, die digitale Bibliothek Europas, hat eine Arbeitsgruppe gegründet, die Metriken definieren und Vorschläge zur Verbesserung der Qualität von Metadaten machen soll. Die GWDG unterstützt diese Bemühungen in einem Forschungsprojekt.

but they did not create plans to measure the issues and fix the problems. This year Europeana Network founded the Data Quality Committee working group which aims to create a plan „to agree on new standards based on these recommendations, and how providers and aggregators can implement these standards.“

Before joining GWDG the author of this article was an Europeana developer who was involved in analysing metadata issues, so he was asked to participate in this committee. In a research project GWDG will support this effort with expertise and computational resources for set up a measuring framework. We hope, that the tools which will be developed here could be reusable in core GWDG services, for example in the planned research data management tool, since the metadata of research data and cultural heritage objects has something common: they describe the underlying content in order to provide access to the „real“ data.

DATA WORKFLOW

We mentioned that each record has a history before they reach Europeana. Within Europeana the records are transformed, normalized, enriched, stored and indexed – each of the processes change them some way or another and it affects data quality as well. It is apparent that the best place where the record can be improved is at the source, at the data provider’s system. The Committee will do recommendations for all phases of the data workflow. These will be incorporated in different documents: most importantly in the guidelines and documentations of EDM schema, and will be discussed in different forums such as at each Europeana Aggregator Forum. (In current data workflow there is an extra organizational level between Europeana and data providers: the aggregators. They are thematical or regional hubs which collect metadata from the organizations and pass to Europeana.)

THE ACTIVITY OF THE WORKING GROUP

The Committee considers that data quality is always relative to intended use and cannot be analyzed or defined in isolation from it, as a theoretical effort. The main focus right now is investigate the use cases regarding the content (re-)use. We hope that the conclusions drawn from Europeana’s use cases could be applicable to the individual domains (archaeology, art, libraries...).

The actual work will be done in the following areas: [5]

1. Mandatory metadata elements for ingestion of EDM data

- Investigating if the current mandatory elements are relevant and sufficient. It also needs to propose how to make legacy data compliant with the agreed list of mandatory elements, or whether to disable it in some (data) exposition channels. Europeana still has several hundred thousand records that don’t have all current mandatory elements.
- Recommendations on measures of completeness for descriptive metadata based on presence/absence of fields.
- Revisiting EDM definitions and guidelines, so that it makes EDM more precise and helps data providers to make consistent, less ambiguous choices.

2. Data checking and normalization

- Looking into ways and rules to normalise metadata. This

includes the use of vocabulary based values or normalised values [6].

- Recommendations for tools and services to validate or detect anomalies in EDM data (either at Europeana or provider/aggregator level) using technology such as XML validation or RDF Shapes.

3. Meaningful metadata values (in the context of use)

- Metadata values intended for human consumption (e.g. titles, descriptions in object metadata or labels attached to URIs) that are not unique or understandable are a concern, to just give one example. The Committee should look into ways of recommending meaningful metadata values and indicators to measure improvements.
- This work includes measures for information value of statements (informativeness, degree of multilinguality...).
- Making recommendations about and assessing the impact on quality of contextual links, including these created via semantic enrichment.

4. Quality of content (digital media)

- The Committee may look at the quality levels for all types of content that are proposed by the Europeana Publishing Framework [7] and work on improvements.

5. Coordination with other quality-related initiatives

- There are some parallel investigations which worth considering for example the Digital Public Library of America [8] initiated discussions on metadata quality withing their database.

The working group uses Basecamp [9] and bi-weekly phone conferences for communication, plus face-to-face meetings are planned twice a year in The Hague as a side event of the regular Aggregators Forum.

METRICS OF METADATA QUALITY

However the working group started working on one metadata quality metric, in the related literature we can find relevant definitions. Here we will show some of them, the ones we plan to implement in the a metadata quality assurance framework [10].

Completeness

Completeness is the ratio of metadata fields fulfilled with values. It might have multiple aspects, since we can organize the elements into different groups according to their functions. The working group set up the following initial list of groups: mandatory elements, descriptiveness, searchability, contextualization, identification, browsing, viewing, re-usability, multilinguality. Each group collects several fields which support the given feature for example identification group contains those fields which play role in identifying the object (title, description, identifier etc.). The individual fields might have different importance as well, and we can assign weighting to them to refine the calculation.

Conformance to expectations

This metric measures the information content (or the entropy) of the individual fields. The basic idea is simple: the more unique

a field value, the more valuable. Measuring projects might follow three different approaches. For categorical values they can measure the ratio of each individual categories. For free text field they can use the whole content of the field and/or the individual terms of them. Modern search engines (such as the Apache Lucene-based Apache Solr and Elasticsearch) provide the „term vector space“ model for the calculation of these values.

Logical consistency and coherence

This metric evaluates how the record follows the rules of the underlying metadata schema. It could be a good measurement in case where the schema contains lots of formal rules. In case of Europeana it has less importance because the schema has changed significantly over the times in this aspect and a very strict metadata schema would not fit for the above mentioned heterogeneity of the database.

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF>
  <ore:Proxy rdf:about="http://data.europeana.eu/proxy/provider/2022105/urn_axmedis_00000_obj_df3788a6_617a_4c3e_8965_df41d6ca929">
    <dc:contributor>Personaggi e interpreti: Il conte D'Almaviva - Mario Zeffiri, Daniele Zanfardino (R, S1, S2); Don Bartolo - Alberto Rinaldi; Rosina - Anna Bonitatibus, Francesca Provvisionato (C, R, S1, S2); Figaro - Christian Senn, José adan Pérez (R, S1, S2); ...</dc:contributor>
    <dc:coverage>Milano,
      Coordinates:
      Lat: 45 28 00 N degrees minutes Lat: 45.4667 decimal degrees
      Long: 009 12 00 E degrees minutes Long: 9.2000 decimal degrees</dc:coverage>
    <dc:creator>E.A.R. Teatro Massimo Bellini</dc:creator>
    <dc:date>2011</dc:date>
    <dc:description xml:lang="it">Locandina, con disegno di dario Fo, dell'opera „Il barbiere di Siviglia“ di Gioacchino Rossini, presentato dal Teatro Massimo Bellini, Catania. Regia di Dario Fo.</dc:description>
    <dc:format xml:lang="it">pdf</dc:format>
    <dc:language xml:lang="it">it</dc:language>
    <dc:rights xml:lang="it">CTFR</dc:rights>
    <dc:subject xml:lang="it">immagine</dc:subject>
    <dc:title xml:lang="it">Il barbiere di Siviglia - Locandina stagione 2011</dc:title>
    <dc:type xml:lang="it">Locandina</dc:type>
    <dcterms:extent xml:lang="it">2 pag</dcterms:extent>
    <dcterms:issued>14/03/1987 (first performance)</dcterms:issued>
    <dcterms:issued>maggio 10, 12, 13, 15, 17, 18, 19 - 2011</dcterms:issued>
    <dcterms:spatial>Teatro Massimo Bellini</dcterms:spatial>
    <dcterms:spatial>Catania</dcterms:spatial>
    <dcterms:spatial>Italia</dcterms:spatial>
    <edm:europeanaProxy>false</edm:europeanaProxy>
    <ore:proxyFor rdf:resource="http://data.europeana.eu/item/2022105/urn_axmedis_00000_obj_df3788a6_617a_4c3e_8965_df41d6ca929"/>
    <ore:proxyIn rdf:resource="http://data.europeana.eu/aggregation/provider/2022105/urn_axmedis_00000_obj_df3788a6_617a_4c3e_8965_df41d6ca929"/>
    <edm:type>TEXT</edm:type>
  </ore:Proxy>
  <edm:Place rdf:about="http://sws.geonames.org/7521391/">
    <wgs84:lat>37.50419</wgs84:lat>
    <wgs84:long>15.09001</wgs84:long>
    <skos:prefLabel>teatro massimo bellini</skos:prefLabel>
    <dcterms:isPartOf rdf:resource="http://sws.geonames.org/3175395/">
  </edm:Place>
  <edm:TimeSpan rdf:about="http://semium.org/time/2011">
    <skos:prefLabel>2011</skos:prefLabel>
    <dcterms:isPartOf rdf:resource="http://semium.org/time/20xx_1_third"/>
    <edm:begin>Sat Jan 01 01:00:00 CET 2011</edm:begin>
    <edm:end>Sat Dec 31 01:00:00 CET 2011</edm:end>
  </edm:TimeSpan>
</rdf:RDF>
```


Occurrence of known patterns

Because of the above mentioned „loose“ nature of the schemas, most of the problems do not break any formal rules, but a metadata or domain expert can tell you that given field values or constellation of individually correct values constitute issues. We can collect them in a „problem catalog“ and check each one whether these patterns occur in the record. The process returns a row binary outputs, each pattern categorizes the record with a zero or a one.

Timeliness

The metric denotes how the quality changed over the times. One of the main reason of the quality measurement is to give data creators feedbacks about the improvable areas, so we can suppose a data improving circle: ingestion – measurement – improvement. The timeliness shows how efficient is this circle.

Provenance

Everybody knows the basics of PageRank algorithm: the more link points to a given site, the more valuable the pages of the site. Provenance is something parallel. It first groups records by the data creators, then calculates the mean value, and in a final step assigns this value to every individual records. This way the good practices are rewarded. [11]

Quality of contextualization

The good metadata contains two types of information. On one hand it precisely describes its object, and on the other hand it provides contextual information as keywords or vocabulary terms – they denotes subjects of the record, such as personal or place names, subject headings, dates. These contextual elements are normalized terms. For example: If we have a painting from Rembrandt, then the record’s creator field contains Rembrandt, and ideally it also has Rembrandt as a contextual field in a normalized name form (and this piece points to a universal dictionary entry, in for example the Virtual International Authority File [12]). We can measure what is the ratio of the content of free text elements and these contextual fields, and whether each potential candidate (personal and geographical names, subject headings) has a corresponding contextual field, and vice versa: we can collect all contextual field, and check whether the each string occurrence is coupled with a contextual field (every record which has the „Rembrandt“ string in the creator field, has an „agent“ field as well for Rembrandt).

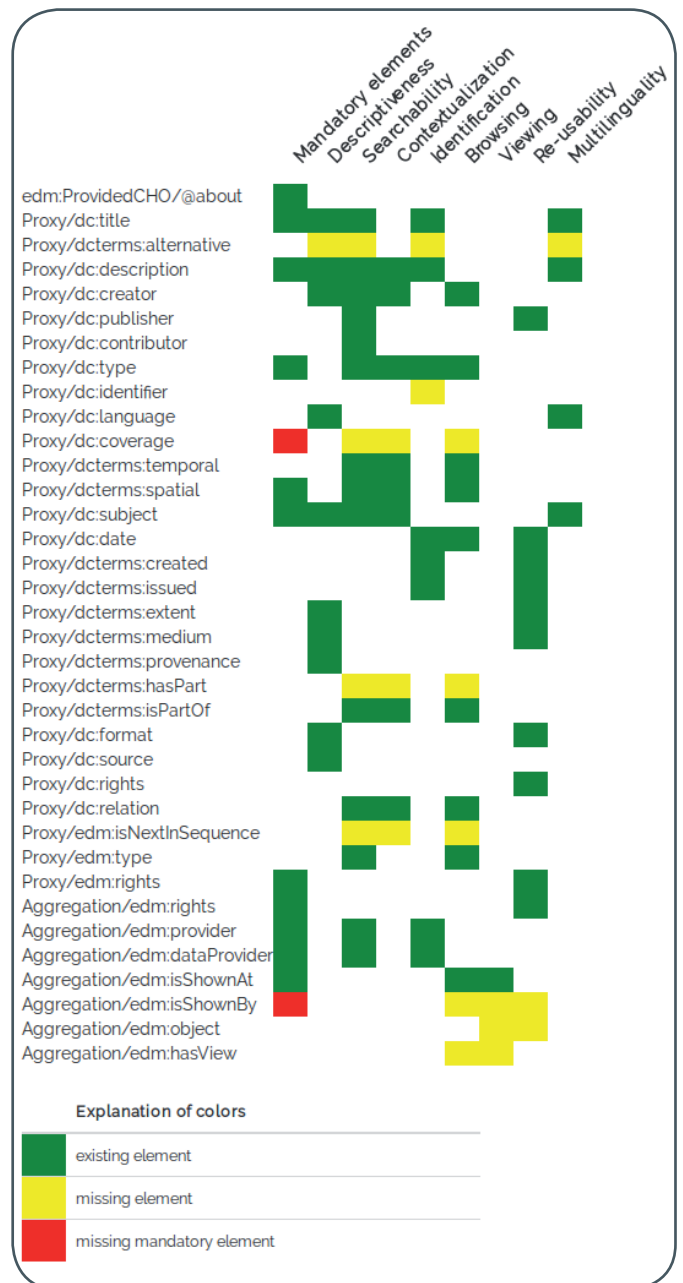
EVALUATION OF THE MEASUREMENTS

If we will have the main metrics we can compare the results with human evaluation. We will ask metadata and domain experts to evaluate selected records. We can use this result set in machine learning algorithms as test set to refine the weightings of the metrics or to exclude those which proved to be irrelevant. We also could find that since our metrics don’t correlate well with the experts’ evaluation we should continue searching for additional – so far not measured – factors.

In either case: The author of this article doesn’t think that we will be able to predict the goodness or badness of metadata in each and every cases. He tends to see the result of the measurements like the „code smell“ [13] – it is not an evidence, but a highly probably symptom of a problem.

APPENDIX 1: AN EXTRACT FROM AN EUROPEANA RECORD

The following sample comes from a Europeana record. We removed some parts of it, to concentrate on those parts which are important for the current topic. The main description takes place in the *ore:Proxy* element. This part is the core of the record, and this is what the data provider submits to Europeana. The schema is an extension of the popular Dublin core standard. It has descriptive elements, such as title and contributor, and contextual elements as well, such as coverage, spatial, subject. Above this core element there are two vocabulary entries (in Europeana slang „contextual entities“): *edm:Place* for a geographical entity, and *edm:TimeSpan* for a time component. Both use an external semantic vocabulary, *geonames.org*, and *semium.org* (the identifiers are URI-s, so there is no guarantee, that they can be resolved as an URL). These contextual elements can be provided by the data creator or they can be the result of a specific step in Europeana’s data processing



Appendix 2: Visualization of Completeness Metric

workflow called „enrichment“, which extracts information from the descriptive elements, and binds them to vocabulary entries.

APPENDIX 2: VISUALIZATION OF COMPLETENESS METRIC

We mentioned that the completeness metric has several sub-dimensions. Here is the first approach for visualizing the existing and missing field. Each row represents a metadata field, and each column represents a sub-dimension. If a field is existing, then the cell is green, otherwise red in the mandatory column, and yellow elsewhere. From the first glance it will be evident which important fields are missing. The metric doesn't cover all the fields, these are the only most important ones.

FOOTNOTES

- [1] <http://europeana.eu>
 [2] <http://pro.europeana.eu/structure/our-network>
 [3] <http://pro.europeana.eu/blogpost/new-metadata-quality-task-force>

- [4] Dangerfield, Marie-Claire et. al (2015): Report and Recommendations from the Task Force on Metadata Quality http://pro.europeana.eu/files/Europeana_Professional/Publications/Metadata%20Quality%20Report.pdf
 [5] The following points are extractions from the Committee's yet unpublished Mission statement document.
 [6] Europeana intensively uses and propagates the Linked Open Data concept. It means that the metadata record includes not just literal values, but concepts defined as vocabulary entries somewhere in the semantic web (such as geographical names, persons, dates). This technique improves the disambiguation and the precision of searches.
 [7] <http://pro.europeana.eu/publication/publishing-framework>
 [8] <http://dp.la/>
 [9] <https://basecamp.com/1768384/projects/11189810>
 [10] <http://pkiraly.github.io/>
 [11] Another positive feedback for the data providers if the site search rank takes into account the result of data quality measurements: If two records has the same relevancy the better quality record will be the first in a search result list.
 [12] <http://viaf.org/>
 [13] https://en.wikipedia.org/wiki/Code_smell ■

Kurz & knapp

Öffnungszeiten des Rechenzentrums um Ostern

Das Rechenzentrum der GWDG ist vom 25.03.2016, Karfreitag, bis zum 28.03.2016, Ostermontag geschlossen.

Falls Sie sich zu der Zeit, in der das Rechenzentrum geschlossen ist, in dringenden Fällen an die GWDG wenden wollen, schicken Sie bitte eine E-Mail an support@gwdg.de. Das dahinter befindliche Ticket-System wird auch während dieser Zeit von Mitarbeiterinnen und Mitarbeitern der GWDG regelmäßig überprüft.

Wir bitten alle Benutzerinnen und Benutzer, sich darauf einzustellen.

Grieger

Zusatztermine für zwei GWDG-Kurse

Kurzfristig wurden folgende zwei neuen Kurstermine in das Kursprogramm des ersten Halbjahres aufgenommen:

- 25. - 26.04.2016: Grundlagen der Bildbearbeitung mit Photoshop
- 03. - 04.05.2016: InDesign – Grundlagen

Nähere Informationen hierzu sind unter <https://www.gwdg.de/kursprogramm> zu finden.

Otto



MS SharePoint

KOLLABORATION LEICHT GEMACHT!

Ihre Anforderung

Sie möchten eine kooperative Kommunikations- und Informationsplattform für Mitarbeiter einrichten, die ständig und von überall verfügbar ist. Sie benötigen ein integriertes Dokumentenmanagementsystem und möchten gemeinsame Besprechungen und Termine planen und verwalten.

Unser Angebot

Wir bieten Ihnen SharePoint als Kollaborationsplattform. Wir können z. B. eine SharePoint Site Collection als gemeinsames Portal für Ihre Arbeitsgruppe oder Ihr Projektteam einrichten. Eine solche Site Collection kann sowohl in Englisch als auch in Deutsch präsentiert werden. Mit einer umfangreichen Auswahl an Schablonen, Apps und Layout-Vorlagen können Sie das Design Ihrer Site Collection anpassen. Der Zugriff erfolgt über GWDG-Benutzerkonten. Weitere Authentifizierungsverfahren sind möglich.

Ihre Vorteile

- > Einheitliches Dokumenten-Managementsystem
- > Umfangreiche Listen und Bibliotheksfunktionen für Dokumente, Bilder oder Dateien

- > Steigern der Produktivität der Mitarbeiter durch vereinfachte tägliche Geschäftsaktivitäten.
- > Einfaches Planen und Protokollieren von Besprechungen
- > Führen nicht öffentlicher Diskussionsrunden
- > Wissensmanagement: Aufbau eines Wikis für Ihre Mitarbeiter
- > Bereitstellung von Informationen und Fachwissen für Mitarbeiter
- > Geringer Entwicklungs- und Pflegeaufwand der SharePoint-Plattform für Benutzer
- > Individuell anpassbares Layout und Design
- > Optimale MS Office-Anbindung
- > Einfache Benutzer- und Gruppenverwaltung

Interessiert?

Der Dienst steht allen Mitgliedern der Max-Planck-Gesellschaft und der Universität Göttingen zur Verfügung. Voraussetzung für die Nutzung ist die Benennung eines Ansprechpartners, der die Administration Ihrer Site Collection übernehmen soll. Wenn Sie SharePoint nutzen möchten, senden Sie bitte eine entsprechende E-Mail an support@gwdg.de. Nähere Informationen zu SharePoint sind auf der u. g. Webseite zu finden.

Dr. SOAP – oder wie ich lernte, den Domino Web-Datenbankzugriff zu lieben

Text und Kontakt:

Thorsten Hindermann
thorsten.hindermann@gwdg.de
0551 201-1837

Der Zugriff auf IBM Notes-Datenbanken kann mittels des standardisierten Web-Service-Protokolls SOAP (Simple Object Access Protocol) erfolgen. In der heutigen Zeit sind mittlerweile viele Entwicklungsumgebungen (engl. Integrated Development Environments, kurz IDE) in der Lage, auf SOAP-basierende Web-Services ohne großen Aufwand abzufragen oder Daten mittels dieser Schnittstellen zu speichern.

AUFSTELLUNG

Um einen Web-Service in einer IBM Notes-Datenbank zu erstellen, steht der *IBM Domino Designer* zur Verfügung. Dieses Programm läuft nur unter dem Betriebssystem Microsoft Windows. Mit Hilfe dieser speziellen Entwicklungsumgebung ist ein Programmierer in der Lage, entsprechende SOAP-Web-Services zu entwickeln. In der aktuellen Domino-Server-Version 9 unterstützt IBM derzeit nur auf SOAP basierende Web-Services. Ob in einer zukünftigen Server-Version dann auch die moderneren REST-Services zur Verfügung stehen, bleibt abzuwarten.

Um mit der Arbeit zu beginnen, ist es notwendig, die entsprechende Datenbank im Designer zu öffnen. In der linken Navigationsansicht sieht der Entwickler nun das virtuelle Dateisystem der Domino-Datenbank.

Im Verzeichnis *Code* -> *Web-Service-Provider* werden ein oder mehrere Web-Service-Infrastruktur(en) angelegt. Für verschiedene Anwendungsfälle kann es sinnvoll sein, dass die SOAP-Daten mit verschiedenen SOAP-Parametern bereitgestellt werden. Im aktuellen Beispiel gibt es eine Web-Service-Infrastruktur für die Zusammenarbeit mit der Java-Entwicklungsumgebung *Netbeans* (*WSINSTITUTEJNB* in der Abb. 1).

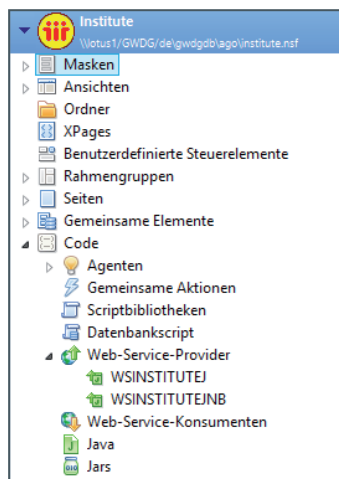


Abb. 1

DIE IBM-ERÖFFNUNG

Als Programmiersprachen stehen aus Gründen der Abwärtskompatibilität *Lotus Script*, eine Basic-Variante, und *Java* zur Verfügung. Die in diesem Artikel gezeigten Quellzeilen-Auflistungen

sind in der von IBM favorisierten Programmiersprache Java entwickelt worden.

Nachdem der Name der Quelldatei umbenannt wurde (siehe Abb. 2), kann die eigentliche Programmierung des Web-Services beginnen. Die Umbenennung ist sinnvoll, da der Name der Quelldatei ein Bestandteil des Zugriffs-URLs auf den Web-Service ist.

Die Absicherung gegenüber unberechtigtem Zugriff wird komplett von der Domino-Datenbank geerbt, in der der Web-Service



Abb. 2

integriert ist. Somit muss sich der Programmierer hier keine weiteren Gedanken um die Sicherheit machen. Dies wird nach den Vorgaben des Datenbankverwalters komplett vom Domino-Server übernommen.

Accessing IBM Notes databases with SOAP

Access to IBM Notes databases can be carried out using the standardized web services protocol SOAP (Simple Object Access Protocol). At the present time many integrated development environments (IDE) will be able to query web services which based on SOAP with little effort or to save data by these interfaces.

```

/**
 * Konstruktor der Klasse. Erste grundlegende Zuweisungen/Initialisierungen werden durchgeführt
 * @author thinder
 */
public WSINSTITUTEJ() {
/**
 * Zugriffsobjekt auf die aktuell Sitzung erstellen
 */
try{
s = WebServiceBase.getCurrentSession();
} catch (Exception e) {
l.toDominoLog(e.toString());
}
/**
 * Zugriffsobjekt auf den aktuellen Inhalt (Context) der Sitzung erstellen
 */
try {
ac = s.getAgentContext();
} catch (NotesException e) {
l.toDominoLog(e.toString());
}
/**
 * Zugriffsobjekt auf die aktuelle Datenbank im aktuellen Inhalt (Context) erstellen.
 */
try {
db = ac.getCurrentDatabase();
} catch (NotesException e) {
l.toDominoLog(e.toString());
}
}
}

```

Auflistung 1

Damit nun der Web-Service erfolgreich die Arbeit aufnehmen kann, muss noch eine richtige Eröffnung der Datenbank für den Web-Service erfolgen. Diese ist in der Auflistung 1 zu sehen.

Vorgabe der Entwicklungsumgebung ist, jede Zugriffszeile in ein `try..catch`-Konstrukt einzubetten, damit im Falle eines Fehlers dieser entsprechend behandelt werden kann.

ÖFFENTLICHKEITSARBEIT...

Die Bereitstellung der eigentlichen Funktionalität, die dann in andere Programme eingebunden werden kann, geschieht über die Deklaration von öffentlichen Methoden. Hierbei kommt der Programmierer sogar ohne die sonst in Java üblichen Annotationen in der Form `@Irgend-Eine-Annotation` aus. Die öffentliche Funktion reicht aus, der Domino-Server kümmert sich bei der Bereitstellung des Web-Services um alles Weitere.

In der Auflistung 2 ist das Auslesen aller Datensätze dieser Domino-Datenbank zu sehen.

Neben den schon gerade beschriebenen `try..catch`-Konstrukten sieht der Leser in dieser Auflistung die Verwendung der Domino API (Application Programmer Interface) zum Zugriff auf die entsprechenden Funktionen der Datenbank; z. B. `v = db.getView(„Kuerzel - alle - intern“)`; , um mit einer Ansicht der Datenbank zu arbeiten, oder

```

/**
 * Die Methode <b>getAllInstituts()</b> gibt alle Institutsnamen aus der Ansicht <b>Kuerzel - alle - intern</b> zurück
 * @return Gibt ein 1-dimensionales Feld vom Typ Zeichenkette zurÃ¼ck, in der Institutsname gespeichert wird.
 * @author thinder
 */
public java.lang.String[] getAllInstitutes() {
/**
 * Zugriff auf die angegebene Ansicht (View) erstellen
 */
try {
v = db.getView(„Kuerzel - alle - intern“);
} catch (NotesException e) {
l.toDominoLog(e.toString());
}
/**
 * Initialisierung der Dimension des Feld vom Typ Zeichenkette
 */
try {
straryAllInstitutes = new String[v.getEntryCount()];
} catch (NotesException e) {
l.toDominoLog(e.toString());
}
/**
 * Zugriffsdurchlauf (Iteration) auf die alle einzelnen Dokumente der Ansicht. Dem aktuellen Dokument wird dann der Institutsname entnommen und in das Feld vom Typ Zeichenkette gespeichert.
 */
try {
for(int i = 0; i < v.getEntryCount(); i++) {
d = v.getNthDocument(i + 1); // In Lotus sind Felder nicht 0, sondern 1, beginnend aufgebaut.
if (d.hasItem(„Inst_Name_1“) ) // Prüfen, ob die angegebene Spalte in dem Dokument existiert
straryAllInstitutes[i] = d.getItemValueString(„Inst_Name_1“);
} else {
straryAllInstitutes[0] = „Es kann kein Institut zurückgegeben werden, da das Item Institut nicht existiert.“;
}
}
} catch (NotesException e) {
l.toDominoLog(e.toString());
straryAllInstitutes[0] += util.newline + e.toString(); // Nicht nur in der Log-Datei des Domino-Server soll der Fehler ausgegeben werden.
}
return straryAllInstitutes; // Das Ergebnis der Methode zurückgeben.
}
}

```

Auflistung 2

`d = v.getNthDocument(i + 1)`; für den Zugriff auf das nächste Dokument. Im Fall einer Domino-Datenbank bezeichnet Dokument das Äquivalent zu Datensatz in anderen Datenbanksystemen.

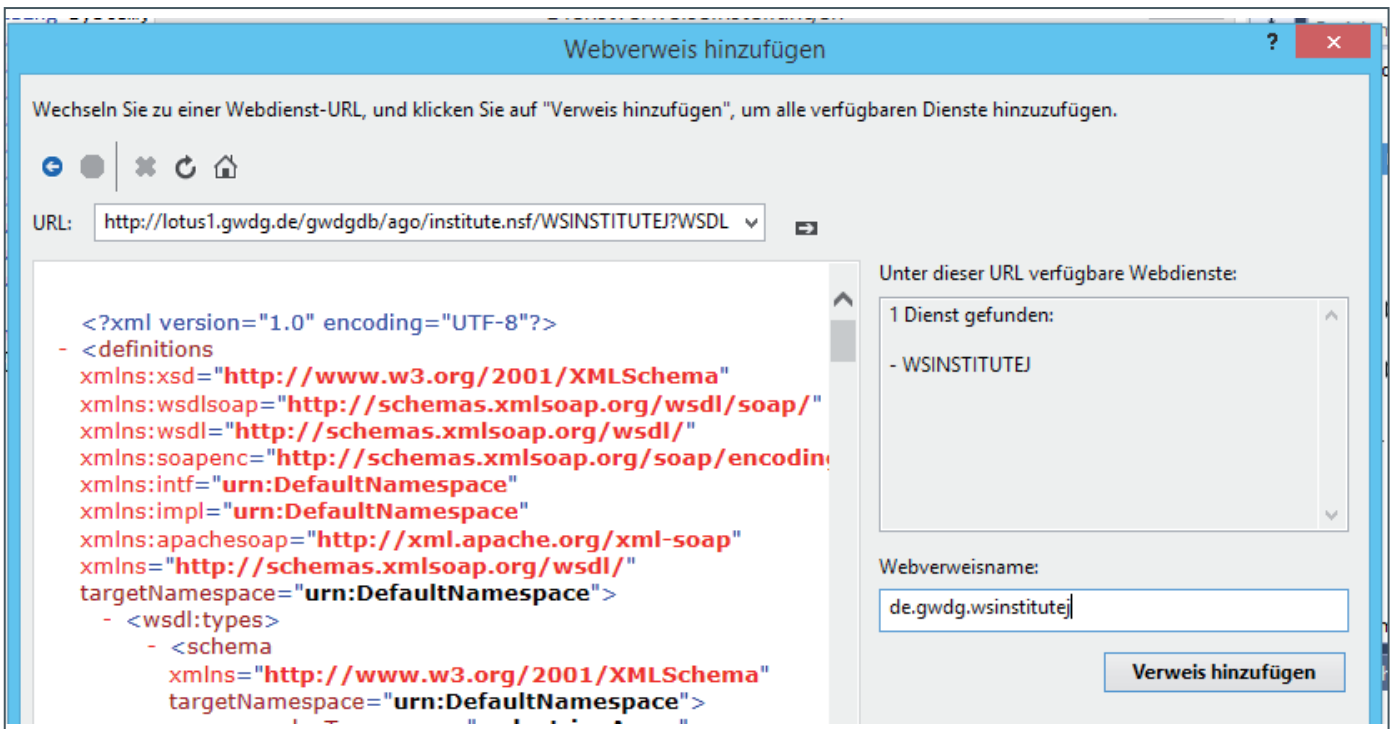


Abb. 3

...LOHNT SICH!

Nachdem auf diese Weise nun alle Methoden für den Web-Service erstellt und programmiert worden sind, sieht das Ergebnis dieses SOAP-Web-Services wie in Auflistung 3 aus.

```
Domino(wsdl)
  getAllInstitutes
  getAllInstitutes
  getAllInstitutesMPG
  getAllInstitutesUNI
  getAllInstitutesGWD
  getAllInstitutesZFR
  getInstituteLongName
  getInstituteInfos
  getAnsprechpartnerSperrEMail
```

Auflistung 3

Für die Verwendung des Web-Services in anderen IDEs und Programmiersprachen ist die WSDL (Web-Service Definition Language, <http://www.w3.org/TR/wsdl>) viel interessanter und wichtiger. Meistens muss der URL mit dem Parameter der WSDL in der entsprechenden IDE angegeben werden, damit diese dann daraus automatisch die entsprechenden Zugriffsmethoden erzeugen kann. Auflistung 4 zeigt einen Ausschnitt dieser WSDL-Ansicht.

Das Ergebnis der Einbindung in eine andere Entwicklungsumgebung, im gezeigten Fall Visual Studio 2015, sieht wie in Abb. 3 aus. Im Endergebnis wandelt Visual Studio die Informationen des WSDL-URL in entsprechende C#- oder Visual-Basic.Net-Programmier-Klassen und -Methoden um. Die Einbindung dieser Klassen und Methoden ist in Abb. 4 und Auflistung 5 zu sehen.

Somit ist am Ende des Artikels ersichtlich, dass mit definierten

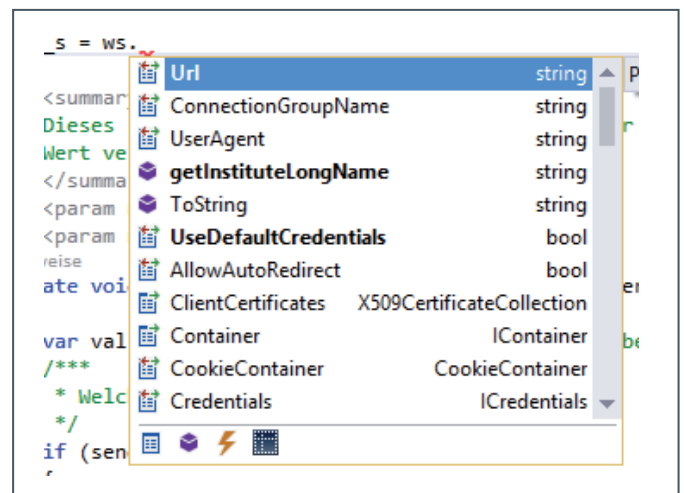


Abb. 4

```
// Zuweisung des Web-Service-Zugriff an eine Variable private
de.gwdg.wsinstitutej.WSINSTITUTEJService ws = new
de.gwdg.wsinstitutej.WSINSTITUTEJService();
// Zeichenketten-Variable für das Ergebnis des Aufruf der
Web-Service-Methode
private String _s;
// Zuweisung des Ergebnis vom Web-Service-Methodenaufruf
_s = ws.getInstituteLongName(„GWDG“);
```

Auflistung 5

Schnittstellen Daten aus Domino-Datenbanken recht einfach von anderen Anwendungen konsumiert werden können, seien dies Desktop- oder mobile Anwendungen. ●


```
<?xml version="1.0" encoding="UTF-8"?>

-<definitions xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:wsd:soap="http://schemas.xmlsoap.org/wsdl/soap/"
xmlns:wsd="http://schemas.xmlsoap.org/wsdl/"
xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
xmlns:intf="urn:DefaultNamespace"
xmlns:impl="urn:DefaultNamespace"
xmlns:apache:="http://xml.apache.org/xml-soap"
xmlns="http://schemas.xmlsoap.org/wsdl/"
targetNamespace="urn:DefaultNamespace">

-<wsdl:types>

-<schema xmlns="http://www.w3.org/2001/XMLSchema"
targetNamespace="urn:DefaultNamespace">
-<complexType name="xsd_stringArray">
-<sequence>
<element name="item" type="xsd:string"
minOccurs="0"
maxOccurs="unbounded"/>
</sequence>
</complexType>
```

Auflistung 4 (Teil 1)

```
<element name="getAllInstitutesReturn"
type="impl:xsd_stringArray"/>
<element name="gwdgGruppierung" type="xsd:string"/>
<element name="getAllInstitutesMPGReturn"
type="impl:xsd_stringArray"/>
<element name="getAllInstitutesUNIReturn"
type="impl:xsd_stringArray"/>
<element name="getAllInstitutesGWDReturn"
type="impl:xsd_stringArray"/>
<element name="getAllInstitutesZFRReturn"
type="impl:xsd_stringArray"/>
<element name="gwdgKuerzel" type="xsd:string"/>
<element name="getInstituteLongNameReturn"
type="xsd:string"/>
<element name="getInstituteInfosReturn"
type="impl:xsd_stringArray"/>
<element name="getAnsprechpartnerSperrEMailReturn"
type="impl:xsd_stringArray"/>
</schema>
</wsdl:types>
```

Auflistung 4 (Teil 2)



Software und Lizenzverwaltung

Der einfache Weg zur Software!

Ihre Anforderung

Sie benötigen eine Software, für die es keine von Ihnen nutzbare Rahmenvereinbarung gibt. Die Anzahl der erforderlichen Lizenzen ist nicht genau festgelegt.

Unser Angebot

Wir verfügen über eine Reihe von Rahmen- und Campusvereinbarungen mit namhaften Softwareherstellern und -lieferanten, über die Software auch in geringerer Stückzahl bezogen werden kann. Wir wickeln für Sie die Beschaffung der erforderlichen Lizenzen ab. Wir können uns bei Vertragsverhandlungen und Bedarfsanalysen engagieren. Zugriffslizenzen können auch über Lizenzserver verwaltet werden.

Ihre Vorteile

> Sie können die benötigte Software in vielen Fällen sofort nutzen.

- > Sie brauchen kein eigenes Ausschreibungs- und Beschaffungsverfahren durchzuführen.
- > Sie ersparen sich die zeitraubenden Verhandlungen mit den Softwareherstellern und -lieferanten.
- > Die Anzahl der benötigten Lizenzen wird Ihnen flexibel zur Verfügung gestellt.
- > Wir können die Nachfrage von verschiedenen Nutzern für neue Lizenzvereinbarungen bündeln.

Interessiert?

Informationen zu bestehenden Lizenzvereinbarungen sind auf der u. g. GWDG-Webseite zu finden. Falls Sie nach spezieller Software suchen, die noch nicht auf unserer Webseite erwähnt ist, kommen Sie bitte auf uns zu. Wir werden prüfen, ob wir eine Vereinbarung abschließen können und bündeln die Nachfrage mit anderen Nutzern.

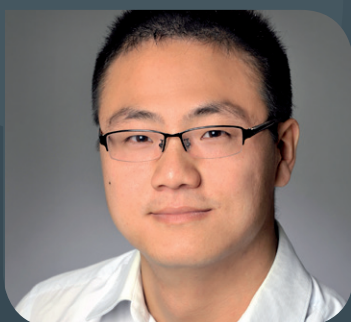
Personalia

NEUE MITARBEITER

SEBASTIAN POHL UND JONAH WINDOLPH

Seit dem 1. Februar 2016 verstärken Herr Sebastian Pohl und Herr Jonah Windolph als studentische Hilfskräfte das Support-Team der GWDG. Schwerpunkte ihrer Arbeitszeit werden die Abendstunden und Wochenenden sein, um dann Anrufe oder Anfragen per Ticket-System von ratsuchenden Nutzern der GWDG entgegenzunehmen und zu bearbeiten. Herr Pohl und Herr Windolph studieren beide an der Georg-August-Universität Göttingen; Herr Pohl im Fach Wirtschaftsinformatik und Herr Windolph im Fach Angewandte Informatik. Wir wünschen beiden eine erfolgreiche Tätigkeit bei der GWDG und freuen uns über die Verstärkung im Support-Bereich.

Heuer



NEUER STIPENDIAT FEI ZHANG

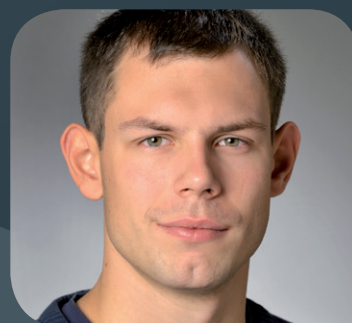
Seit dem 1. November 2015 ist Herr Fei Zhang als Stipendiat in der Arbeitsgruppe „eScience“ (AG E) tätig und arbeitet im Bereich „Cloud Computing“. Herr Zhang bringt einen Bachelor-Abschluss in Netzwerktechnik (2011) und einen Master-Abschluss in Informatik (2013) der National University of Defense Technology, Changsha, Hunan, China mit. Von 2012 bis 2013 absolvierte er ein einjähriges Praktikum am National Supercomputer Center in Tianjin, China. Seit September 2014 ist er Doktorand bei Professor Ramin Yahyapour an der Georg-August-Universität Göttingen. Sein Dissertationsthema lautet „VM Migration over WAN“. Herr Zhang bringt sein Expertenwissen aus dem Bereich Cloud Computing und hier insbesondere der Migration Virtueller Maschinen (VM) in die AG E ein. Herr Zhang ist per E-Mail unter fei.zhang@gwdg.de und telefonisch unter 0551 201-2195 erreichbar.

Wieder

AUSBILDUNG ERFOLGREICH ABGESCHLOSSEN EDUARD DREGER

Mit der Präsentation seines Abschlussprojekts „Implementierung einer REST-Schnittstelle für die UNI-IT der Universität Göttingen zur externen Nutzung eines mTAN-Verfahrens des GWDG-Kundenportals“ bei der IHK hat Herr Eduard Dreger am 26. Januar 2016 seine zweieinhalbjährige Ausbildung zum Fachinformatiker Anwendungsentwicklung erfolgreich abgeschlossen. Wie üblich wird er zunächst für sechs Monate bei der GWDG weiterbeschäftigt. Während dieser Zeit wird er hauptsächlich an der Weiterentwicklung des Kundenportals mitwirken. Wir gratulieren Herrn Dreger ganz herzlich zur bestandenen Prüfung und wünschen ihm alles Gute für seinen Einstieg ins Berufsleben.

Grieger





Using the Parallel Processing Power of the GWWDG Scientific Compute Cluster

Upcoming Introductory and Parallel Programming Courses

GWWDG operates a scientific compute cluster with currently 14,800 cores and a total compute power of 200 Teraflops (2.0×10^{14} floating point operations per second), which can be used by all scientists of the institutes of GWWDG's supporting organisations, University of Göttingen and Max Planck Society.

In order to facilitate the access to and the efficient use of these computing resources, GWWDG offers introductory and parallel programming courses, held at GWWDG's site 'Am Faßberg'.

The next courses in 2016 are

> April 18th, 9:30 am - 4:00 pm

Using the GWWDG Scientific Compute Clusters – an Introduction

This course explains all steps for accessing GWWDG's clusters, to compile and install software, and to work with the batch system for the execution of application jobs. The course is intended for new or inexperienced users of the clusters.

> April 19th - 20th, 9:15 am - 5:00 pm

Parallel Programming with MPI (Including MPI for Python)

This course introduces the message passing interface (MPI) for programming parallel applica-

tions in FORTRAN, C, and in Python. All concepts will be illustrated with hands on exercises. Examples of parallel applications will be presented and analysed.

> May 10th, 9:15 am - 4:30 pm

High-level, High-performance Technical Computing with Julia

Julia is a modern programming language combining high-level dynamic programming with high performance. The course covers the basics of Julia including numerical computing, parallel computing, and statistical methods.

These three courses are repeated regularly. Other courses on parallel computing, dealing with more specialized topics can be arranged on demand. The possible subjects include parallel programming for shared memory systems and for graphics processors, and using extensions of C or Fortran with high level parallel constructs.

More Information about the courses held regularly or on demand at www.gwdg.de/scientific-computing-courses.

Information for registering for the courses at www.gwdg.de/courses.

If you have any further questions please contact support@gwdg.de.

INFORMATIONEN:
support@gwdg.de
0551 201-1523

März bis
Dezember 2016

Kurse



KURS	VORTRAGENDE/R	TERMIN	ANMELDEN BIS	AE
EINFÜHRUNG IN WINDOWS 10	Buck	03.03.2016 9:00 – 12:30 Uhr	25.02.2016	2
DIE SHAREPOINT-UMGEBUNG DER GWGD	Buck	10.03.2016 9:00 – 12:30 und 13:30 – 15:30 Uhr	03.03.2016	4
PHOTOSHOP FÜR FORTGESCHRITTENE	Töpfer	15.03. – 16.03.2016 9:30 – 16:00 Uhr	08.03.2016	8
INDESIGN – AUFBAUKURS	Töpfer	05.04. – 06.04.2016 9:30 – 16:00 Uhr	29.03.2016	8
MAC OS X IM WISSENSCHAFTLICHEN ALLTAG	Bartels	13.04. – 14.04.2016 9:30 – 16:30 Uhr	06.04.2016	8
USING THE GWGD SCIENTIFIC COMPUTE CLUSTER – AN INTRODUCTION	Dr. Boehme, Ehlers	18.04.2016 9:30 – 16:00 Uhr	11.04.2016	4
PARALLELRECHNERPROGRAMMIERUNG MIT MPI	Prof. Haan	19.04. – 20.04.2016 9:15 – 17:00 Uhr	12.04.2016	8
DIE SHAREPOINT-UMGEBUNG DER GWGD	Buck	21.04.2016 9:00 – 12:30 und 13:30 – 15:30 Uhr	14.04.2016	4
GRUNDLAGEN DER BILDBEARBEITUNG MIT PHOTOSHOP	Töpfer	25.04. – 26.04.2016 9:30 – 16:00 Uhr	18.04.2016	8
EINFÜHRUNG IN DIE STATISTISCHE DATENANALYSE MIT SPSS	Cordes	27.04. – 28.04.2016 9:00 – 12:00 und 13:00 – 15:30 Uhr	20.04.2016	8
INDESIGN - GRUNDLAGEN	Töpfer	03.05. – 04.05.2016 9:30 – 16:00 Uhr	26.04.2016	8

KURS	VORTRAGENDE/R	TERMIN	ANMELDEN BIS	AE
HIGH-LEVEL, HIGH-PERFORMANCE TECHNICAL COMPUTING WITH JULIA	Chronz	10.05.2016 9:15 – 16:30 Uhr	03.05.2016	4
ADMINISTRATION VON PCS IM ACTIVE DIRECTORY DER GWDG	Buck	12.05.2016 9:00 – 12:30 und 13:30 – 15:30 Uhr	05.05.2016	4
ANGEWANDTE STATISTIK MIT SPSS FÜR NUTZER MIT VORKENNTNISSEN	Cordes	18.05. – 19.05.2016 9:00 – 12:00 und 13:00 – 15:30 Uhr	11.05.2016	8
EINFÜHRUNG IN DAS IP-ADRESSMANAGEMENTSYSTEM DER GWDG FÜR NETZWERKBEAUFTRAGTE	Dr. Beck	24.05.2016 10:00 – 12:00 Uhr	17.05.2016	2
DIE SHAREPOINT-UMGEBUNG DER GWDG	Buck	26.05.2016 9:00 – 12:30 und 13:30 – 15:30 Uhr	19.05.2016	4
UNIX FÜR FORTGESCHRITTENE	Dr. Sippel	30.05. – 01.06.2016 9:15 – 12:00 und 13:15 – 15:30 Uhr	23.05.2016	12
OUTLOOK – E-MAIL UND GROUPWARE	Helmvoigt	02.06.2016 9:15 – 12:00 und 13:00 – 16:00 Uhr	26.05.2016	4
QUICKSTARTING R: EINE ANWENDUNGSORIENTIERTE EINFÜHRUNG IN DAS STATISTIKPAKET R	Cordes	08.06. – 09.06.2016 9:00 – 12:00 und 13:00 – 15:30 Uhr	01.06.2016	8
DIE SHAREPOINT-UMGEBUNG DER GWDG	Buck	16.06.2016 9:00 – 12:30 und 13:30 – 15:30 Uhr	09.06.2016	4
DATENSCHUTZ - VERARBEITUNG PERSONENBEZOGENER DATEN AUF DEN RECHENANLAGEN DER GWDG	Dr. Grieger	22.06.2016 9:00 – 12:00 Uhr	15.06.2016	2
STATISTIK MIT R FÜR TEILNEHMER MIT VORKENNTNISSEN – VON DER ANALYSE ZUM BERICHT	Cordes	06.07. – 07.07.2016 9:00 – 12:00 und 13:00 – 15:30 Uhr	29.06.2016	8
HIGH-LEVEL, HIGH-PERFORMANCE TECHNICAL COMPUTING WITH JULIA	Chronz	01.09.2016 9:15 – 16:30 Uhr	25.08.2016	4
GRUNDLAGEN DER BILDBEARBEITUNG MIT PHOTOSHOP	Töpfer	06.09. – 07.09.2016 9:30 – 16:00 Uhr	29.08.2016	8
INSTALLATION UND ADMINISTRATION VON WINDOWS 10	Buck	15.09.2016 9:00 – 12:30 und 13:30 – 15:30 Uhr	08.09.2016	4
INDESIGN – GRUNDLAGEN	Töpfer	20.09. – 21.09.2016 9:30 – 16:00 Uhr	13.09.2016	8
DIE SHAREPOINT-UMGEBUNG DER GWDG	Buck	29.09.2016 9:00 – 12:30 und 13:30 – 15:30 Uhr	22.09.2016	4
PHOTOSHOP FÜR FORTGESCHRITTENE	Töpfer	19.10. – 20.10.2016 9:30 – 16:00 Uhr	12.10.2016	8

KURS	VORTRAGENDE/R	TERMIN	ANMELDEN BIS	AE
MAC OS X IM WISSENSCHAFTLICHEN ALLTAG	Bartels	26.10. – 27.10.2016 9:30 – 16:30 Uhr	19.10.2016	8
INDESIGN – AUFBAUKURS	Töpfer	02.11. – 03.11.2016 9:30 – 16:00 Uhr	26.10.2016	8
EINFÜHRUNG IN DIE STATISTISCHE DATENANALYSE MIT SPSS	Cordes	08.11. – 09.11.2016 9:00 – 12:00 und 13:00 – 15:30 Uhr	01.11.2016	8
EINFÜHRUNG IN DAS IP-ADRESSMANAGEMENTSYSTEM DER GWDG FÜR NETZWERKBEAUFTRAGTE	Dr. Beck	15.11.2016 10:00 – 12:00 Uhr	08.11.2016	2
ADMINISTRATION VON PCS IM ACTIVE DIRECTORY DER GWDG	Buck	17.11.2016 9:00 – 12:30 und 13:30 – 15:30 Uhr	10.11.2016	4
QUICKSTARTING R: EINE ANWENDUNGSORIENTIERTE EINFÜHRUNG IN DAS STATISTIKPAKET R	Cordes	23.11. – 24.11.2016 9:00 – 12:00 und 13:00 – 15:30 Uhr	16.11.2016	8
UNIX FÜR FORTGESCHRITTENE	Dr. Sippel	28.11. – 30.11.2016 9:15 – 12:00 und 13:15 – 15:30 Uhr	21.11.2016	12
OUTLOOK – E-MAIL UND GROUPWARE	Helmvoigt	01.12.2016 9:15 – 12:00 und 13:00 – 16:00 Uhr	24.11.2016	4
HIGH-LEVEL, HIGH-PERFORMANCE TECHNICAL COMPUTING WITH JULIA	Chronz	06.12.2016 9:15 – 16:30 Uhr	29.11.2016	4
ANGEWANDTE STATISTIK MIT SPSS FÜR NUTZER MIT VORKENNTNISSEN	Cordes	07.12. – 08.12.2016 9:00 – 12:00 und 13:00 – 15:30 Uhr	30.11.2016	8
DIE SHAREPOINT-UMGEBUNG DER GWDG	Buck	15.12.2016 9:00 – 12:30 und 13:30 – 15:30 Uhr	08.12.2016	4

Teilnehmerkreis

Das Kursangebot der GWDG richtet sich an alle Mitarbeiterinnen und Mitarbeiter aus den Instituten der Universität Göttingen und der Max-Planck-Gesellschaft sowie aus einigen anderen wissenschaftlichen Einrichtungen.

Anmeldung

Anmeldungen können schriftlich per Brief oder per Fax unter der Nummer 0551 201-2150 an die GWDG, Postfach 2841, 37018 Göttingen oder per E-Mail an die Adresse support@gwdg.de erfolgen. Für die schriftliche Anmeldung steht unter <https://www.gwdg.de/antragsformulare> ein Formular zur Verfügung. Telefonische Anmeldungen können leider nicht angenommen werden.

Kosten bzw. Gebühren

Unsere Kurse werden wie die meisten anderen Leistungen der GWDG in Arbeitseinheiten (AE) vom jeweiligen Institutskontingent abgerechnet. Für die Institute der Universität Göttingen und der Max-Planck-Gesellschaft erfolgt keine Abrechnung in EUR.

Absage

Sie können bis zu acht Tagen vor Kursbeginn per E-Mail an support@gwdg.de oder telefonisch unter 0551 201-1523 absagen. Bei späteren Absagen werden allerdings die für die Kurse berechneten AE vom jeweiligen Institutskontingent abgebucht.

Kursorte

Alle Kurse finden im Kursraum oder Vortragsraum der GWDG statt. Die Wegbeschreibung zur GWDG sowie der Lageplan sind unter <https://www.gwdg.de/lageplan> zu finden.

Kurstermine

Die genauen Kurstermine und -zeiten sowie aktuelle kurzfristige Informationen zu den Kursen, insbesondere zu freien Plätzen, sind unter <https://www.gwdg.de/kursprogramm> zu finden.



Gesellschaft für wissenschaftliche
Datenverarbeitung mbH Göttingen