

GWDG  
**NACHRICHTEN**  
08|13

Next Generation  
Sequencing

Softwarelösungen zur  
Plagiatsprävention

Datenmanagement und  
virtuelle Forschungsum-  
gebungen

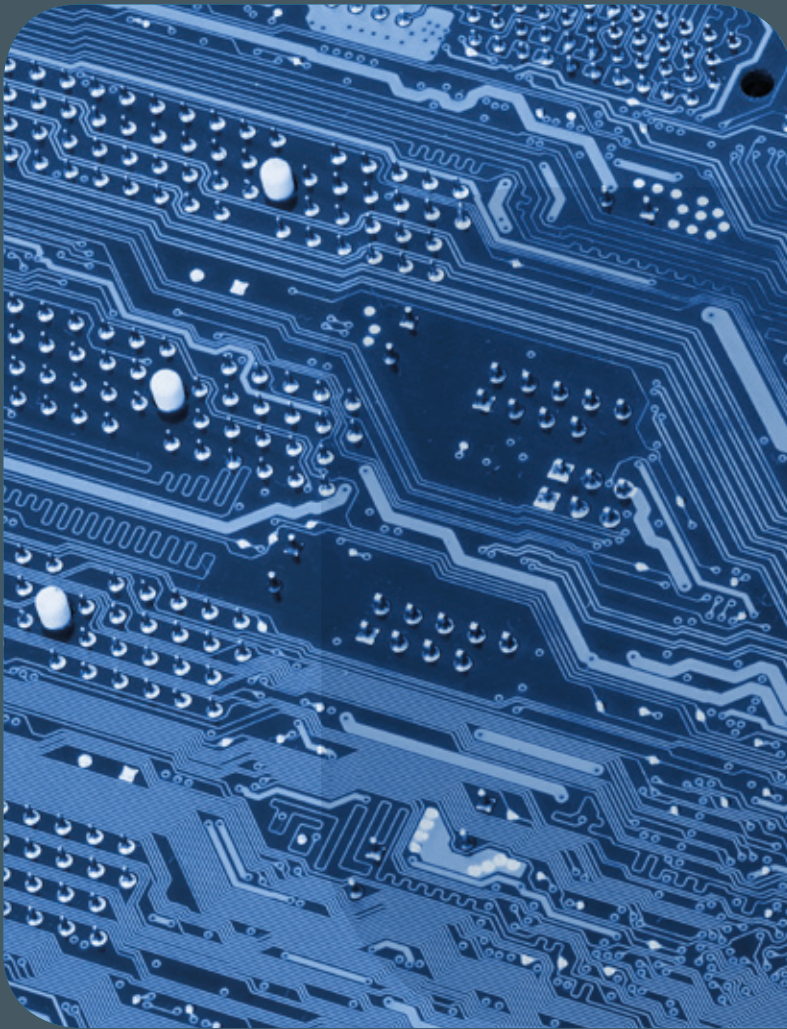
IdM als Dienstleistung

Android und die Sicherheit

ZEITSCHRIFT FÜR DIE KUNDEN DER GWDG

COPY

PASTE



## GWDG NACHRICHTEN

# 08|13 Inhalt

.....

**4 Next Generation Sequencing – Schnipseljagd im Genom** **6 Tipps & Tricks** **7 Test von Softwarelösungen zur Plagiatsprävention** **8 Forschungsprojekte der GWDG – Teil 2: Datenmanagement und virtuelle Forschungsumgebungen** **11 Identity Management als Dienstleistung** **16 Android und die Sicherheit** **17 Personalia** **18 Kurse**

### Impressum

.....  
Zeitschrift für die Kunden der GWDG

ISSN 0940-4686  
36. Jahrgang  
Ausgabe 8/2013

**Erscheinungsweise:**  
monatlich

[www.gwdg.de/gwdg-nr](http://www.gwdg.de/gwdg-nr)

**Auflage:**  
500

**Fotos:**  
© pizuttipics - Fotolia.com (1)  
© Tom - Fotolia.com (4)  
© ILMaster - Fotolia.com (8-10)  
© MPLbpc-Medienservice (3, 17)  
GWDG (18)

#### 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](mailto:thomas.otto@gwdg.de)

**Herstellung:**  
Maria Geraci  
E-Mail: [maria.geraci@gwdg.de](mailto:maria.geraci@gwdg.de)

**Druck:**  
GWDG / AG H  
E-Mail: [printservice@gwdg.de](mailto:printservice@gwdg.de)



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

## *Liebe Kunden und Freunde der GWDG,*

*das Abhören von Kommunikation durch die NSA u. a. in Deutschland liefert im Moment Stoff für zahlreiche Schlagzeilen. Einige Kommentare ähneln sich jedoch: Ist diese Überwachung wirklich so überraschend? Haben wir das nicht immer so vermutet? Ja sicher!*

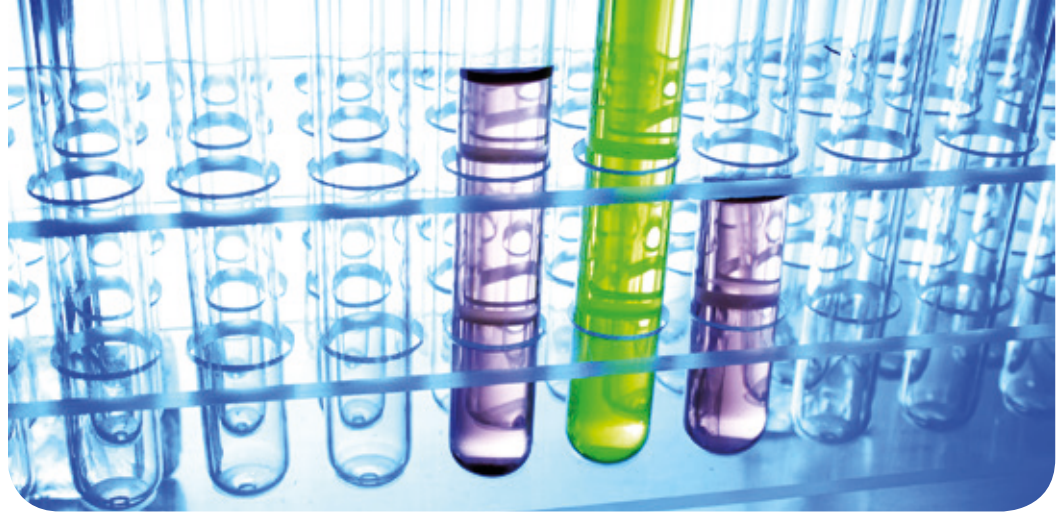
*Dennoch überrascht zum einen das Ausmaß der Datensammlung und zum anderen die offenbar enge Zusammenarbeit der großen Firmen mit der NSA und Gewährung eines direkten Zugriffs. Es stellt sich die Frage sowohl nach der Verhältnismäßigkeit als auch nach der rechtlichen Kontrolle solcher Maßnahmen.*

*Bei einem Termin in Seattle konnte ich vor kurzem einen Datenschutz- und Sicherheitsbeauftragten von einer der aktuell im Kreuzfeuer stehenden Firmen sprechen. Sein Kommentar, dass man die aus Deutschland wiederkehrenden Fragen nach Patriot Act und Einsichtsnahmen durch amerikanische Behörden für überbewertet hält, zeigt die gelebte Doppelmoral. Für eine weitere Verbreitung von Cloud-Angeboten sollte man sich daher genau überlegen, bei wem man seine Daten speichern will und mit wem man diese dann auch teilt. Als Alternative zu kommerziellen Angeboten bietet die GWDG zahlreiche Cloud-Angebote an, wobei die Daten und Systeme in unseren eigenen Rechenzentren verbleiben. Und das ist vielleicht auch gut so.*

*Ich wünsche Ihnen viel Freude beim Lesen dieser Ausgabe der GWDG Nachrichten!*

**Ramin Yahyapour**

*GWDG – IT in der Wissenschaft*



# Next Generation Sequencing – Schnipseljagd im Genom

**Text und Kontakt:**  
Dr. Rainer Bohrer  
rainer.bohrer@gwdg.de  
0551 201-1829

Für die statistischen Rechnungen der Gensequenzanalyse mit Methoden des Next Generation Sequencing (NGS) steht jetzt eine Palette von Programmen auf unserem neuen Hochleistungsrechencluster im Batchbetrieb (max. 64 Cores und 256 GB RAM pro Slot) zur Verfügung: ABySS, BEAST, Bowtie2, BWA, cufflinks, GATK, GenomeTools, Mothur, MrBayes, NCBI-BLAST, Oases, R, SAMtools, SINA, SOAP2-aligner, SOAP2-denovo, SOAP2-denovo-trans, STAR, tophat, Velvet. Weiterhin besteht die Möglichkeit, mit Hilfe eines GWDG-Skripts BLAST-Analysen - auch mit mehreren Millionen Sequenzen - unter Nutzung von hunderten von Batchslots - durchzuführen.

## EINE ATEMBERAUBENDE ENTWICKLUNG

Anfang dieses Jahrtausends (2001/2003) feierte man die „Entschlüsselung“ des menschlichen Genoms nach über 10-jähriger Arbeit von mehr als 1.000 Wissenschaftlern mit Hilfe der Shotgun-Methode. Bei dieser Methode werden kurze DNA-Stränge von bis zu 800 Basen Länge mit klassischen Methoden der Bioinformatik so überlappend aneinandergesetzt, dass man den kompletten DNA-Strang eines Genoms erhält. Die neue Generation der Sequenzierungsmaschinen (Next Generation Sequencing, NGS) hat sich aber so stark weiter entwickelt, dass z. B. das komplette menschliche Genom in wenigen Stunden sequenziert werden kann. Sequenzieren heißt allerdings nur, dass man zwar die Basenabfolge also den Code kennt, ihn aber deshalb noch lange nicht versteht. Erst der Vergleich von Sequenzabschnitten verwandter Organismen, bei denen man die Aufgabe der Sequenzen zu kennen glaubt, legt dann Interpretationen für die Funktion im neu untersuchten Organismus nahe.

## DIE SACHE HAT EINEN HAKEN

Doch leider haben die neuen Sequenzierungsmethoden des NGS einen entscheidenden Nachteil: sie liefern nur extrem kurze Basensequenzen, je nach Methode zwischen 50-300 Bausteinen. Diese sind viel zu kurz, um sie nach der oben beschriebenen klassischen Methode aufgrund gleichartiger Sequenzabschnitte so zu

sortieren, dass letztlich ein kompletter DNA-Strang (*contig*) entsteht. So schön schnell und relativ billig diese Geräte sind, eine Analyse kostet nur 150-250 €, so liefern sie doch sehr große Datenmengen, die erst mit sehr ausgefeilten Statistiken so aufbereitet werden müssen, das letztlich doch wieder ein einziger DNA-Strang (*contig*) rekonstruiert werden kann. Für die statistische Bearbeitung dieser großen Datenmengen (pro Analyse 100-400 GByte!) sind zum einen sehr große Massenspeicher (HDD oder SSD) und zum anderen sehr große Rechenleistungen (20-30 Cores) mit bis zu 1-2 Terabyte (!) RAM erforderlich.

## NextGenerationSequencing – hunting in the genome

A set of 20 bioinformatic programs used for statistical calculations in research with methods of Next Generation Sequencing (NGS) are now available at our new HPC based batch cluster (max. 64 Cores and 256 GB RAM per slot) as there are: ABySS, BEAST, Bowtie2, BWA, cufflinks, GATK, GenomeTools, Mothur, MrBayes, NCBI-BLAST, Oases, R, SAMtools, SINA, SOAP2-aligner, SOAP2-denovo, SOAP2-denovo-trans, STAR, tophat, Velvet. Furthermore it is possible to run massive BLAST analysis (up to millions of sequences) using hundreds of batch slots of this HPC system controlled by a GWDG script.

Da diese Analysemethoden erst seit wenigen Jahren marktreif sind und sich ständig weiterentwickeln, entsteht auch laufend neue Auswertungssoftware, denn die Statistiken müssen sehr genau auf die methodischen Schwächen der einzelnen Sequenzierungstechniken angepasst werden.

Bei der Auswertung geht der aktuelle Trend zum einen dahin, die vom Sequenzierautomat gelieferten Sequenzschnipsel mit verschiedenen statistischen Auswertungsverfahren zu bearbeiten, und zum anderen, unterschiedliche Sequenzierautomaten zu verwenden. So will man die methodischen Schwächen der Sequenzierungstechnik und der unterschiedlichen statistischen Modelle möglichst überwinden. Bioinformatik und klassische Statistik werden dadurch in diesem Wissenschaftsbereich zu einem entscheidenden Faktor.

## BIG DATA - DIE GROSSE HERAUSFORDERUNG

Die bei NGS anfallenden riesigen Datenmengen stellen einzelne Arbeitsgruppen – aber z. T. auch ganze Institute – vor ungeheure Probleme bei der Bereitstellung der erforderlichen Rechenleistung, der schnellen und effektiven Speicherung und erst recht der Langzeitarchivierung. Seit Anfang 2012 wurden deshalb verschiedene Initiativen von Seiten der Wissenschaftler in Göttingen ergriffen – insbesondere durch Prof. Dr. Burkhard Morgenstern (Bioinformatik) und Dr. André Fischer (ENI-G) –, die seit dieser Zeit von der GWDG Unterstützung erfahren haben.

Nun sind wir in der Lage, unseren Kunden eine gute Unterstützung im Bereich NGS anzubieten, da wir den nun schon seit einem Jahr betriebenen Linux-basierten, leistungsstarken Bioinformatik-Server, [gwdu115.gwdg.de](http://gwdu115.gwdg.de) (32 Cores, 128 GByte RAM), mit unserem neuen High-Performance-Cluster ergänzen können. Auf diesem HPC-Cluster stehen einige kommandozeilen-orientierte frei verfügbare Programme aus der NGS-Welt für die Arbeit im Rahmen des Batchsystems zur Verfügung. Dabei kommen 15 Rechenknoten mit je 64 Cores und je 256 GByte RAM zum Einsatz. In naher Zukunft wird es möglich sein, das RAM von bis zu 4 Knoten einem einzigen Knoten zur Verfügung zu stellen, und somit Prozesse mit Speicheranforderungen von bis zu 1 GByte zufrieden zu stellen, wie es bei der de-novo Assemblierung großer Genome oftmals erforderlich ist.

## WELCHE PROGRAMME STEHEN ZUR VERFÜGUNG?

Durch das neue Launching-System unseres HPC-Clusters erfährt man mit der Eingabe von `module avail` auf der Kommandozeile eines der beiden Frontendrechner ([gwdu101.gwdg.de](http://gwdu101.gwdg.de) oder [gwdu102.gwdg.de](http://gwdu102.gwdg.de)), dass z. Zt. die folgenden Programme zur Verfügung stehen: ABySS, BEAST, Bowtie2, BWA, cufflinks, GATK, GenomeTools (incl. readjoiner), Mothur, MrBayes, NCBI-BLAST, Oases, SAMtools, SINA, SOAP2-aligner, SOAP2-denovo, SOAP2-denovo-trans, STAR, tophat, Velvet.

ABySS und das SOAP2-Paket sind zwei der drei Gewinner im „Assemblathon 2010“ der Zeitschrift Nature (<http://www.nature.com/news/2011/110323/full/471425a.html>). ABySS ist für denovo-Transskriptom-Assembling in einer vergleichenden Studie recht gut bewertet worden (gute Resultate bei relativ geringem Verbrauch von Hauptspeicher und Rechenressourcen, siehe <http://www.biomedcentral.com/1471-2105/12/S14/S2>). STAR wird

sehr erfolgreich z. B. an der ETH Zürich eingesetzt, weil es sehr schnell und effektiv ist.

Die einfache Syntax des Launch-Befehls `module` ist hier kurz erläutert:

BEFEHL	ERLÄUTERUNG
<code>module avail</code>	Zeigt alle verfügbaren Module an
<code>module load NCBI-BLAST</code>	Lädt das Module NCBI-BLAST (incl. batch_blast.script)
<code>module load OASES VELVET</code>	Lädt die beiden Module OASES und VELVET
<code>module unload OASES</code>	Löscht das Module OASES, behält aber VELVET
<code>module purge</code>	Löscht alle geladenen Module
<code>module help</code>	Zeigt die Hilfe zu „module“ an
<code>man module</code>	Zeigt die komplette Manpage zu „module“

Das klassische Statistik-Programm *R* steht auch zur Verfügung, muss aber nicht erst mit `module` geladen werden.

Da fast alle Programme ganz neu installiert sind und zum Test spezielle Ausgangsdaten benötigen, die der GWDG meist nicht zur Verfügung stehen, kann nicht ausgeschlossen werden, dass das eine oder andere Programm noch nicht ganz fehlerfrei funktioniert. Wir bitten deshalb um Nachsicht und möglichst schnelle Rückmeldung an uns, damit wir das Problem beheben können.

Weitere inhaltliche Unterstützung an der Universität Göttingen kann man erhalten durch die Core Facility der Medizinischen Statistik. Bitte wenden Sie sich dafür an Dr. Klaus Jung ([Klaus.Jung@ams.med.uni-goettingen.de](mailto:Klaus.Jung@ams.med.uni-goettingen.de)) oder [Sekretariat.AMS@med.uni-goettingen.de](mailto:Sekretariat.AMS@med.uni-goettingen.de).

## BLAST-ANALYSEN „IM GROSSEN STIL“

Für die Nutzung des altbewährten BLAST-Algorithmus „im großen Stil“ steht darüber hinaus ein sehr einfach zu bedienendes Skript (`batch_blast.script`) zur Verfügung, welches über den Befehl `module load NCBI-BLAST` neben dem NCBI-BLAST-Programm-paket zugänglich gemacht wird. Beim ersten Aufruf des Skriptes `batch_blast.script` wird eine Konfigurationsdatei heruntergeladen, die mit einem Linux-Editor (wie z. B. emacs, xemacs, joe, demnächst auch nano) an nur sehr wenigen Stellen angepasst werden muss. Das Konfigurationsfile `bbs.config` ist gut kommentiert und auch Linux-Anwender mit geringen Kenntnissen sollten damit problemlos zu Rande kommen. Aber auch wenn Sie überhaupt noch keine Erfahrung im Umgang mit Linux haben, können wir Ihnen helfen, wenn Sie eine E-Mail an [support@gwdg.de](mailto:support@gwdg.de) schicken (Subject „Bioinformatik“). Das `batch_blast.script` ist theoretisch in der Lage bis zu 8 Mio. Sequenzdateien (im Multiple FastA-Format) mit Hilfe des Batchsystems mit einer Datenbank, wie z. B. NR, abzugleichen. Da jede der 8 Mio. Sequenzdateien mehrere Sequenzen enthalten kann, sollten damit alle denkbaren Anforderungen befriedigt werden können. Bevor man das Skript mit den vielen Sequenzdateien füttert, muss in der Regel die eigene Sequenzdatei in viele kleinere Dateien zerlegt und auf Directories (= Ordner) aufgeteilt werden. Dafür steht ein weiteres Skript als Werkzeug zur Verfügung, das `split_mfa.script`. Dieses zerlegt zuverlässig die Gesamtdatei in kleinere Häppchen und verteilt bis zu jeweils 8.000 Files auf Directories von 000 bis max. 999. Vor der Benutzung der Skripten empfiehlt es sich, die jeweilige Hilfefunktion mit `split_mfa.script` (ohne Parameter) bzw. `batch_blast.script -help` aufzurufen.

Unter [support@gwdg.de](mailto:support@gwdg.de) (Subject: Bioinformatik - NCBI-BLAST) stehen wir Ihnen gerne auch für weitere Fragen zu Verfügung.

## WIE ERHÄLT MAN ZUGRIFF AUF DEN HPC-CLUSTER?

Grundsätzlich kann jeder mit seinem GWDG-Account für das Rechnen auf den verschiedenen HPC-Clustern zugelassen werden. Schicken Sie dafür bitte eine entsprechende E-Mail an [support@gwdg.de](mailto:support@gwdg.de) mit dem Subject „HPC-Cluster“. Da es sich bei den beiden Frontend-Rechnern ([gwdu101.gwdg.de](http://gwdu101.gwdg.de) und [gwdu102.gwdg.de](http://gwdu102.gwdg.de)) um Linuxsysteme handelt, ist für den Zugriff von Macintosh- oder Windows-PCs eine Terminalverbindungs-Software erforderlich, am besten verwendet man dafür einen XWindow-Terminalserver (s. unter: <http://www.gwdg.de/index.php?id=1318>). Beide Frontend-Rechner sind nicht direkt, sondern nur über die Login-Maschine der GWDG [gwdu05.gwdg.de](http://gwdu05.gwdg.de) (oder ersatzweise [gwdu06.gwdg.de](http://gwdu06.gwdg.de)) zugänglich.

## WEITERE BIOINFORMATIK-PROGRAMME

Auf dem weltweit erreichbaren Linux-Server [gwdu115.gwdg.de](http://gwdu115.gwdg.de) stehen darüber hinaus viele weitere Bioinformatik-Programme zur Verfügung: ACT+Artemis, AMPHI, BEAST, BLAST, Chlo-roP, ClustalW, ClustalX, Consed, EMBOSS, FastA, GCG-Package (nur Kommandozeile), HMMER, MIRA, MrBayes, Phrap, Phred,

Phred2FastA, PHYLIP, RECON, SignalP, Staden-Package, STAR, T-Coffee, TargetP, TMHMM, Vienna-RNA-Package.

## GWDG-UNTERSTÜTZUNG BEI KOMMERZIELLEN SEQUENZANALYSEPROGRAMMEN

Weiterhin unterstützt die GWDG verschiedene Institute der Universität und der Max-Planck-Gesellschaft durch die Bereitstellung von leistungsfähigen virtuellen Maschinen mit Spezialsoftware, wie z. B. für CLC Genomics Workbench.

Wir betreiben auch Lizenzserver, wie z. B. für „Geneious Pro“ von Biomatters. Diese in der Universität und in der Max-Planck-Gesellschaft weit verbreitete Software ist zwar als „Geneious Pro“ für NGS-Zwecke nicht geeignet (die NGS-fähige Servervariante „Geneious Server“ haben wir nicht installiert), aber Geneious Pro ist in der täglichen Arbeit mit Sequenzen sehr beliebt, da es relativ einfach zu bedienen und sehr leistungsfähig ist.

Wer Interesse an einem Spezialserver hat oder eigene „Geneious Pro“-Lizenzen erwerben und in unseren Lizenzpool (z. Zt. 49 Lizenzen) einbringen möchte, wende sich bitte an [support@gwdg.de](mailto:support@gwdg.de) mit dem Betreff „Bioinformatik“. Selbstverständlich sind wir offen für Hinweise zur Installation weiterer Programme. Im Übrigen sei darauf hingewiesen, dass sowohl auf dem HPC-Cluster als auch der [gwdu115.gwdg.de](http://gwdu115.gwdg.de) eigene Programme innerhalb des eigenen Speicherbereiches kompiliert und/oder installiert werden dürfen. ■

# Tipps & Tricks

## Microsoft SharePoint 2013: Single Sign-On

Für die Anmeldung auf einer SharePoint 2013 Site Collection werden dieselben Login-Informationen wie für die Anmeldung am Active Directory verwendet. Möchte man mit seinem, bereits am Active-Directory angemeldetem Nutzerkonto auf eine SharePoint-2013-Seite zugreifen, ist es nicht notwendig, sich erneut anzumelden. Stattdessen kann das New-Technology-LAN-Manager-Protokoll (NTLM) verwendet werden. NTLM ist ein von Microsoft entwickelter Standard, der, einmal konfiguriert, die automatische Anmeldung an ausgewählten Systemen übernimmt. Nutzer, die sich am Active Directory anmelden, verwenden diesen Standard bereits automatisch zur Authentifizierung innerhalb des GWDG-Netzwerkes, ohne dass ein Eingreifen des Nutzers notwendig ist. In der Regel gilt das auch für die Anmeldung am SharePoint 2013. Sollte die automatische Anmeldung am SharePoint 2013 jedoch noch nicht funktionieren, wird hier beschrieben, wie die notwendigen Einstellungen manuell vorgenommen werden können.

Das NTLM-Protokoll wird von den meisten Browsern unterstützt. Je nach Browser unterscheidet sich jedoch die Konfiguration. Im Folgenden wird die Konfiguration für ausgewählte gängige Browser beschrieben:

### Microsoft Internet Explorer

Im Internet Explorer befinden sich die Einstellungen unter

„Extras > Internetooptionen > Sicherheit > Lokales Intranet“.

An dieser Stelle müssen zwei Einstellungen vorgenommen werden. Zunächst muss unter „Sites > Erweitert > Hinzufügen“ der URL <https://sharepoint.gwdg.de> zur Liste der als Intranet definierten Webseiten hinzugefügt werden.

Anschließend muss unter „Stufe anpassen > Benutzerauthentifizierung“ der Punkt „Automatisches Anmelden nur in der Intranetzone“ ausgewählt werden. Diese Einstellung ist in der Regel nicht notwendig, da es sich hierbei um den Systemstandard handelt. Lediglich wenn es zu Problemen kommt, sollten diese Einstellungen überprüft werden.

### Google Chrome

Google Chrome verwendet die Windows-Internetooptionen zur Bestimmung der als Intranet definierten Webseiten. Hier müssen also dieselben Einstellungen wie für den Internet Explorer vorgenommen werden.

### Mozilla Firefox

Im Mozilla Firefox kann das Konfigurationsmenü geöffnet werden, indem in der Adresszeile `about:config` eingegeben wird. Anschließend muss der Parameter `network.automatic-ntlm-auth.trusted-uris` durch Doppelklicken editiert werden. Im sich öffnenden Fenster können nun eine Reihe von URLs (hier <https://sharepoint.gwdg.de>) angegeben werden, für die das NTLM-Protokoll verwendet werden soll. Einzelne URLs müssen durch ein Komma getrennt werden.

Lübben

# Test von Softwarelösungen zur Plagiatsprävention

## Text und Kontakt:

Regina Bost  
regina.bost@gwdg.de  
0551 201-1831

Anke Bruns  
anke.bruns@gwdg.de  
0551 201-1519

Roland Groh  
roland.groh@gwdg.de  
0551 201-1838

Vom 16. September 2013 bis zum 15. November 2013 stellt die GWDG im Rahmen einer Testphase allen Interessierten kostenlos die Softwareprodukte Turnitin und Ithenticate zur Verfügung.

„Turnitin“ und „Ithenticate“ sind zwei Softwarelösungen des Herstellers iParadigms (<http://iparadigms.com/>), die es ermöglichen, Texte mit einem Pool anderer wissenschaftlicher Texte zu vergleichen. Dies erleichtert die Erkennung von Plagiaten und ermöglicht so die Überprüfung der Einhaltung von wissenschaftlichen Standards.

Während „Turnitin“ für die Originalitätsprüfung studentischer Arbeiten im Rahmen von Lehrveranstaltungen entwickelt wurde, eignet sich „Ithenticate“ besonders für die Beurteilung von Texten für Forschungseinrichtungen, wissenschaftliche Verlage oder Graduiertenkollegs.

Als Datenbasis für die Originalitätsprüfungen, auf die beide Anwendungen zugreifen können, stehen je nach Bedarf studentische Arbeiten (Global Student Paper Database), Texte wissenschaftlicher Verlage (Publishers' Database) und eine Sammlung mit einem von iParadigms entwickelten Webcrawler erschlossener Internetquellen zur Verfügung. Institutionen können auf Wunsch eigene Textsammlungen anlegen. Neben dem Abgleich mit deutschsprachigen Texten lassen sich mithilfe des integrierten Übersetzungswerkzeugs auch Übereinstimmungen mit englischsprachigen Texten ermitteln.

Die hochgeladenen Arbeiten können dauerhaft in verschlüsselter Form in der Global Student Paper Database oder in eigenen Institutional Repositories gespeichert oder aber nach Fertigstellung des Originalitätsberichts komplett gelöscht werden. Personenbezogene Daten, die im Zuge der Prüfung von Dokumenten erhoben werden, sind nur für die beurteilende Person sichtbar und werden unmittelbar nach der abgeschlossenen Prüfung gelöscht.

„Turnitin“ und „Ithenticate“ werden über eine Webschnittstelle bedient, so dass für die Nutzung außer einem gängigen Webbrowser keine weitere Softwareinstallation benötigt wird. „Turnitin“ entspricht dem Learning Tools Interoperability (LTI) Standard und ist in verschiedene Learning Management Systeme integrierbar. Schnittstellen zu weiteren Systemen können mithilfe eines SDK für Java und PHP entwickelt werden.

Bei Interesse können Sie sich bis zum 5. September 2013 bei der unten angegebenen Kontaktadresse für die Testphase anmelden. Wir werden Sie nach Ablauf der Testphase um Ihre Einschätzung des getesteten Produkts bitten. Bitte melden Sie sich auch, falls Sie grundsätzlich an der Plagiaterkennungssoftware interessiert sind, der Zeitraum für eigene Tests jedoch unpassend ist. So kann der Bedarf an einer derartigen Lösung besser abgeschätzt werden. Sollte die Testphase gut angenommen werden, ist geplant, „Turnitin“ und „Ithenticate“ als Dienste der GWDG unseren Kunden zur Nutzung anzubieten. Dieses Angebot wird allen Einrichtungen der Universität Göttingen zur Verfügung stehen.

Wenn Sie interessiert sind, die Software zu testen, senden Sie eine E-Mail an: [test-plagiatserkennung@gwdg.de](mailto:test-plagiatserkennung@gwdg.de).

## Testing of software solutions for plagiarism prevention

From September, 16 to November, 15 in 2013 the GWDG provides in a test phase the software products Turnitin and Ithenticate to all interested employees of the Max Planck Society and the University Göttingen.

Turnitin and Ithenticate are softwares for comparing texts with a pool of other scientific texts. This facilitates the detection of plagiarism and enables the verification of compliance with scientific standards. While Turnitin is designed to test the originality of student's work in the context of courses, Ithenticate is particularly suitable for the assessment of texts for research institutions, scientific publishers or graduate school.

If you are interested in testing the software, please contact the project team at GWDG under :  
[test-plagiatserkennung@gwdg.de](mailto:test-plagiatserkennung@gwdg.de)

# Forschungsprojekte der GWDG – Teil 2: Datenmanagement und virtuelle Forschungsumgebungen

## Text und Kontakt:

Dr. Ulrich Schwardmann  
ulrich.schwardmann@gwdg.de  
0551 201-1542

Die GWDG unterstützt ihre Kunden nicht nur durch die Bereitstellung von IT-Diensten, sondern betreibt selbst Forschung zu ausgewählten Themen der Angewandten Informatik. Aktuell ist die GWDG hier an einer Reihe von Projekten auf nationaler sowie europäischer Ebene beteiligt, die im Rahmen einer Artikelserie vorgestellt werden. Die ersten beiden Teile stellen die aktuellen Projekte kurz vor, während weitere Artikel ausgewählte Forschungsthemen, Vorhaben und Ergebnisse im Detail beleuchten.

## PERICLES

Das Projekt PERICLES (Promoting and Enhancing Reuse of Information throughout the Content Lifecycle taking account of Evolving Semantics) hat Forschung zu neuen Ansätzen zur Erhaltung digitaler Information zum Ziel, um deren Zugriff und Nutzbarkeit für künftige Generationen sicherzustellen.

PERICLES hat ein besonderes Auge auf die Herausforderungen der digitalen Langzeiterhaltung, die durch hoch komplexes und dynamisches digitales Material entstehen, wie sie sich auf zwei Gebieten finden lassen: auf der einen Seite digitale Kunst, wie zum Beispiel interaktive softwarebasierte Installationen, oder digitale Medien der Sammlungen und Archive der Tate Gallery. Auf der anderen Seite experimentelle wissenschaftliche Daten von der European Space Agency und International Space Station.

## Research projects of the GWDG – Part 2: Data management and virtual research environments

GWDG participates in various national and international research projects dealing with selected topics from the area of applied computer science. These projects can be clustered into three main categories: cloud computing, virtual research environments, and data management. A series of articles will introduce research topics and results, starting with two contributions that briefly introduce all current research project conducted by GWDG.

## Beiträge der GWDG

Die GWDG arbeitet dabei mit der Georg-August-Universität Göttingen als Projektpartner im Arbeitspaket WP7 „Integration and test beds“ zusammen und stellt dort ihre Expertise zum Datenmanagement zum Aufbau einer Projekt-Arbeitsumgebung zur Verfügung.

Pericles – Promoting and Enhancing Reuse of Information throughout the Content Lifecycle taking account of Evolving Semantics

**Themenschwerpunkt:** Digitale Langzeitarchivierung

**Laufzeit:** 02.2013 – 01.2017

**Instrument:** European Commission under the 7th Framework Program

**Partner:** King's College London, United Kingdom, The University of Edinburgh, United Kingdom, Space Applications Services NV, Belgique, Georg-August-Universität Göttingen Stiftung Öffentlichen Rechts, Deutschland, Centre for Research and Technology Hellas, Hellas, The University of Liverpool, United Kingdom, Xerox SAS, France, Institut d'Aeronomie Spatiale de Belgique, Belgique, Hoegskolan i Boras, Sverige, Dotsoft Oloklromenes Efarmoges Diadiktioy kai Vaseon Dedomenom AE, Hellas, The Board of Trustees of the Tate Gallery, United Kingdom

**Webseite:** <http://www.pericles-project.eu/>



## DARIAH-DE

Das DARIAH-Projekt (Digital Research Infrastructure for the Arts and Humanities) hat zum Ziel, Kultur und Geisteswissenschaften durch eine digitale Infrastruktur zu unterstützen. Als Teil der europäischen ESFRI-Initiative arbeiten viele Institute auf nationaler und internationaler Ebene zusammen. Gemeinsam bilden sie das europaweit größte geisteswissenschaftliche Projekt für den Aufbau von Infrastrukturen in diesen Forschungsgebieten. Alleine in Deutschland sind 17 Partner beteiligt. Ein genauerer Überblick wurde in dem Artikel „Forschungsinfrastrukturen für die e-Humanities – Beteiligung der GWDG am Projekt DARIAH-DE“ in den GWDG-Nachrichten 9/2011 gegeben.

### Beiträge der GWDG

Die GWDG arbeitet gemeinsam mit mehreren Rechenzentren u. a. daran, die besonderen Anforderungen an die Verfügbarkeit und Referenzierbarkeit von Forschungsdaten, aber auch an neue computergestützte Forschungsmethoden zu unterstützen. Die Stärken der GWDG liegen hier insbesondere bei der dauerhaften Referenzierbarkeit von Daten mit Hilfe von PIDs (Persistent Identifiers) und der Langzeitverfügbarkeit von Forschungsdaten, damit diese auch in vielen Jahren noch wiederverwendbar sind.

### DARIAH-DE Digital Research Infrastructure for the Arts and Humanities

**Themenschwerpunkt:** Datenmanagement und Persistent Identifier

**Laufzeit:** 01.03.2011 – 28.02.2014

**Instrument:** BMBF

**Partner:** Berlin-Brandenburgische Akademie der Wissenschaften, DAASI International GmbH, Deutsches Archäologisches Institut, Technische Universität Darmstadt - Interdisziplinäre Arbeitsgruppe Digital Humanities, Musikwissenschaftliches Seminar Detmold/Paderborn, Göttingen Centre for Digital Humanities, GWDG, Universität Köln – Historisch-Kulturwissenschaftliche Informationsverarbeitung, Institut für Europäische Geschichte Mainz, Jülich Supercomputing Centre, Karlsruher Institut für Technologie, Otto-Friedrich-Universität Bamberg – Lehrstuhl für Medieninformatik, MPG – Max Planck Digital Library, MPG – Rechenzentrum Garching, Salomon Ludwig Steinheim-Institut für deutsch-jüdische Geschichte an der Universität Duisburg-Essen, Niedersächsische Staats- und Universitätsbibliothek, Julius-Maximilians-Universität Würzburg – Institut für deutsche Philologie – Lehrstuhl für Computerphilologie und Neuere Deutsche Literaturgeschichte

**Webseite:** <https://de.dariah.eu/>

## CLARIN-D

Die GWDG ist ebenso an CLARIN (COMMON LANGUAGE RESOURCE AND TECHNOLOGY INFRASTRUCTURE) beteiligt, dem europäischen Schwesterprojekt von DARIAH. CLARIN konzentriert sich speziell auf die Sprachwissenschaften und bietet noch weitere Herausforderungen an die Datenverarbeitung im Bezug auf Textannotation oder für Werkzeuge für semantische Analysen.

### Beiträge der GWDG

Die GWDG arbeitet bei diesem Projekt im Rahmen eines Unterauftrags mit an der Weiterentwicklung der EPIC API Version 2. Das Projekt wird vom Bundesministerium für Bildung und Forschung (BMBF) im Zeitraum gefördert.

### Clarín - D

**Themenschwerpunkt:** Persistent Identifier

**Laufzeit:** 01.04.2011 – 31.03.2016

**Instrument:** BMBF

**Partner:** Seminar für Sprachwissenschaft, Abt. Computerlinguistik, Tübingen, ASV - Institut für Informatik, Abt. Automatische Sprachverarbeitung, Leipzig, BBAW - Berlin-Brandenburgische Akademie der Wissenschaften, Berlin, IMS - Institut für Maschinelle Sprachverarbeitung, Stuttgart, IDS - Institut für Deutsche Sprache, Mannheim, BAS - Bayerisches Archiv für Sprachsignale, München, The Language Archive, Max Planck Institut für Psycholinguistik, Nijmegen, Hamburger Zentrum für Sprachkorpora, Englische Sprach- und Übersetzungswissenschaft, Saarbrücken.

**Webseite:** <http://de.clarin.eu/>

### INF ADIR: DAS ASTROFIT DATEN-INFRASTRUKTURPROJEKT (TEILPROJEKT SFB 963)

Die GWDG ist am Teilprojekt „INF ADIR: Das AstroFIT Daten-Infrastrukturprojekt“ des Sonderforschungsbereiches (SFB) 963 „Astrophysikalische Strömungsinstabilität und Turbulenz“ an der Georg-August-Universität Göttingen beteiligt. Mit diesem Teilprojekt wird die zentrale Infrastruktur für die Speicherung und das Management der Daten des gesamten SFBs bereitgestellt. Sowohl die halbautomatische Integration, Analyse, Visualisierung und der Austausch der diversen Beobachtungs- und Experimentdaten als auch die Kalibrierung von Simulationsmodellen werden hierbei unterstützt. Auf diese Weise wird generell wie auch speziell die Zusammenarbeit auf dem Gebiet von technischen und methodischen Aspekten von numerischen Simulationen gefördert.

### Beiträge der GWDG

Die GWDG bietet im Rahmen der Projektstätigkeit eine Projektmanagementplattform auf Basis von ChiliProject an (Vorstellung ChiliProject siehe GWDG Nachrichten 6-7/2013). Die Plattform dient zum Austausch von Dokumenten im SFB, der Organisation von Arbeitsgruppen und projektspezifischen Tätigkeiten, sowie der Sammlung und Dokumentation von Wissen durch das integrierte Wiki. Zusätzlich können die Mitglieder des SFBs ihrer Programmquelltext in einer zentralen GIT-Repository-Struktur ablegen und gemeinsame Software-Entwicklung standortübergreifend betreiben. Die Einführung der Projektmanagementplattform und der GIT-Repositories wurde durch Workshops begleitet aus dem INF-Projekt heraus begleitet. Die Plattform zur Langzeitarchivierung und Publizierung der Datensätze des SFBs befindet sich momentan im technischen Aufbau, so dass die GWDG konzeptionell und inhaltlich das Datenmanagement im SFB unterstützt.

## INF ADIR: AstroFIT

### Daten-InfRastrukturprojekt

**Themenschwerpunkt:** Datenmanagement und Infrastruktur

**Laufzeit:** 01.01.2012 - 31.12.2015

**Instrument:** SFB-Projekt der Deutschen Forschungsgemeinschaft (DFG)

**Partner:** Institut für Astrophysik, Göttingen, Institut für Geophysik, Göttingen, Institut für Nicht-Lineare Dynamik, Göttingen, Geowissenschaftliches Zentrum Abteilung Isotopengeologie, Göttingen, Institut für Numerische und Angewandte Mathematik, Göttingen, Institut für Geophysik und extraterrestrische Physik TU Braunschweig, Institut für Theoretische Physik TU Braunschweig, Max-Planck-Institut für Sonnensystemforschung, Göttingen, Max-Planck-Institut für Dynamik und Selbstorganisation, Göttingen, DLR Institut für Aerodynamik und Strömungstechnik, Göttingen, GWDG

**Webseite:** <https://www.uni-goettingen.de/de/215327.html>

## DIGITAL HUMANITIES RESEARCH COLLABORATION (DHFV)

Die Digital Humanities Research Collaboration (DHFV) ist das erste Projekt, das am Göttinger Centre for Digital Humanities (GCDH) angesiedelt ist. Das GCDH ist eine fakultätsübergreifende Einrichtung der Georg-August-Universität Göttingen. Es koordiniert die Forschungs-, Lehre- und Infrastrukturaktivitäten am Göttingen Research Campus im Bereich der Digital Humanities. Die GWDG ist einer der Partner im DHFV-Projekt und dabei zusammen mit der Niedersächsischen Staats- und Universitätsbibliothek Göttingen (SUB) am Teilprojekt 3a „Infrastruktur“ beteiligt. Die Ziele dieses Teilprojektes konzentrieren sich auf zwei Hauptbereiche: zum einen das Konzept einer nachhaltigen und umfassenden Forschungsinfrastruktur und zum anderen das Datenmanagement.

## Beiträge der GWDG

Die GWDG kümmert sich im Teilprojekt 3a zusammen mit der SUB Göttingen um die Konzeption und Zusammenstellung einer nachhaltigen und umfassenden Forschungsinfrastruktur für Digital Humanities. Hierbei werden auf verschiedene Infrastrukturprojekte im Göttinger Research Campus und aus den Portfolios der SUB und GWDG zurückgegriffen, um Synergieeffekte am Standort Göttingen und in den Infrastrukturprojekten wie DARIAH-DE zu entfalten und bestehende Infrastruktur nachzunutzen. Hieraus entstehen Architekturskizzen und ein Dienstekatalog, die auf gemeinsamen Standards, Best-Practices, Schnittstellen und Formaten in den Teilprojekten des DHFV aufsetzen. Zusätzlich beteiligt sich die GWDG an Aufgabenstellungen zur Formulierung von Datenmanagementplänen und Langzeitarchivierungsunternehmungen im DHFV.

## Digital Humanities Research Collaboration (DHFV)

**Themenschwerpunkt:** Forschungsinfrastruktur und Datenmanagement

**Laufzeit:** 01.01.2012 – 31.12.2014

**Instrument:** Niedersächsischen Ministerium für Wissenschaft und Kultur aus dem Programm „Niedersachsen Vorab der Volkswagen-Stiftung“

**Partner:** Geisteswissenschaftliche und Naturwissenschaftliche Institute der Georg-August Universität Göttingen, Niedersächsischen Staats- und Universitätsbibliothek Göttingen (SUB)

**Webseite:** <http://www.gcdh.de/en/projects/dh/>

# Identity Management als Dienstleistung

## Text und Kontakt:

Andreas Ißleiber  
andreas.issleiber@gwdg.de  
0551 201-1815

Ein Identity Management wird zur Automatisierung von Aufgaben aus dem Bereich der Benutzerverwaltung wie das Erzeugen, Löschen oder Ändern von Attributen eingesetzt. Die GWDG bietet einen neuen Dienst „IdM as a service“ an, welcher den Instituten die Möglichkeit gibt, die lokale Benutzerverwaltung im Institut zu vereinfachen. Mit der Anbindung an das IdM der GWDG bekommt der Benutzer eines Instituts Zugriff auf zentrale Dienste wie Exchange 2010, GWDG Cloud Share, Eduroam, VMware, CloudServices.

## EINLEITUNG

Viele Institute stehen vor der Herausforderung, Accounts, Identitäten, Berechtigungen sowie Ressourcen revisionssicher und effizient zu verwalten. Ein Identity Management (IdM) soll hier Abhilfe schaffen und hat sich mittlerweile in vielen Bereichen etabliert. Viele existierende IdM-Systeme sind ausgereift und insbesondere in größeren Umgebungen mit hohem Nutzen einsetzbar.

Bereits seit 2005 unterstützt die GWDG als Kompetenzpartner die Universität Göttingen sowie die Max-Planck-Institute in der Verarbeitung und Vereinfachung von Prozessen im Rahmen des Identity Managements. Innerhalb dieses Zeitraums konnten viele lokale Verzeichnisdienste von Universität und Max-Planck-Instituten an das IdM-System der GWDG angebunden werden. Als Betreiber des zentralen Identitätsmanagements (Identity Management) „MetaDir“ sind eine Vielzahl von Workflows aus unterschiedlichsten Bereichen erfolgreich umgesetzt worden und haben nicht unerheblich für eine Vereinfachung der Verzeichnis- und Benutzerverwaltung am Institut gesorgt, ohne dabei die Autonomie auf Seiten des Instituts aufgeben zu müssen.

## ZIELE DES IDENTITY MANagements

Ein Identity Management umfasst alle organisatorischen und technischen Prozesse der Verwaltung von Accounts, Rollen und Berechtigungen.

### Wesentliche Ziele für ein IdM sind unter anderem:

- Aggregation von existierenden Identitäten und Accounts
- Die Schaffung von Konvergenz in den Bereichen Verzeichnisdienste und Benutzerkonten
- Abbildung von Prozessen für die Benutzeranlage und Deprovisionierung
- Regelung von Zugriffsrechten in den angebundenen Verzeichnissen
- Abbildung von Rollen in den Zielverzeichnissen

In der GWDG werden diese Ziele bei der Anbindung von Verzeichnisdiensten verfolgt und umgesetzt.

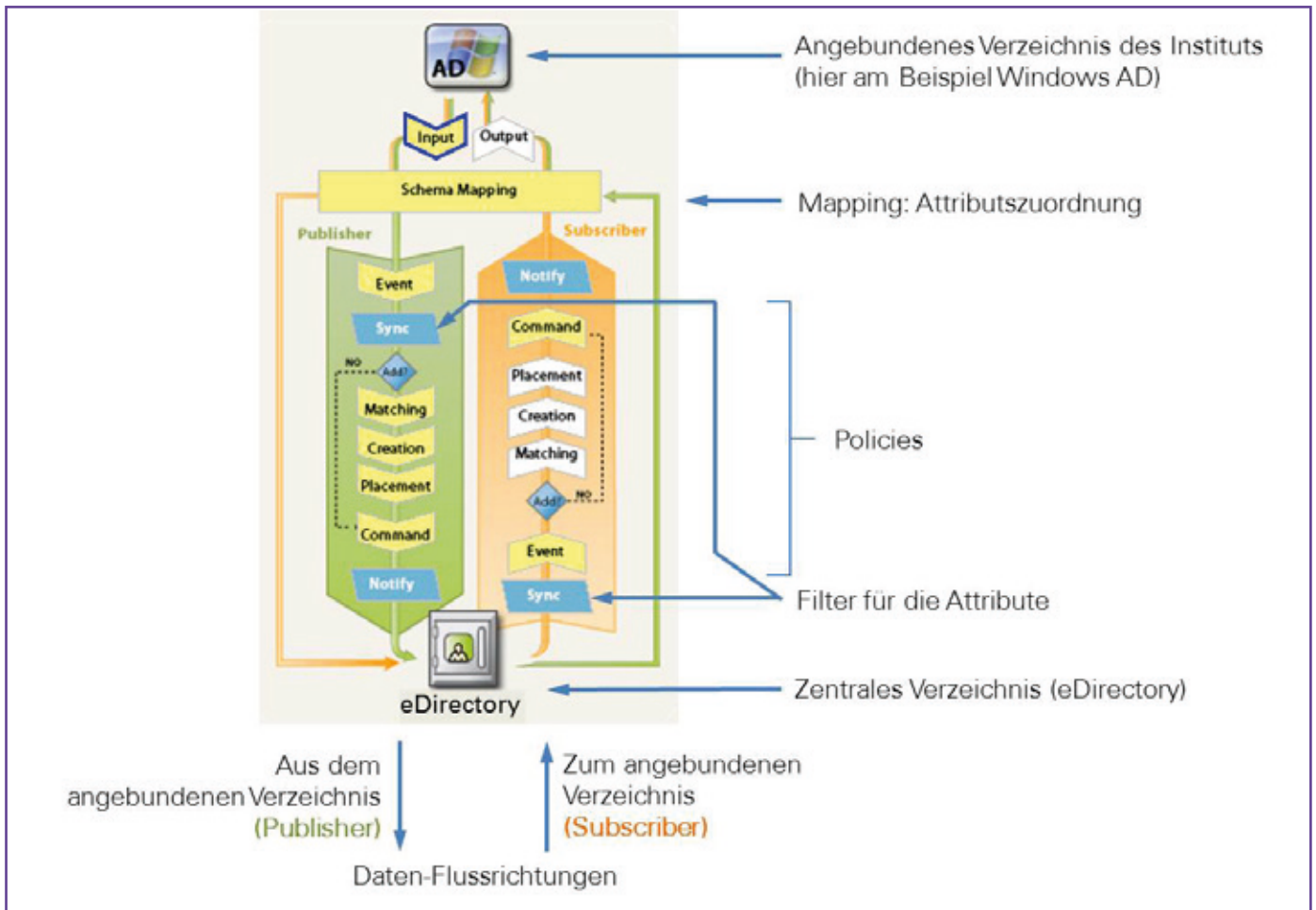
## METADIRECTORY

Ein MetaDirectory (MetaDir) ist die entscheidende technische Komponente des gesamten Identity Managements. Bei dem in der GWDG eingesetztem MetaDirectory handelt es sich um ein sehr ausgereiftes und etabliertes Produkt. Das hier eingesetzte eDirectory von Novell/NetIQ erlaubt eine sehr fein granulierbare Rechteverteilung sowie eine Abbildung von Rollen mit diversen Vererbungsstrategien in einem zentralen Verzeichnis.

Das MetaDirectory verbindet über Konnektoren die angebundenen Verzeichnisse, welche wiederum als Quelle oder Zielverzeichnisse (oder auch beides) fungieren. Im IdM der GWDG definieren dabei Filter für ein angebundenes Verzeichnis, welche Attribute in welche Richtung übertragen werden, und Policies, was mit den Attributen geschehen soll. Auf Seiten des MetaDirectory wird die gesamte Logik in Form eines sogenannten „Treibers“ definiert. Diese IdM-Treiber sind für die unterschiedlichsten Verzeichnisdienste und Anwendungen verfügbar (Windows AD, LDAP, SQL-Datenbanken, Plaintext u.v.m.). Die Flussrichtung eines Treibers kann uni- oder bidirektional definiert sein, sodass Daten zwischen einem angebundenen Verzeichnis und dem MetaDirectory auch in beide Richtungen ausgetauscht werden können. Der Subscriber eines IdM-Treibers beschreibt die Flussrichtung vom MetaDirectory zum Verzeichnisdienst des Institut und der Publisher entsprechend die Rückrichtung.

## Identity management as a service

Identity management is used to automate administrative tasks, such as creating or deleting accounts or change users attributes. The GWDG provides a new service called “IdM as a service” which enables the Institutes to manage very easily their own directory services. With “IdM as a service” the users also get access to central resources of the GWDG (e.g. Exchange 2010, GWDG Cloud Share, Eduroam, VMware, CloudServices).



1\_Schematische Darstellung eines IdM-Treibers

## DAS METADIRECTORY BEI DER GWDG

Das MetaDirectory selbst, basiert auf einem zentralen Verzeichnisdienst (eDirectory), in dem alle Identitäten, Gruppen, Rollen und deren Attribute abgelegt sind.

Seit der Erstinbetriebnahme 2005 ist das MetaDirectory der GWDG mittlerweile zu einem zentralen und unverzichtbaren Werkzeug für die Bewältigung vieler komplexer Prozesse im Rahmen der Benutzerverwaltung der Universität und den Max-Planck-Instituten geworden. Mit momentan ca. 90.000 Identitäten gehört es auch zu den größten Verzeichnisdiensten im wissenschaftlichen Umfeld in Niedersachsen. Derzeit sind 48 Treiber im MetaDirectory aktiv und versorgen 41 angebundene Verzeichnisse. Diese reichen von HIS und FlexNow für die Studierenden der Universität Göttingen, über die Anbindung spezieller Datenbanken bis hin zu den zahlreichen lokalen Verzeichnisdiensten der Max-Planck-Institute.

Auch bei der Treiberentwicklung der GWDG für das IdM, ist Standardisierung und Dokumentation ein wesentliches Ziel. Im Laufe der Anbindung weiterer Max-Planck-Institute haben wir, basierend auf den unterschiedlichen Bedürfnissen, einen standardisierten Treiber entwickeln können, der die meisten Erfordernisse der Institute abdeckt. Dieser vordefinierte Treiber ist damit die Grundlage für eine rasche Anbindung weiterer Institute an das IdM.

## IDM AS A SERVICE

IdM as a service ist einer der neueren Dienste aus dem Dienstleistungskatalog der GWDG. Er orientiert sich an den Bedürfnissen der Max-Planck-Institute sowie der Universität Göttingen und ist deshalb auf die Umgebung in Instituten und deren Anforderungen zugeschnitten. Mit IdM as a service kann die lokale Benutzerverwaltung der Institute entscheidend entlastet werden.

### Unsere Dienstleistung besteht im Wesentlichen aus

- Analyse der lokalen Verzeichnisdienste sowie der Benutzerverwaltung und ggf. mit Vorschlägen zur deren Optimierung
- Abbildung der lokalen Prozesse des Instituts im Rahmen der IdM Anbindung
- Bereitstellung eines zentralen Web-Portals (<https://idm.gwdg.de>) zur Administration von Gruppen und Accounts
- Nutzung von zentralen Diensten der GWDG basierend auf der Anbindung an das IdM
- Umsetzung von Gruppenmitgliedschaften und Rollen im zentralen IdM
- Gemeinsame Dokumentation der Prozesse bei der Anbindung an das IdM, zusammen mit dem Institut
- Standardisierung und Optimierung von Prozessabläufen

Die Anbindung der lokalen Benutzerverwaltung eines Instituts an das IdM bringt unmittelbare Vorteile:

- **Die Autonomie der Benutzerverwaltung bleibt auf Seiten des Instituts erhalten.** Die Kontrolle über alle Prozesse der Benutzerverwaltung bleibt beim Institut.
- **Direkte Nutzung zentraler Dienste der GWDG**  
Die Nutzung zentraler Dienste bei der GWDG ist durch die Anbindung unmittelbar gegeben. Dabei kann der im Institut verwendete Benutzername (oder alternativ auch der UPN oder die E-Mail-Adresse) sowie das Passwort für die Anmeldung an den Diensten verwendet werden. Als Dienste sind unter anderem
  - › Exchange 2010
  - › Spam-Filterung
  - › Eduroam
  - › Sharepoint
  - › GWDG Cloud Share
  - › Virtuelle Server (VMware)
  - › Virtuelle WebServer (Apache-Cluster)

direkt durch die Anbindung des lokalen Institutsverzeichnis am IdM nutzbar. Der Zugang zu den einzelnen

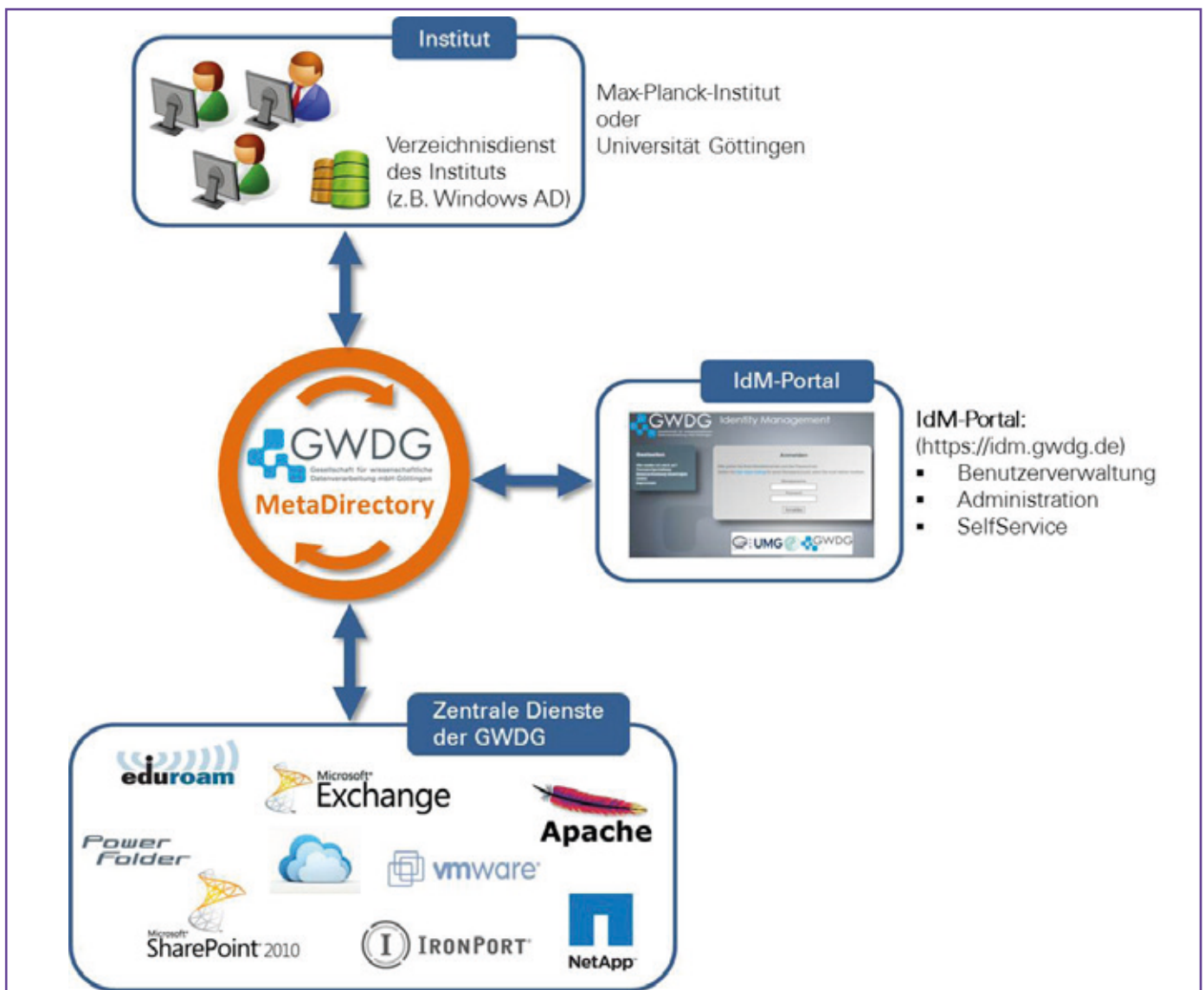
Diensten kann aber auch für jeden Benutzer individuell durch das Institut über das IdM gesteuert werden.

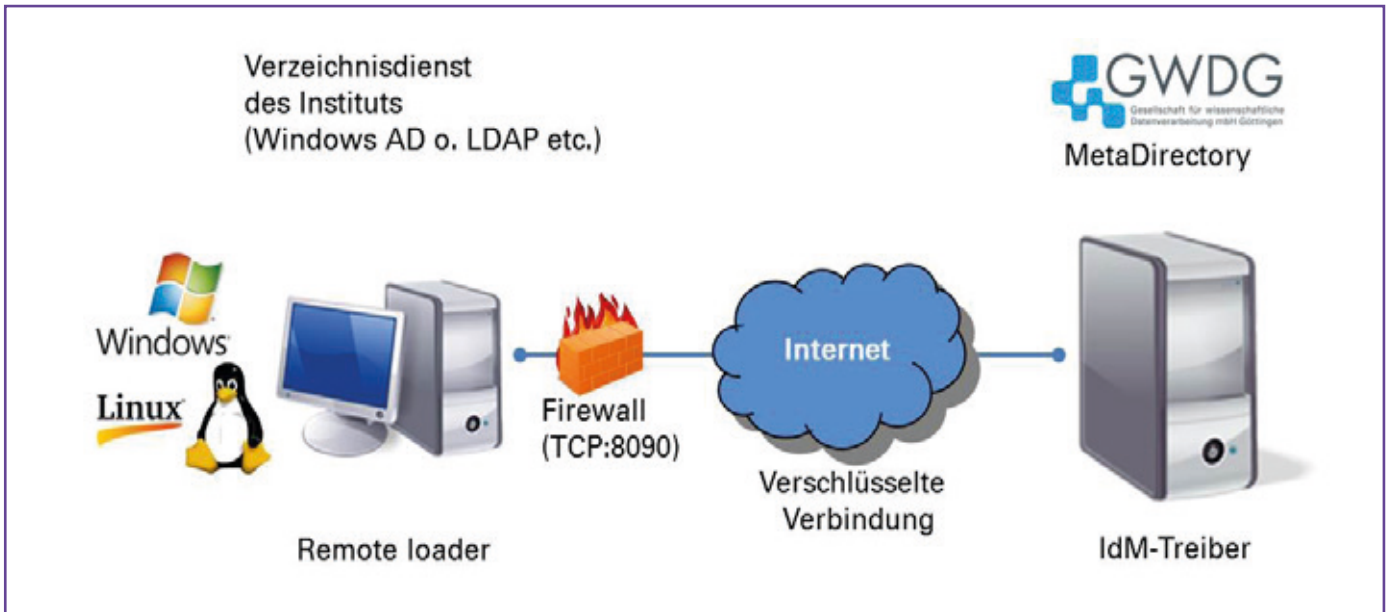
- **Web-Portal für die Administration von Accounts im Institut über das IdM**

Ein zentrales Web-Portal für den Administrator von Benutzern steht für alle relevanten Prozesse (Anlage, Änderung, Löschen, Sperren) für das Institut zur Verfügung. Hierbei kann die Arbeitsumgebung für jedes Institut den individuellen Bedürfnissen angepasst werden (vgl. GWDG-Nachrichten 3/2013). Die Anbindung eines Instituts kann auch so gestaltet werden, dass alle Bereiche der Benutzerverwaltung des Instituts über das Portal erfolgen.

- **Selfservice für den Endanwender**

Der Endanwender kann selbst über das IdM-Portal im Rahmen des SelfService z. B. sein Passwort ändern. Dieses wird dann im IdM und (auf Wunsch) auch automatisch im Verzeichnis des Instituts geändert. Dabei wird auch die Einhaltung der Passwort-Policies des Instituts durch das IdM gewährleistet.





3\_Anbindung eines lokalen Verzeichnisdienst am MetaDirectory

· **Rekonstruktion des lokalen Verzeichnisses durch das IdM im Fehlerfall (Disaster Recovery)**

Ist eine vollständige Anbindung des Instituts an das IdM erfolgt, so kann bei besonderen gravierenden Störfällen am Institut selbst, das Institutsverzeichnis durch das IdM wiederhergestellt werden.

Immer mehr Max-Planck-Institute nutzen unseren Dienst IdM as a service, um ihre lokale Benutzerverwaltung zu entlasten und vor allen Dingen ihren Benutzern einen einfachen Zugang zu den zentralen Diensten der GWWDG zu ermöglichen. Überdies können durch das IdM (MetaDirectory) für die Max-Planck-Institute zukünftige Dienste zeitnah und automatisch an den Endanwender gebracht werden. Eine Vorbereitung für ein einheitliches Format des Benutzernamens ist ein weiterer Schritt zu einem globalen Single-Sign-On z. B. für die Max-Planck-Institute.

**WIE FUNKTIONIERT DIE ANBINDUNG AN DAS IDM?**

Die Anbindung des lokalen Verzeichnisses eines Instituts an das zentrale IdM ist relativ einfach und schnell realisierbar. Die GWWDG bietet vorbereitete Treiber für Windows AD (2003-2012) sowie LDAP (in allen Varianten) an.

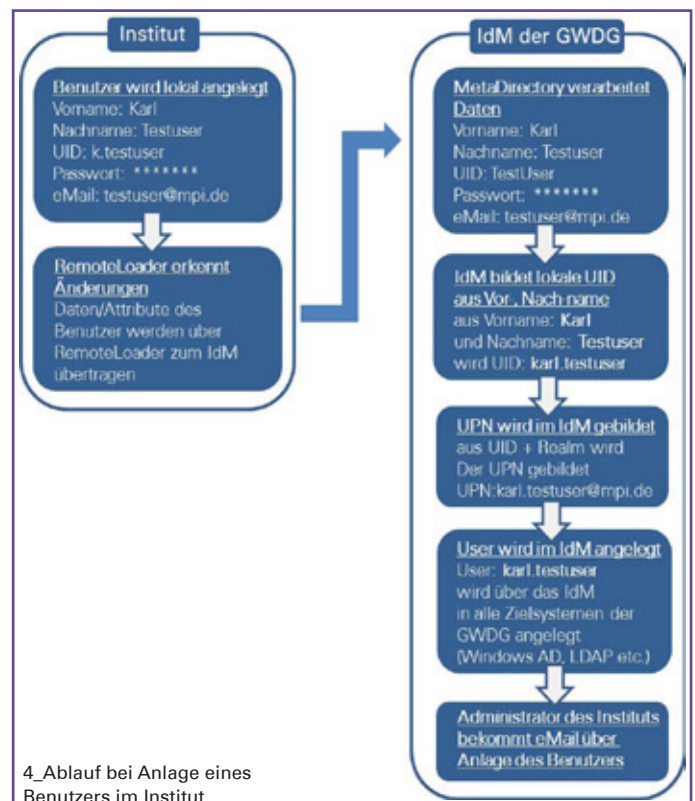
**Was ist auf Seiten des Instituts zu tun?**

Damit die Kommunikation zwischen dem institutseigenen Verzeichnisdienst und dem MetaDirectory möglich ist, wird eine kleine Software (Remote Loader) auf dem lokalen System installiert. Über diesen Remote Loader, welcher für Windows (alle Varianten) sowie LINUX verfügbar ist, werden im Betrieb die Daten zwischen den beiden Systemen über eine durch ein Zertifikat verschlüsselte Verbindung sicher übertragen. Auf der Institutsseite muss lediglich noch die institutseigene Firewall so eingerichtet werden, dass eine Verbindung vom MetaDirectory der GWWDG über TCP-Port 8090 zu dem Verzeichnisdienst möglich ist. Die GWWDG unterstützt natürlich das Institut auch bei der lokalen Installation des RemoteLoaders.

**Funktionsweise des Remote Loaders und des Treibers**

Sämtliche Änderungen am lokalen Verzeichnisdienst (Neuanlage, Löschen oder Änderungen an Attributen) werden vom Remote Loader erkannt und entsprechend den definierten Policies des Treibers zum MetaDirectory übertragen. Das kann, wenn gewünscht, auch das Passwort sein, welches im lokalen Verzeichnis für einen Benutzer gesetzt ist. Für eine Passwortsynchronisation ist dann eine weitere kleine Software (pwfilter) auf Seiten des Servers im Institut zu installieren.

Im Treiber ist definiert, aufgrund welcher Regeln nun welche Aktionen ausgeführt werden und welche Attribute in welche Richtung übertragen werden sollen.



4\_Ablauf bei Anlage eines Benutzers im Institut

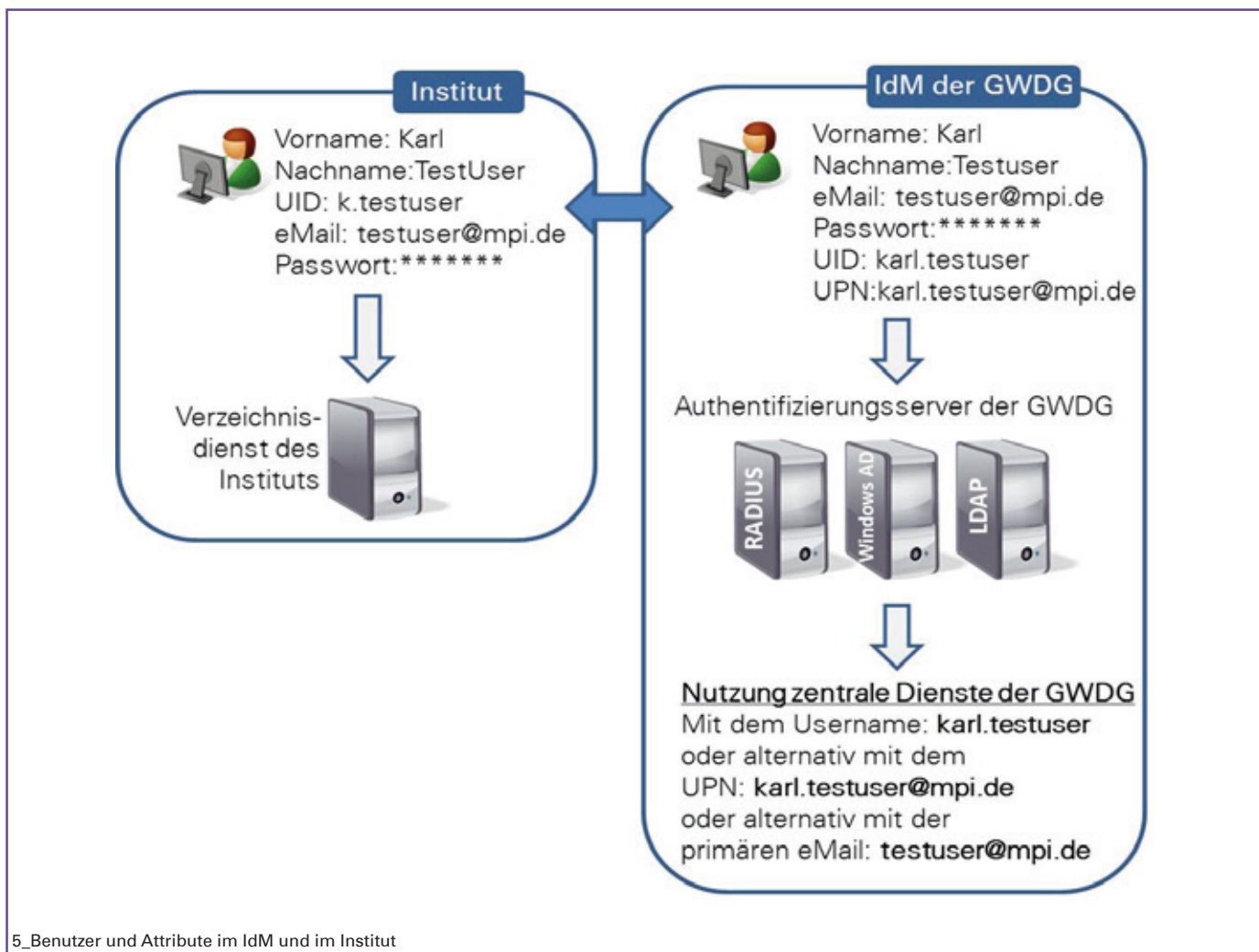


Abbildung 4 zeigt exemplarisch den Ablauf bei Anlage eines Benutzers im lokalen Verzeichnis eines an das IdM der GWDG angebotenen Instituts. Im IdM der GWDG wird aus dem Vor- sowie Nachnamen eines Benutzers eine UID gebildet. Diese UID muss im IdM der GWDG eindeutig sein. Die direkte Übernahme der UID aus dem Quell-Verzeichnis des Instituts ist wenig sinnvoll, da in den Instituten unterschiedlichste Varianten der UID existieren und gerade bei häufigen Nachnamen (z. B. Hans Meier -> UID = meier) eine Eindeutigkeit auf Seiten des IdM nicht gewährleistet ist. Schließlich sind am IdM der GWDG mittlerweile über 40 Verzeichnisse angebotenen, bei denen es auch bereits einen (Hans Meier -> UID = meier) geben kann. Die Bildung der UID im IdM hat keinerlei Rückwirkung auf das Format der UID im Institut. Dieses kann so bei behalten bleiben.

Einen Überblick über den Datensatz für den angelegten Benutzer zeigt Abbildung 5. Hier sind die Attribute des angelegten Benutzers auf beiden Seiten dargestellt. Der Benutzer wird gleichzeitig auch im MetaDirectory angelegt und in der Folge in die anderen am MetaDirectory angeschlossenen Zielsysteme geschrieben. In der Regel sind dieses die Authentifizierungsserver der GWDG (Windows AD Cluster, LDAP Cluster). Dieser Vorgang ist in wenigen Sekunden abgeschlossen und der Benutzer eines Instituts kann bereits ca. 1 Minute nach Anlage seines Account im lokalen Verzeichnis des Instituts, die zentralen Dienste der GWDG nutzen.

## HABEN WIR IHR INTERESSE GEWECKT?

Die Anbindung von Instituten an das IdM der GWDG ist bereits gängige Praxis und eine Voraussetzung für den raschen Zugang zu zentralen Ressourcen und Diensten der GWDG. IdM as a service erlaubt insbesondere auch die Lösung lokaler Probleme bei der Benutzerverwaltung. Wenn im Institut mehrere, auch unterschiedliche, Verzeichnisdienste synchronisiert werden müssen (Windows AD, LDAP), ist IdM as a service ein effektives Werkzeug, um solche Aufgaben zu lösen. Unsere Vorschläge für die Anbindung garantieren gleichzeitig ein Verbleib der vollständigen Kontrolle über das lokale und übergreifende Identity Management beim Administrator des Instituts.

Durch die Anbindung der lokalen Benutzerverwaltung an das IdM erreicht das Institut neben der Flexibilisierung von Prozessen der Benutzerverwaltung vor allen Dingen den automatisierten Zugang ihrer Benutzer zu wichtigen Diensten. Die Vorteile für das Institut und die lokalen Administratoren sind offensichtlich.

Wir unterstützen Sie gerne bei der Bewältigung der Aufgaben im Bereich der Benutzerverwaltung.

Sie erreichen das IdM-Team per E-Mail über: [ids@gwdg.de](mailto:ids@gwdg.de) oder über [support@gwdg.de](mailto:support@gwdg.de)

# Android und die Sicherheit

## Text und Kontakt:

Michael Reimann  
michael.reimann@gwdg.de  
0551 201-1826

Das mobile Betriebssystem Android zeichnet sich durch einen hohen Verbreitungsgrad aus, inzwischen nutzen es weltweit ca. 900 Millionen Anwender. Damit weckt es inzwischen die Begehrlichkeit der Angreifer, und zwar in einem ähnlichen Umfang wie wir es im PC-Umfeld auch von dem Windows-Betriebssystem her kennen. Was hat das für die Anwender zu bedeuten?

Laut eines Berichts der Sicherheitsexperten von Kaspersky Labs soll der überwiegende Anteil der mobilen Schadsoftware auf Android zielen. iOS und Windows Phone spielen bei dieser Aufstellung eine eher untergeordnete Rolle, was allerdings nicht bedeutet, dass sich deren Nutzer in Sicherheit wiegen können. Nur bietet hier die Einbindung in ein geschlossenes Öko-System durchaus Vorteile. Die von vielen Android-Nutzern geschätzte Freiheit, Apps auch aus anderen Quellen als dem Google Play Store installieren zu können, erweist sich schnell als ein nicht zu unterschätzendes Gefährdungsrisiko, wie man gerade bei der neuesten "MasterKey"-Sicherheitslücke sehen kann.

## „MASTERKEY“-SICHERHEITSLÜCKE

Ein Forscherteam des Sicherheitsunternehmens Bluebox hatte Anfang Juli eine gefährliche Sicherheitslücke aufgedeckt, die fast alle Android-Geräte ab der Version 1.6 betrifft und die die komplette Übernahme des Gerätes erlauben würde. Angreifer könnten so Schadcode in eine App einschleusen, ohne dass die kryptographische Signatur verändert wird. Damit könnte auch das Betriebssystem diese Manipulation nicht mehr erkennen, was genau durch diese Signatur eigentlich verhindert werden soll. Jede Anwendung ließe sich so in einen bösartigen Trojaner verwandeln. Google hat diese Schwachstelle zwar inzwischen korrigiert und dieses Update an die Hersteller weitergereicht, doch muss es jetzt erst auf die Geräte kommen, manche werden es erfahrungsgemäß vermutlich nie erhalten. Bislang (Stand Mitte Juli) hat einzig das Samsung

Galaxy S4 diese Korrektur bekommen. Erfreulicherweise zeigte sich, dass Apps aus dem Google Play Store von dieser Sicherheitslücke offenbar nicht betroffen seien. Die Nutzer von Android-Geräten sind daher gut beraten, bei der Installation von Apps immer darauf zu achten, aus welcher Quelle diese stammen.

## WEITERE ANGRIFFE

Aber nicht nur vorhandene Sicherheitslücken im System werden vermehrt für die Verbreitung von Schadsoftware ausgenutzt. Auch andere als harmlose Apps getarnte Schädlinge, die beispielsweise als SMS-Trojaner unerlaubt Kurzmitteilungen an teure Premium-Nummern versenden oder einen Fremdzugriff ermöglichen und vertrauliche Daten ausspähen, tauchten in der Vergangenheit immer wieder auf. Manche davon schafften es sogar in den Google Play Store, wo sie dann aber bald aufgrund der dort ständig verbesserten Überwachungsmaßnahmen wieder entfernt wurden.

## SCHUTZSOFTWARE

Nun dramatisieren die Hersteller von Schutzsoftware in ihren Warnungen immer gerne ein wenig, denn sie wollen ihre Produkte schließlich ja auch verkaufen. Dennoch ist hier auch für viele Sicherheitsforscher der Trend unübersehbar, dass Android auf Grund seiner gewaltigen Verbreitung vermehrt als Plattform zur Verbreitung von Schadsoftware genutzt wird. Diese beunruhigende Erkenntnis nahmen nun zwei Institutionen unabhängig voneinander zum Anlass, sich die derzeit unter Android verfügbaren Schutzprogramme näher anzusehen: Stiftung Warentest (Ausgabe 7/2013) unter: <http://www.test.de/Sicherheits-Apps-fuer-Smartphones-Kein-Zugriff-fuer-Gauner-4560470-0/> und AV-Test in einem seit Januar 2013 laufenden Dauertest unter: <http://www.av-test.org/news/news-single-view/artikel//sicherheits-apps-im-dauerstress/>.

In diesen Tests mussten die Sicherheitsanwendungen vorrangig auf dem Smartphone gespeicherte Schädlinge enttarnen, sie sollten aber auch vor Phishingseiten - betrügerische Internetseiten, die zur Eingabe von Zugangs- oder Kreditkartendaten auffordern - warnen. Zudem wurde untersucht, ob von ihnen private Daten, Positionsdaten oder gar Geräteerkennung übermittelt werden. Gemäß Stiftung Warentest hält „Sicherheits-Software für Android [...] nur selten, was die Hersteller in der Werbung versprechen.

## Android and security

The mobile operating system Android is used now by about 900 million users worldwide. So that it now raises the desirability of the attacker, in a similar extent, as we know it also from the Windows operating system. What does this mean for the user? The mobile operating system Android is used now by about 900 million users worldwide. So that it now raises the desirability of the attacker, in a similar extent, as we know it also from the Windows operating system. What does this mean for the user?



Der Preis ist bei weitem kein Indikator für die Qualität“. Wer sich dennoch mit dem Gedanken trägt, sein Android-Gerät mit einer solchen Schutzsoftware auszustatten, sollte die genannten Testergebnisse in seine Überlegung mit einbeziehen. Eine übrigens von fast allen Produkten angebotene Zusatzfunktion ist die Hilfestellung bei Verlust oder Diebstahl eines Smartphones. Hier seien die Kunden des Exchange 2010 Servers der GWDG daran erinnert, dass sie dort Funktionen wie die Fernlöschung ebenfalls bereits nutzen können. Eine Anleitung findet sich am Ende der folgenden Seite: <http://www.gwdg.de/index.php?id=2610>

## FAZIT

Für viele von uns ist das Smartphone zu einem unverzichtbaren Begleiter geworden. Es zu schützen, ist damit zu einer wichtigen Aufgabe geworden. Ob man dabei zu einem der verfügbaren Sicherheitsprodukte greifen will, muss jeder für sich selber entscheiden. Folgende grundsätzliche Empfehlungen sollten dennoch stets berücksichtigt werden:

- Apps nur aus vertrauenswürdigen Quellen installieren
- (Menü: Einstellungen > Sicherheit > Unbekannte Herkunft)
- Nutzen der Bildschirm-Sperre, um es Dieben nicht allzu leicht zu machen
- (Menü: Einstellungen > Sicherheit > Display Sperre > PIN/ Passwort)
- Durchführen regelmäßiger Sicherheitsupdates, sofern der Hersteller diese anbietet
- (Menü: Einstellungen > Über das Telefon > Systemupdates)
- Diese und weitere Sicherheits-Tipps lassen sich auch der folgenden Aufstellung des BSI (Bundesamt für Sicherheit in der Informationstechnik) entnehmen: [https://www.bsi-fuer-buerger.de/BSIFB/DE/MobileSicherheit/mobileSicherheit\\_node.html](https://www.bsi-fuer-buerger.de/BSIFB/DE/MobileSicherheit/mobileSicherheit_node.html)

Aktuelle (Sicherheits-)Meldungen zu den mobilen Geräten können jederzeit auch auf der Mailingliste GWDG-Mobil verfolgt werden. Näheres dazu: <http://www.gwdg.de/index.php?id=118> ■

# Personalia

## NEUER MITARBEITER CLAUDIUS LAUMANN

Seit dem 1. Juli 2013 hat die Arbeitsgruppe „Anwendungs- und Informationssysteme“ einen neuen Mitarbeiter: Herr Claudius Laumanns. Herr Laumanns hat an der Georg-August-Universität Göttingen Sozialwissenschaften studiert und darüber hinaus im IT-Bereich folgende Zertifikate erworben: Microsoft Certified IT Professional (MCITP) für Konfiguration, Netzwerk und Administration von Active Directory und Windows Server 2008 R2. MCITP Enterprise Administrator für Windows Server 2008 R2. Linux Professional Institute Certification 1 und 2 (LPIC1 und LPIC2). CompTIA Security+ Industriestandard für Security Administration. Herr Laumanns arbeitet im Bereich Beratung und Betreuung von Apple-Macintosh-Anwendern. Er ist unter der Tel.-Nr. 0551 201-1830 zu erreichen, seine E-Mail-Adresse lautet: [claudius.laumanns@gwdg.de](mailto:claudius.laumanns@gwdg.de) Heise





**INFORMATIONEN:**  
support@gwdg.de  
0551 201-1523

August bis  
Dezember 2013

# Kurse

KURS	VORTRAGENDE/R	TERMIN	ANMELDEN BIS	AE
INDESIGN – GRUNDLAGEN	Töpfer	06.08. – 07.08.2013 9:30 – 16:00 Uhr	31.07.2013	8
EINFÜHRUNG IN WINDOWS 7	Buck	13.08.2013 9:00 – 12:30 und 13:30 – 15:30 Uhr	06.08.2013	4
INSTALLATION UND ADMINIS- TRATION VON WINDOWS 7	Buck	20.08.2013 9:00 – 12:30 und 13:30 – 15:30 Uhr	13.08.2013	4
DIE SHAREPOINT-UMGEBUNG DER GWDG	Buck	21.08.2013 9:00 – 12:30 und 13:30 – 15:30 Uhr	14.08.2013	4
ADMINISTRATION VON PCS IM ACTIVE DIRECTORY DER GWDG	Buck	27.08.2013 9:00 – 12:30 und 13:30 – 15:30 Uhr	20.08.2013	4
GRUNDLAGEN DER BILDBEAR- BEITUNG MIT PHOTOSHOP	Töpfer	04.09. – 05.09.2013 9:30 – 16:00 Uhr	28.08.2013	8
INDESIGN – GRUNDLAGEN	Töpfer	10.09. – 11.09.2013 9:30 – 16:00 Uhr	03.09.2013	8
GRUNKURS UNIX/LINUX MIT ÜBUNGEN	Hattenbach	17.09. – 19.09.2013 9:15 – 12:00 und 13:30 – 16:00 Uhr	10.09.2013	12
USING THE GWDG SCIENTIFIC COMPUTE CLUSTER	Dr. Boehme, Ehlers	23.09.2013 9:00 – 12:00 Uhr	16.09.2013	2
PARALLELRECHNERPRO- GRAMMIERUNG MIT MPI	Dr. Haan	23.09.2013 14:00 – 17:00 Uhr 24.09. – 25.09.2013 9:15 – 17:00 Uhr	16.09.2013	10

KURS	VORTRAGENDE/R	TERMIN	ANMELDEN BIS	AE
<b>OUTLOOK – E-MAIL UND GROUPWARE</b>	Helmvoigt	26.09.2013 9:15 – 12:00 und 13:00 – 16:00 Uhr	19.09.2013	4
<b>PHOTOSHOP FÜR FORTGESCHRITTENE</b>	Töpfer	01.10. – 02.10.2013 9:30 – 16:00 Uhr	24.09.2013	8
<b>CLIENT-MANAGEMENT MIT BARAMUNDI</b>	Becker, Körmer, Quentin, Rosenfeld	17.10.2013 9:00 – 12:30 und 13:30 – 15:30 Uhr	10.10.2013	4
<b>INDESIGN – AUFBAUKURS</b>	Töpfer	22.10. – 23.10.2013 9:30 – 16:00 Uhr	15.10.2013	8
<b>UNIX FÜR FORTGESCHRITTENE</b>	Dr. Sippel	04.11. – 06.11.2013 9:15 – 12:00 und 13:15 – 15:30 Uhr	28.10.2013	12
<b>EINFÜHRUNG IN DIE STATISTISCHE DATENANALYSE MIT SPSS</b>	Cordes	13.11. – 14.11.2013 9:00 – 12:00 und 13:00 – 15:30 Uhr	06.11.2013	8
<b>EINFÜHRUNG IN DAS IP-ADRESSMANAGEMENTSYSTEM DER GWDG FÜR NETZWERKBEAUFTRAGTE</b>	Dr. Beck	28.11.2013 10:00 – 12:00 Uhr	21.11.2013	2
<b>UNIX/LINUX-ARBEITSPLATZRECHNER – INSTALLATION UND ADMINISTRATION</b>	Gerdes, Dr. Heuer, Körmer, Dr. Sippel	02.12. – 03.12.2013 9:15 – 12:00 und 13:30 – 16:00 Uhr	25.11.2013	8
<b>UNIX/LINUX-SERVER – GRUNDLAGEN DER ADMINISTRATION</b>	Gerdes, Dr. Heuer, Körmer, Dr. Sippel	04.12. – 05.12.2013 9:15 – 12:00 und 13:30 – 16:00 Uhr	27.11.2013	8
<b>UNIX/LINUX – SYSTEMSICHERHEIT FÜR ADMINISTRATOREN</b>	Gerdes, Dr. Heuer, Körmer, Dr. Sippel	06.12.2013 9:15 – 12:00 und 13:30 – 15:00 Uhr	29.11.2013	4
<b>ANGEWANDTE STATISTIK MIT SPSS FÜR NUTZER MIT VORKENNTNISSEN</b>	Cordes	11.12. – 12.12.2013 9:00 – 12:00 und 13:00 – 15:30 Uhr	04.12.2013	8

#### 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](mailto:support@gwdg.de) erfolgen. Für die schriftliche Anmeldung steht unter <http://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 Institutskontingents 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](mailto:support@gwdg.de) oder telefonisch unter 0551 201-1523 absagen. Bei späteren Absagen werden allerdings die für die Kurse berechneten AE vom jeweiligen Institutskontingents abgebucht.

#### Kursorte

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



Gesellschaft für wissenschaftliche  
Datenverarbeitung mbH Göttingen