

GWDG NACHRICHTEN 11|20

HLRN-IV-System „Emmy“

Nationales
Hochleistungsrechnen

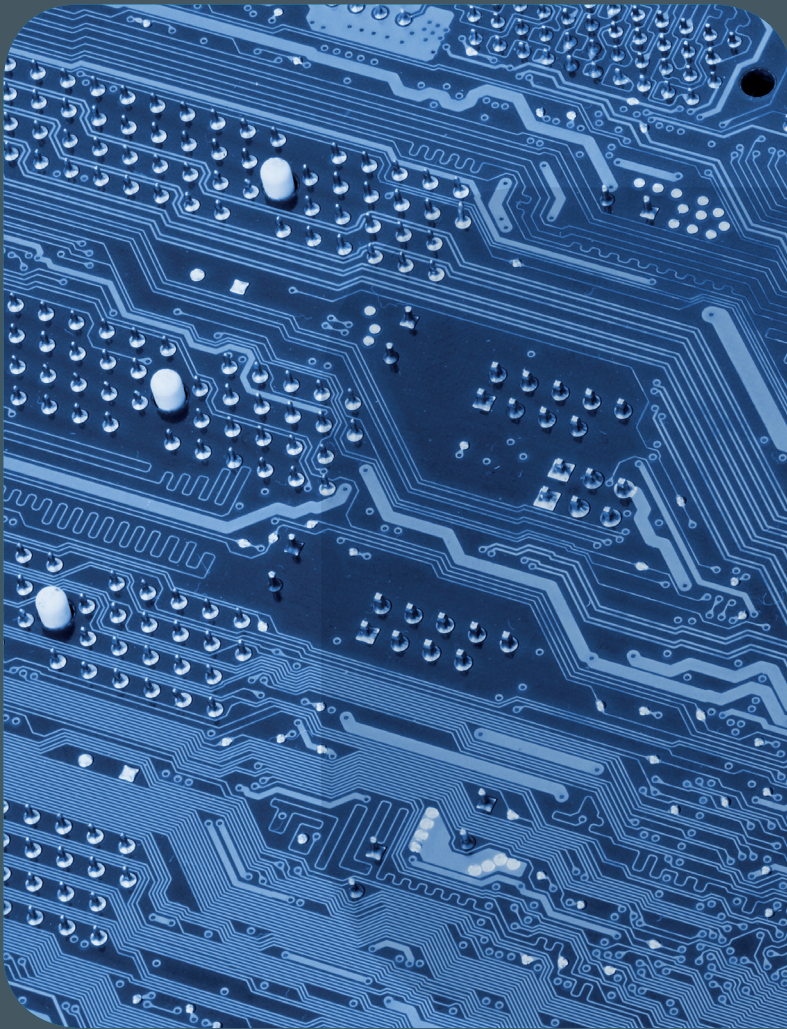
E-Mail-Verschlüsselung

37. DV-Treffen der
Max-Planck-Institute

MINE

ZEITSCHRIFT FÜR DIE KUNDEN DER GWDG





GWGD NACHRICHTEN

11|20 Inhalt

.....

**4 Phase 2 des Göttinger HLRN-IV-Systems
„Emmy“ in Betrieb genommen 7 Göttinger
Supercomputer „Emmy“ ist schnellster Rechner
in Norddeutschland 9 Göttingen wird Standort
für nationales Hochleistungsrechnen**

**11 Kurz & knapp 12 E-Mail-Verschlüsselung mit
X.509-Zertifikaten – Teil 5: Thunderbird, Notes
und Mutt 17 Virtuell verbunden – 37. DV-Treffen
der Max-Planck-Institute 20 MINE – a Text and
Data Mining Platform 23 Academy**

Impressum

.....

Zeitschrift für die Kunden der GWGD

ISSN 0940-4686
43. Jahrgang
Ausgabe 11/2020

Erscheinungsweise:
10 Ausgaben pro Jahr

www.gwdg.de/gwdg-nr

Auflage:
550

Fotos:

© momius - Fotolia.com (8)
© Nataliya Kalabina - Fotolia.com (10)
© Fotogestoeber - Fotolia.com (16)
© DamienGeso - Fotolia.com (22)
© Robert Kneschke - Fotolia.com (23)
© CoolIT (5, Mitte)
© StulzTI (5, unten)
© Gerd Schnapka - GNZ Berlin (18)
© MPLbpc-Medienservice (3)
© GWGD (1, 2, 6, 17)

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 Kund*innen und Freund*innen der GWDG,*

diese Ausgabe der GWDG-Nachrichten hat einen thematischen Schwerpunkt im High Performance Computing, wozu es mehrere erfreuliche Gründe gibt.

Zum einen ist unser neues HPC-System „Emmy“ als Teil des Norddeutschen Verbundes für Hoch- und Höchstleistungsrechnen (HLRN) endlich im Vollausbau in den Regelbetrieb gegangen. Gleich zum Start konnte Emmy in der TOP500-Liste der schnellsten Rechner Platz 47 weltweit belegen. In Deutschland landet Emmy auf Rang fünf und ist damit der leistungsfähigste Rechner in Norddeutschland.

Zum anderen gab es die Entscheidung der Gemeinsamen Wissenschaftskonferenz (GWK) von Bund und Ländern, dass wir ab 2021 eines der acht neuen Zentren für Nationales Hochleistungsrechnen sein werden. Damit einher geht eine Finanzierung von 72 Mio. Euro für die ersten zehn Jahre, die nicht nur für den Rechnerbetrieb, sondern erstmals auch signifikant für Personal für HPC-Forschung und -Beratung zur Verfügung stehen.

Damit schließt sich ein längerer Kreis, der vor acht Jahren begonnen hat, und ist ein schöner Erfolg für die gesamte GWDG. Ab 2021 erhalten wir mit der Inbetriebnahme des Neubaus des gemeinsamen Rechenzentrums dann auch ein räumliches Upgrade mit neuen Rechnerraumkapazitäten.

Damit haben wir zwar das 50-jährige Jubiläum der GWDG in diesem Jahr nicht angemessen feiern können, aber nun doch weitere gute Gründe, um dies noch bei nächster Gelegenheit nachzuholen zu dürfen.

Ramin Yahyapour

GWDG – IT in der Wissenschaft

Phase 2 des Göttinger HLRN-IV-Systems „Emmy“ in Betrieb genommen

Text und Kontakt:

Marcus Vincent Boden
marcus-vincent.boden@gwdg.de
0551 201-2191

Dr. Christian Köhler
christian.koehler@gwdg.de
0551 201-2193

Sebastian Krey
sebastian.krey@gwdg.de
0551 39-30277

Seit Ende Oktober 2020 ist die zweite Ausbaustufe von „Emmy“, dem Göttinger Teil des HLRN-IV-Systems, in Betrieb. Die Installation erfolgte im neuen Modularen Rechenzentrum (MDC) der Universität Göttingen, zusammen mit der neuesten Hardwaregeneration des Scientific Compute Clusters der GWDG. Auch die erste Ausbaustufe des HLRN-IV ist an den neuen Standort umgezogen. Die MDC-Lösung zeichnet sich durch eine besonders effiziente Warmwasserdirektkühlung für einen Großteil der hier installierten Systeme aus.

ÜBERBLICK HLRN

Als Norddeutscher Verbund zur Förderung des Hoch- und Höchstleistungsrechnens (HLRN) arbeiten die Bundesländer Berlin, Bremen, Hamburg, Mecklenburg-Vorpommern, Niedersachsen und Schleswig-Holstein seit 2001 zusammen, seit 2012 ist auch Brandenburg Teil des Verbundes. Betreiberzentren für die HPC-Systeme des HLRN sind seit 2018 die Universität Göttingen und das Zuse-Institut Berlin (ZIB). Beide Zentren werden ab 2021 Standorte des neu formierten Verbundes Nationales Hochleistungsrechnens (NHR) und erhalten Fördermittel in Höhe von jeweils 72 Millionen Euro über zehn Jahre.

Der Betrieb der ersten Phase des HLRN-IV begann in 2018. Während am ZIB zunächst das ursprünglich für den HLRN-III beschaffte Cray XC30/XC40-System „Konrad“ [1] als HLRN-IV Phase 1 weiterbetrieben wurde, nahm Göttingen den Betrieb nach Installation der ersten Phase des HLRN-IV-Systems „Emmy“ auf. Dieses ist nach Emmy Noether benannt, die als Mathematikerin an der Universität Göttingen als erste Frau in Deutschland habilitierte, und wurde durch den französischen IT-Dienstleister Atos SE installiert. Das System verfügte zunächst über 449 Dual-Socket-Knoten auf Basis von Intel Xeon Gold 6148 CPUs (Skylake-Mikroarchitektur). Es stehen insgesamt 40 CPU-Kerne pro Knoten bereit. Im Einzelnen handelt es sich um

- 432 Standard-Knoten mit je 192 GB Speicher und 480 GB lokaler SSD,
- 16 Fat-Knoten mit je 768 GB Speicher, 480 GB lokaler SSD und
- einen GPU-Knoten mit 768 GB Speicher, 240 GB lokaler SSD und vier nVidia Tesla V100 GPUs (jeweils 32 GB VRAM).

Letzterer ist vor allem für Machine Learning und interaktive Visualisierungsaufgaben geeignet. Die Dateisysteme bestehen aus einem 8,1 PB fassenden auf Lustre basierenden DDN- ExaScaler-Filesystem, einem als NFS exportierten 320 TB DDN- GridScaler-Filesystem und einem DDN IME Burst Buffer für I/O- intensive

Anwendungen. Die technischen Daten zum Göttinger HLRN-IV-System „Emmy“ sowie dem GWDG Scientific Compute Cluster (SCC) sind im Profil der GWDG bei der Gauß-Allianz [2] zu finden.

HLRN-IV PHASE 2

Seit Ende Oktober 2020 ist die zweite Phase des Systems „Emmy“ in Betrieb. Diese besteht aus insgesamt 974 Dual-Socket-Knoten der Intel Cascade Lake Advance Performance (CLX-AP) Generation [3]. Es sind jeweils zwei Intel Xeon Platinum 9242 CPUs installiert, so dass pro Knoten insgesamt 96 Kerne zur Verfügung stehen. Dabei sind die Knoten im Einzelnen wie folgt ausgestattet:

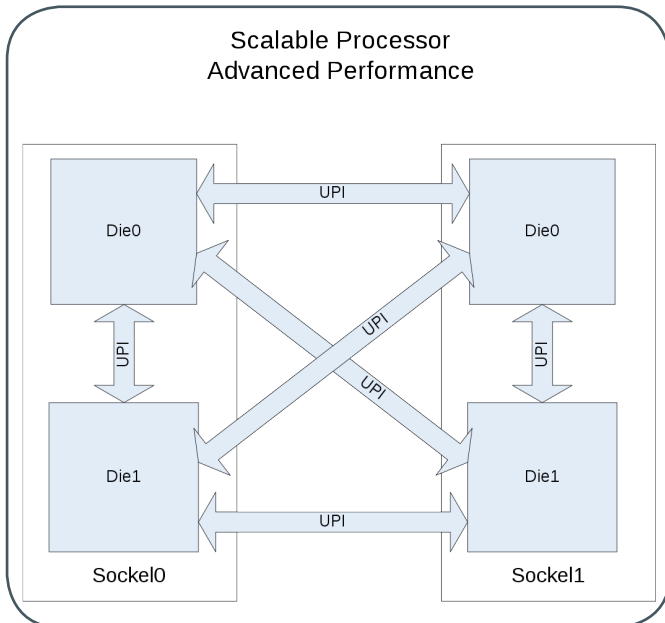
- 956 Standard-Knoten mit je 384 GB Speicher,
- 16 Fat-Knoten mit je 768 GB Speicher und
- 2 Superfat-Knoten mit je 1.536 GB Speicher.

Diese Anzahl an CPU-Kernen (bei Intel als „Extreme Core Count“ (XXC) bezeichnet) wurde durch die Kombination zweier Dies (= ununterbrochener Abschnitt eines Wafers bei der Halbleiterherstellung) auf einem Substrat erzielt, die in ein gemeinsames Package verbaut werden. Das Design hat Konsequenzen für klassische HPC-Anwendungen wie Simulationen mit hohem Kommunikationsbedarf

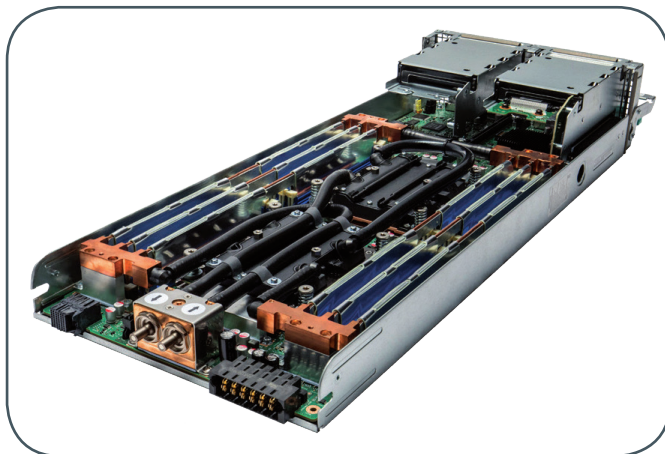
HLRN-IV Phase 2 in Göttingen in Operation

Since the end of October 2020 the second phase of „Emmy“, the Göttingen part of the HLRN-IV system, is in operation. The installation was done in the new modular data center (MDC) of the University of Göttingen, in conjunction with the most recent hardware generation of the GWDG Scientific Compute Cluster (SCC). The first phase of HLRN-IV has also been moved to the new site. The MDC solution features particularly efficient direct warm-water cooling for the majority of the systems installed here.

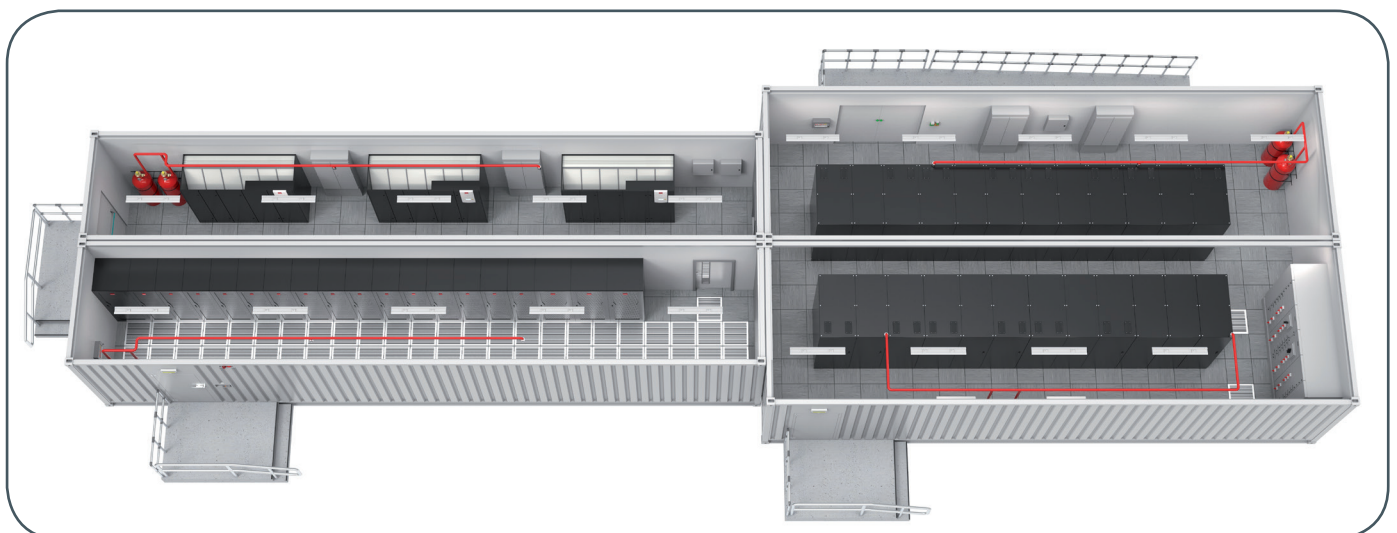
zwischen den Prozessen: Während die CPUs in Dual-Socket-Sky-lake-Knoten, wie sie in Phase 1 zum Einsatz kommen, durch drei Intel Ultra Path Interconnect (UPI) Links (jeweils 10,4 GT/s (Gigatransfers pro Sekunde), entsprechend einer unidirektionalen Bandbreite von 20,8 GB/s) verbunden sind, ist die gleiche Anzahl an Links



1_Cascade Lake AP CPU [4]



2_Warmwassergekühlter Knoten HLRN-IV Phase 2



3_Modulares Rechenzentrum (MDC), links der luftgekühlte, rechts der warmwassergekühlte Bereich

nun pro Die verfügbar (siehe Abbildung 1) und teilt sich somit auf die Verbindung zur anderen Hälfte des gleichen Packages sowie zu den beiden Dies am anderen CPU-Sockel auf, so dass sich die Bandbreite zwischen je zwei Dies auf ein Drittel reduziert. Aus Sicht der Codeoptimierung betrachtet, handelt es sich also bei den Knoten der Phase 2 um 4-Socket-Server, obwohl man sie auf den ersten Blick für Dual-Socket-Systeme (siehe Abbildung 2) halten könnte. Daher ist besonderes Augenmerk auf die geschickte (und für das Betriebssystem nicht mehr dynamisch zu ändernde (Pinning)) Platzierung der Prozesse einer Anwendung zu legen, um nicht durch unnötige Zugriffe auf den an ein anderes Die angebotenen Speicher kostbare UPI-Bandbreite zu binden und die Zugriffslatenz zu erhöhen.

Beide Phasen sind durch eine gemeinsame 100 Gbps Intel Omni-Path Fabric vernetzt [5], an die auch die neueste Generation der Systeme des GWDG Scientific Compute Clusters (SCC) angeschlossen ist. Im Oktober 2020 wurde auf beiden Phasen des HLRN-IV-Systems „Emmy“ zusammen mit dem neuen Hardwareupgrade des GWDG SCC ein TOP500-Benchmark durchgeführt und das Ergebnis in der Liste der 500 weltweit schnellsten Systeme im November veröffentlicht [6]. Danach erreicht das Gesamtsystem Platz 47 weltweit und landet auf Rang fünf in Deutschland und ist damit der leistungsfähigste Rechner in Norddeutschland (siehe die entsprechende GWDG-Presseinformation in dieser Ausgabe der GWDG-Nachrichten).

Für das Management (Software, Fabric etc) und weitere Dienste kommen sowohl dedizierte Service-Knoten als auch virtuelle Maschinen aus der Cloud-Umgebung der GWDG zum Einsatz. Für Archiv- und Backupdienste ist das System an die Infrastruktur der GWDG angebunden. Das System ist über eine dedizierte, durch den DFN-Verein bereitgestellte 10G-Verbindung, den „HLRN-Link“, mit dem Berliner HLRN-IV-System „Lise“, benannt nach der Physikerin Lise Meitner, verbunden. Mit der Inbetriebnahme von „Lise“ wurde der HLRN-IV im Dezember 2019 am ZIB feierlich eingeweiht [7].

Das HPC-Team der GWDG unterstützt den HLRN-Verbund nicht nur durch den Betrieb und die Weiterentwicklung des Göttinger Systems, sondern auch durch die Fachberatung der Anwenderschaft, die Pflege des Softwareangebots und die Mitarbeit in den Gremien des HLRN. Insbesondere erfolgt die kontinuierliche Abstimmung mit dem ZIB, der überregionalen Fachberatung und dem Beirat Informationssicherheit.



4_Freiluftkühlung am MDC

MODULARES RECHENZENTRUM

Speziell für den HLRN-IV und die neueste Generation des Scientific Compute Clusters der GWDG [8] haben die Universität Göttingen und die GWDG neue Rechenzentrumsflächen in Form eines modularen Rechenzentrums (Modular Data Center, MDC) am Hans-Adolf-Krebs-Weg bereitgestellt. Die Nutzung des MDC begann im Juli 2020, ein Jahr nach der Auftragsvergabe, mit dem Umzug der ersten Phase des HLRN-IV aus den Rechnerräumen in der ehemaligen Fernmeldezentrale und dem parallelen Aufbau der zweiten Phase. Ende Oktober bzw. Anfang November sind die zweite Ausbaustufe des HLRN-IV und die Erweiterung des SCC in den neuen Räumlichkeiten in Betrieb gegangen. Die Errichtung des MDC war aufgrund des stark erhöhten Flächenbedarfs für HPC-Systeme noch vor Verfügbarkeit des neuen Rechenzentrumsgebäudes notwendig geworden. Das MDC ist speziell an den Bedarf der HLRN-Systeme und des SCC angepasst und erreicht daher bei vergleichsweise geringen Investitionskosten für diese Systeme dennoch eine sehr hohe, zeitgemäße Energieeffizienz.

Das MDC ist in zwei Räume aufgeteilt (siehe Abbildung 3). Der linke Raum ist luftgekühlt und bietet Platz für 19 Racks (17 x 42 HE und 3 x 45 HE) mit einer baulichen Trennung in Kalt- und Warmgang. Hier sind die Stagesysteme, die Director Switches der Omni-Path Fabric, die Phase 1 des HLRN-IV sowie eine GPU-Partition des SCC untergebracht. Der rechte Raum bietet Platz für 14 geschlossene Racks mit Warmwasserdirektkühlung (48 HE plus 6 x 1 HE für Switches an den Seiten, 96 Computeknoten). Hiervon werden aktuell zwölf Stellplätze genutzt (elf Racks für die Phase 2 des HLRN-IV sowie ein Rack für den SCC). Die mit Warmwasser gekühlten Systeme geben über eine freie Direktkühlung (siehe Abbildung 4) mit 32°C Vorlauftemperatur 80 % der Verlustleistung an das Wasser ab. Die übrigen 20 % werden über InRow-Kühler abgeführt. Bei der Leistungsprüfung der Kühlung wurden die eingebauten Systeme 48 Stunden unter HPL-Volllast betrieben. Dabei lag die Leistungsaufnahme des gesamten MDC bei knapp unter 1,5 MW, wovon 1,36 MW auf die IT-Systeme entfiel, was einer Gesamt-PUE (PUE = Power Usage Efficiency) unter Volllast von 1,1 entspricht. Eine getrennte Betrachtung nach den beiden Kühltechniken ergibt eine PUE von 1,03 für die

Warmwasserkühlung und 1,23 für die Luftkühlung. Bei sehr niedriger Last steigen die PUE-Werte auf 1,07 bzw. 1,36 und 1,18 für das gesamte MDC.

FACHBERATUNG

Neben dem Betrieb der Hochleistungsrechner durch die Universität Göttingen und das Zuse-Institut Berlin stellt das HLRN-Kompetenznetzwerk, das sich u. a. aus den beteiligten Lehrstühlen, Forschungsinstituten und lokalen Tier-3-Rechenzentren der beteiligten Bundesländer sowie HLRN-Gremien zusammensetzt, einen zentralen Baustein des Verbunds dar [9]. Aus Nutzer*innensicht ist dabei das Berater*innennetzwerk hervorzuheben: Lokale Berater*innen leisten an den HLRN-Standorten insbesondere Unterstützung für die Nutzer*innen beim Zugang zu den Systemen des HLRN.

Das Fachberater*innennetzwerk ist überregional in Zusammenarbeit (und teilweise personeller Überschneidung) mit den Betreiberzentren aktiv bei der Installation und Pflege der Anwendungssoftware für alle durch den Verbund bedienten wissenschaftlichen Gebiete. Dies umfasst insbesondere Erdsystemwissenschaften, computerbasierte Chemie und Materialwissenschaften, Lebenswissenschaften und Strömungswissenschaften, so dass auf den Systemen des HLRN u. a. Ozeanmodelle, Quantenchemie-Codes und Molekulardynamik-Simulationen bzw. strömungsmechanische Modelle gerechnet werden. Unterstützung wird dabei geleistet von der Beratung beim Entwurf von Anträgen auf Rechenzeit, der Handhabung der Systeme und Anwendungssoftware, bis hin zur Zusammenarbeit zwischen Nutzer*innen, Fachberater*innen und Systemanbietern bei der Codeoptimierung für jede neue Architektur.

LINKS

- [1] <https://www.hlrn.de/supercomputer/hlrn-iii-system/>
- [2] <https://gauss-allianz.de/de/profile>
- [3] https://en.wikichip.org/wiki/intel/microarchitectures/cascade_lake
- [4] Bildnachweis: [https://de.wikipedia.org/wiki/Intel_Xeon_\(Cascade_Lake\)#/media/Datei:Intel_ScalableProcessor-Advanced_Performance.svg](https://de.wikipedia.org/wiki/Intel_Xeon_(Cascade_Lake)#/media/Datei:Intel_ScalableProcessor-Advanced_Performance.svg), Lizenz: Attribution-ShareAlike 3.0 Unported (CC BY-SA 3.0)
- [5] Siehe auch die GWDG Nachrichten 4/2018
- [6] <https://www.top500.org/lists/top500/>
- [7] <https://www.hlrn.de/eroeffnung2019/>
- [8] <https://www.gwdg.de/hpc-on-campus/scc>
- [9] <https://www.hlrn.de/ueber-uns/netzwerk/>



Göttinger Supercomputer „Emmy“ ist schnellster Rechner in Norddeutschland

Text:
GWDG-Pressinformation 3/2020
vom 17.11.2020 (mit kleinen Änderungen)

In der neuesten Ausgabe der TOP500-Liste der schnellsten Rechner der Welt belegt das in Göttingen installierte System „Emmy“ weltweit Platz 47 und in Deutschland Rang fünf. Emmy ist damit der schnellste Rechner in Norddeutschland und ein System des Norddeutschen Verbundes für Hoch- und Höchstleistungsrechnen (HLRN), das von der GWDG und der Universität Göttingen betrieben wird. Der Superrechner „Emmy“ hat von einer theoretischen Spitzenleistung von 8,78 PFlop/s (Billiarden Rechenoperationen pro Sekunde) zusammen mit dem HPC-System der GWDG in dem für diesen Vergleich genutzten Linpack-Benchmark 5,95 PFlop/s erreicht.

Spitzenplatzierung für Göttinger Supercomputer: In der neuesten Ausgabe der Top500-Liste der schnellsten Rechner der Welt belegt das in Göttingen installierte System „Emmy“ weltweit Platz 47. In Deutschland landet Emmy auf Rang fünf und ist damit der leistungsfähigste Rechner in Norddeutschland. Emmy ist ein System des Norddeutschen Verbundes für Hoch- und Höchstleistungsrechnen (HLRN), das von der Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) und der Universität Göttingen betrieben wird.

„Die Leistungsfähigkeit von Emmy erlaubt enorme Fortschritte in der Forschung in Computational Science, Künstlicher Intelligenz und Data Science“, sagt Prof. Dr. Ramin Yahyapour, CIO der Universität Göttingen und CEO der GWDG. Erst am vergangenen Freitag hatte Göttingen den Zuschlag als Standort im Verbund Nationales Hochleistungsrechnen (NHR) mit einer Förderung von 72 Millionen Euro über zehn Jahre erhalten.

Superrechner „Emmy“ ist nach Emmy Noether benannt, die als Mathematikerin an der Universität Göttingen als erste Frau in Deutschland habilitierte. Von einer theoretischen Spitzenleistung von 8,78 PFlop/s (Billiarden Rechenoperationen pro Sekunde) hat Emmy zusammen mit dem HPC-System der GWDG in dem für diesen Vergleich genutzten Linpack-Benchmark 5,95 PFlop/s erreicht. Die sieben Bundesländer Berlin, Brandenburg, Bremen, Hamburg, Mecklenburg-Vorpommern, Niedersachsen und Schleswig-Holstein betreiben seit 2001 (Brandenburg seit 2012) im HLRN-Verbund gemeinschaftlich ein verteiltes Supercomputersystem an den beiden Betreiberzentren Universität Göttingen und Zuse-Institut Berlin. Emmys HLRN-Schwestersystem in Berlin, „Lise“ (benannt nach Lise Meitner), liegt mit einer Leistung von 5,36 PFlop/s auf Platz 56 der aktuellen Top500-Liste, was Platz acht im deutschen Vergleich entspricht.

Der von der Firma Atos auf Basis von Intel-Prozessor- und Netzwerktechnologie für den HLRN designte Superrechner liefert

seine Rechenleistung mit 974 Computeknoten mit je zwei Intel Xeon Platinum 9242 Cascade Lake AP Prozessoren sowie 449 Computeknoten mit je zwei Intel Xeon Gold 6148 SkyLake SP. Es stehen also pro Knoten 96 bzw. 40 CPU-Kerne zur Verfügung. Die Computeknoten sind mit 192 GB bis 1.536 GB Arbeitsspeicher ausgestattet und es stehen ein geteiltes 8,1 PiB großes Lustre-Dateisystem (DDN ExaScaler) sowie ein 48 TiB NVMe Burstbuffer (DDN Infinite Memory Engine) zur Verfügung. Die einzelnen Knoten des Systems sind durch ein Intel Omni-Path-Hochgeschwindigkeitsnetzwerk mit 100 GBit/s miteinander verbunden.

Die Cascade Lake AP-Knoten sind mit einer Warmwasserdirektkühlung ausgestattet, wodurch ein besonders energieeffizientes Design erreicht werden konnte. Emmy ist in einem modularen, speziell an die Anforderungen des Systems angepassten

Supercomputer „Emmy“ in Göttingen is North Germany's Fastest Computer

Top ranking for a supercomputer in Göttingen: In the recent edition of the Top500 list of the world's fastest computers, the „Emmy“ system installed in Göttingen is ranked 47th in the world. In Germany, Emmy ranks fifth, and is thus the most powerful computer in North Germany. Emmy is a system of the North German Supercomputing Alliance (HLRN), which is operated by the Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) and the University of Göttingen. The supercomputer „Emmy“ has reached 5.95 PFlop/s from a theoretical peak performance of 8.78 PFlop/s (quadrillion computing operations per second) together with the HPC system of the GWDG in the Linpack benchmark used for this comparison.

Rechenzentrum von Stulz Technology Integration Ltd aufgestellt, welches basierend auf Containern in einer kompakten Bauform zwei Maschinenräume bereitstellt. Einer der Räume bietet Platz für 14 Racks mit Warmwasserdirektkühlung und der zweite Raum ist mit 19 luftgekühlten Racks ausgestattet. Durch größtmöglichen Einsatz von freier Kühlung konnte eine Power Usage Efficiency (PUE, Energieeffizienz) von 1,03 unter Vollast für den Warmwasserbereich erreicht werden. Je näher der PUE-Wert bei 1 liegt, desto besser die Energieeffizienz. In der Bewertungsskala laut Green Grid gelten PUE-Werte kleiner 1,5 als effizient und Werte kleiner 1,2 als sehr effizient.

Die TOP500 ist eine Liste der 500 schnellsten Computersysteme der Welt. Sie wird seit 1986 zweimal im Jahr herausgegeben. Die Sortierung ergibt sich aus dem High Performance Linpack-Benchmark und stellt damit eine Rangfolge der leistungsfähigsten Maschinen zur Lösung linearer Gleichungssysteme dar. Die gesamte Rangliste ist unter <https://www.top500.org> zu finden.

Atos ist ein weltweit führender Anbieter für die digitale Transformation mit 110.000 Mitarbeitern in 73 Ländern und einem Jahresumsatz von 12 Milliarden Euro. Als europäischer Marktführer für Cloud, Cybersecurity und High Performance Computing bietet die Atos Gruppe ganzheitliche Lösungen für Orchestrated Hybrid

Cloud, Big Data, Business-Anwendungen und Digital Workplace. Der Konzern ist der weltweite Informationstechnologie-Partner der Olympischen und Paralympischen Spiele und firmiert unter den Marken Atos, Atos|Syntel und Unify. Atos ist eine SE (Societas Europaea) und an der Pariser Börse als eine der 40 führenden französischen Aktiengesellschaften (CAC40) notiert.

Intel (NASDAQ: INTC) ist eines der führenden Unternehmen in der Entwicklung zukunftsweisender Technologien, von denen Menschen auf der ganzen Welt profitieren. Inspiriert durch das Mooresche Gesetz arbeitet Intel ständig an der Weiterentwicklung ihrer Halbleiter-Produkte, um Kunden bei der Bewältigung ihrer größten Herausforderungen bestmöglich zu unterstützen. Intel implementiert intelligente Technologien und Lösungen in der Cloud, im Netzwerk und in allen Geräten dazwischen. Dadurch nutzt Intel das volle Potenzial von Daten, die Unternehmen und Gesellschaft verbessern.

Die GWDG betreibt neben Emmy, dem Göttinger HLRN-IV-System, mit dem „Scientific Compute Cluster“ einen weiteren Hochleistungsrechner für den Göttinger Wissenschaftsstandort. Ein von der GWDG für das Deutsche Zentrum für Luft- und Raumfahrt (DLR) betriebener Hochleistungsrechner ist in der Vorbereitung. ■



FTP-Server

Eine ergiebige Fundgrube!

Ihre Anforderung

Sie möchten auf das weltweite OpenSource-Softwareangebot zentral und schnell zugreifen. Sie benötigen Handbücher oder Programmbeschreibungen oder Listings aus Computerzeitschriften. Sie wollen Updates Ihrer Linux- oder FreeBSD-Installation schnell durchführen.

Unser Angebot

Die GWDG betreibt seit 1992 einen der weltweit bekanntesten FTP-Server mit leistungsfähigen Ressourcen und schneller Netzanbindung. Er ist dabei Hauptmirror für viele Open-Source-Projekte.

Ihre Vorteile

- > Großer Datenbestand (65 TByte), weltweit verfügbar
- > Besonders gute Anbindung im GÖNET

- > Aktuelle Software inkl. Updates der gebräuchlichsten Linux-Distributionen
- > Unter pub befindet sich eine aktuell gehaltene locatedb für schnelles Durchsuchen des Bestandes.
- > Alle gängigen Protokolle (http, https, ftp und rsync) werden unterstützt.

Interessiert?

Wenn Sie unseren FTP-Server nutzen möchten, werfen Sie bitte einen Blick auf die u. g. Webseite. Jeder Nutzer kann den FTP-Dienst nutzen. Die Nutzer im GÖNET erreichen in der Regel durch die lokale Anbindung besseren Durchsatz als externe Nutzer.

>> www.gwdg.de/ftp-server

Göttingen wird Standort für nationales Hochleistungsrechnen

Text:
Presseinformation Nr. 166 der
Georg-August-Universität Göttingen
vom 13.11.2020 (mit kleinen Änderungen)

Die GWDG als Hochschulrechenzentrum der Universität Göttingen wurde als eines von acht Rechenzentren für den neu gegründeten Verbund Nationales Hochleistungsrechenzentren (NHR) ausgewählt. Damit erhalten die Universität Göttingen und die GWDG Fördermittel in Höhe von 72 Millionen Euro über zehn Jahre und können den Ausbau des Hochleistungsrechnens, das als Schlüsseltechnologie für wissenschaftlichen und technischen Fortschritt gilt, als zentralen Bestandteil ihrer Standortstrategie weiter vorantreiben.

Göttingen wird einer der Standorte des neu gegründeten Verbunds Nationales Hochleistungsrechnen (NHR). Das hat die Gemeinsame Wissenschaftskonferenz (GWK) von Bund und Ländern heute bekanntgegeben. Damit verbunden sind Fördermittel in Höhe von insgesamt rund 72 Millionen Euro über zehn Jahre. Der Antrag wurde im Namen der sieben norddeutschen Bundesländer des Norddeutschen Verbunds für Hoch- und Höchstleistungsrechnen (HLRN) für die beiden Standorte Universität Göttingen/Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) und das Zuse-Institut Berlin (ZIB) für ein gemeinsames NHR-Zentrum gestellt. Berlin erhält eine ebenso hohe Förderung.

Prof. Dr. Ramin Yahyapour, CIO der Universität Göttingen und CEO der GWDG, Antragsteller für Göttingen, freut sich über die Anerkennung des HLRN-Konzeptes mit zwei zusammenarbeitenden Standorten für High Performance Computing (HPC) in Berlin und Göttingen: „Diese Förderung erlaubt es uns, rechnergestützte Forschung auf höchstem Niveau insbesondere in den Kompetenzschwerpunkten Lebenswissenschaften, Klimaforschung, Chemie und Ingenieurwissenschaften weiter auszubauen.“

Universitätspräsident Prof. Dr. Reinhard Jahn sieht in der Förderung die Strategie der Universität bestätigt: „In vielen wissenschaftlichen Disziplinen ist Spitzenforschung heutzutage zunehmend rechen- und datenintensiv, weshalb wir diesen Bereich sukzessive ausgebaut haben. Hier ist Göttingen nun hervorragend aufgestellt.“

Niedersachsens Wissenschaftsminister Björn Thümler gratuliert den Spitzenforscherinnen und -forschern: „Die Erfolgsgeschichte des Hochleistungsrechnens in Norddeutschland und speziell am Standort Göttingen wird mit der Aufnahme in den NHR-Verbund fortgeschrieben. Mit dieser Entscheidung wird Forscherinnen und Forschern der Rahmen für die Spitzenforschung von morgen bereitet. Wir sind hocherfreut!“

Die GWK hatte sich im November 2018 auf die gemeinsame Förderung eines koordinierten Verbunds des Nationalen Hochleistungsrechnens durch Bund und Länder verständigt, um der

zunehmenden Bedeutung und der steigenden Nachfrage nach Hochleistungsrechnern gerecht zu werden. Insgesamt wurden acht Rechenzentren in die Förderung aufgenommen; die Fördermittel stehen für Investitionen, Betriebs- und Personalkosten zur Verfügung. Ein Schwerpunkt im NHR-Verbund wird die Stärkung der Methodenkompetenz durch koordinierte Aus- und Weiterbildung der Nutzerinnen und Nutzer sowie insbesondere des wissenschaftlichen Nachwuchses sein.

Die sieben Bundesländer Berlin, Brandenburg, Bremen, Hamburg, Mecklenburg-Vorpommern, Niedersachsen und Schleswig-Holstein sind seit 2001 (Brandenburg seit 2012) im Norddeutschen Verbund zur Förderung des Hoch- und Höchstleistungsrechnens (HLRN) zusammengeschlossen, der Bestandteil der nationalen HPC-Infrastruktur ist. Der HLRN-Verbund betreibt gemeinschaftlich ein verteiltes Supercomputersystem an der Universität Göttingen und dem Zuse-Institut Berlin.

Die Universität Göttingen verfolgt als Teil ihrer Standortstrategie einen Ausbau der Informatik, der Datenwissenschaften und des Hochleistungsrechnens. Hochleistungsrechnen gilt als Schlüsseltechnologie für wissenschaftlichen und technischen Fortschritt.

Göttingen Becomes a Center for National High-Performance Computing

The GWDG as the university computing center for the University of Göttingen was selected as one of eight computer centers for the newly founded National High Performance Computing Alliance (NHR). This means that the University of Göttingen and the GWDG will receive funding of 72 million euros over ten years and will be able to further push forward the expansion of high-performance computing – key technology for scientific and technical progress – as a central component of its location strategy.

Hierzu gehört der Zugang zu großen Rechnerkapazitäten, aber auch die Entwicklung von geeigneter Software. Zur Standortstrategie gehört auch eine Kooperation mit dem Deutschen Zentrum für Luft- und Raumfahrt (DLR), in der ab 2021 wesentliche Anteile der Investitionen des DLR in Hochleistungs-Rechnersysteme am Standort Göttingen gebündelt werden

Die GWDG betreibt neben Emmy, dem Göttinger HLRN-IV-System, mit dem „Scientific Compute Cluster“ einen weiteren Hochleistungsrechner für den Göttinger Wissenschaftsstandort und wird auch den geplanten DLR-Hochleistungsrechner betreiben. In der aktuellen TOP500-Liste der schnellsten Rechner der Welt (<https://www.top500.org>) belegt Emmy Platz 47 weltweit. In Deutschland erreicht Emmy Rang fünf und ist damit der

leistungsfähigste Rechner in Norddeutschland (siehe hierzu auch den Artikel in dieser Ausgabe der GWDG-Nachrichten).

Weitere Informationen

- Presseinformation des Niedersächsischen Ministeriums für Wissenschaft und Kultur vom 13.11.2020: <https://www.mwk.niedersachsen.de/startseite/aktuelles/presseinformationen/niedersachsen-wird-teil-des-deutschen-hochleistungsrechnens-194549.html>
- Pressemitteilung 11/2020 der Gemeinsamen Wissenschaftskonferenz vom 13.11.2020: <https://www.gwk-bonn.de/fileadmin/Redaktion/Dokumente/Pressemitteilungen/pm2020-11.pdf> ■



GWDG Pad

Kollaboration leicht gemacht!

Ihre Anforderung

Sie möchten allein oder gemeinsam mit Ihrem Team unkompliziert an Textdokumenten arbeiten oder Präsentationen erstellen und dabei auf eine Vielzahl nützlicher Funktionen zurückgreifen. Ihre Änderungen sollen sowohl für Sie als auch Ihre Teamkolleg*innen direkt und in Echtzeit einsehbar sein. Sie möchten die Lese- oder Schreibberechtigung für Ihre Dokumente einschränken können, sodass Sie Ihre Daten vor unbefugtem Zugriff schützen können. Zudem wollen Sie in der Auswahl Ihrer Endgeräte flexibel sein, sowohl mobiler Zugriff als auch Desktop-Varianten sollen unterstützt werden.

Unser Angebot

Auf Basis der freien Software CodiMD bieten wir Ihnen einen Dienst, mit dem Sie schnell und unkompliziert Dokumente erstellen, mit anderen Personen teilen und gemeinsam bearbeiten können.

Ihre Vorteile

- > Kollaborativer Echtzeit-Editor
- > Übersicht über alle Ihre Dokumente nach Login

- > Unterstützung von UML-Diagrammen, mathematischen Formeln, Syntax-Highlighting, Musiknoten und vielem mehr
- > Modi zum Erstellen und Vorführen von Präsentationen
- > Einbinden externer Ressourcen wie Videos, PDF-Dateien oder SlideShare
- > Autovervollständigungs-Funktion für Markdown-Ausdrücke
- > Zugriffsbeschränkungen für jedes Dokument einstellbar
- > Veröffentlichung von Dokumenten möglich
- > Webbrowser ausreichend zur Benutzung, keine weitere Installation von Software nötig

Interessiert?

Wenn Sie unseren Dienst „GWDG Pad“ unter <https://pad.gwdg.de> nutzen möchten, benötigen Sie lediglich einen aktuellen Webbrowser. Um eigene Dokumente erstellen zu können, ist zusätzlich die Verwendung eines gültigen GWDG-Accounts oder die einmalige Registrierung unter <https://www.gwdg.de/registration> oder erforderlich.

>> www.gwdg.de/gwdg-pad

Kurz & knapp

Neue Standard-Raumeinstellung bei Aufzeichnungen in BigBlueButton

Seit dem 22. Oktober 2020 bietet Greenlight in den Raumeinstellungen der von der GWDG angebotenen Videokonferenzlösung „BigBlueButton“ (<https://meet.gwdg.de>) die Option „Aufnahme des Raumes erlauben“ an. Wenn diese Option nicht aktiviert ist, ist es nicht möglich, eine Aufnahme zu starten. Bei allen bisher bestehenden Räumen wurde die Aufnahmefunktion deaktiviert und muss daher manuell aktiviert werden.

Sie können diese Option sowohl bei bestehenden Räumen als auch bei neuen Räumen aktivieren.

1. Klicken Sie dazu auf die drei Punkte neben einem bestehenden Raum und dann auf „Raumeinstellungen“ oder klicken Sie alternativ „Raum erstellen“.
2. Aktivieren Sie „Aufnahme des Raumes erlauben“.

Wenn die Option „Aufnahme des Raumes erlauben“ in den Raumeinstellungen ausgewählt ist, erscheint nach Betreten des Raumes der Button „Aufzeichnung starten“. Wird dieser gedrückt, wird die Sitzung aufgezeichnet. Um die Aufzeichnung zu unterbrechen, drücken Sie den Button bitte erneut.

Wenn Sie die Einstellungen ändern, während eine Sitzung läuft, müssen sie die Sitzung zunächst beenden und dann neu starten, damit der Button zur Aufnahme angezeigt wird.

Wenn Sie die Option „Aufnahme des Raumes erlauben“ deaktivieren, bleiben selbstverständlich alle bisherigen Aufzeichnungen des Raumes erhalten.

Warum ist „Aufnahme des Raumes erlauben“ in Greenlight (meet.gwdg.de) nicht immer aktiviert?

Um Aufnahmen durchzuführen zu können, ist es bei BigBlueButton technisch erforderlich, dass der Server die Sitzung zwischenspeichert, auch wenn der Button „Aufzeichnung starten“ nicht betätigt wurde. Auch wenn die GWDG bei Nicht-Aufzeichnung diese Daten automatisiert nach der Sitzung löscht, sollte die Erhebung von persönlichen Daten grundsätzlich auf ein Minimum beschränkt sein. Daher hat sich die GWDG dazu entschieden, die Möglichkeit zur Aufzeichnung in einem Raum mit einem Opt-In zu versehen.

Hellkamp

Lars Bernard (TU Dresden) in den Vorsitz des Rates für Informationsinfrastrukturen gewählt – Petra Gehring und Stefan Liebig im Amt bestätigt

(RfII-Pressmitteilung vom 09.11.2020)

Der Rat für Informationsinfrastrukturen (RfII) hat auf seiner virtuell ausgerichteten 19. Sitzung am 5./6. November 2020 einen neuen Vorsitz gewählt. Die beiden amtierenden Vorsitzenden, Petra Gehring und Stefan Liebig (Stellvertretung), wurden in ihren Ämtern bestätigt. Ergänzend wurde mit Lars Bernard erstmals ein zweiter Stellvertreter hinzugewählt. Alle drei gehören dem RfII seit 2014 an.

Petra Gehring ist Professorin für Philosophie an der TU

Darmstadt und seit 2017 Vorsitzende des Beratungsgremiums. Stefan Liebig ist Direktor des SOEP, Vorstandsmitglied im DIW Berlin und hält eine Professur für Soziologie an der Humboldt-Universität Berlin. Seit 2017 ist er stellvertretender Vorsitzender des RfII.

Lars Bernard ergänzt den Vorsitz als Inhaber einer Professur für Geoinformatik an der TU Dresden. Seit August 2020 bekleidet er auch das Amt des Chief Officer Digitalisierung und Informationsmanagement und ist Mitglied des erweiterten Rektorats seiner Universität. Zuvor war er Sprecher des Bereichs Bau und Umwelt sowie Dekan der Fakultät Umweltwissenschaften der TU Dresden.

Die 19. Sitzung markierte den Beginn der zweiten Hälfte der aktuellen Mandatsperiode des RfII. Dies haben die Ratsmitglieder zum Anlass genommen, über Stand und Perspektiven der RfII-Arbeit zu diskutieren. „Wir sehen einen Druck zur Beschleunigung des digitalen Wandels, der eine schwer überschaubare Zahl immer neuer Initiativen hervorbringt. Hier sieht der RfII seine Stärke darin, Impulse für die Selbstorganisation zu geben und Brücken zwischen den Akteuren zu schlagen“ so Petra Gehring. Die Gründung der NFDI, der Nationalen Forschungsdateninfrastruktur, die auf eine Empfehlung des RfII zurückgeht, sei ein wichtiger Schritt zur Konsolidierung eines Teilbereichs innerhalb der Wissenschaft gewesen. Aber auch jenseits des Wissenschaftssystems finden Parallelprozesse für den Aufbau von Dateninfrastrukturen statt, sowohl national wie europäisch, zu denen die Wissenschaft Anschluss halten und Positionen entwickeln muss. In diesen Prozessen will der RfII weitere Impulse setzen.

In der 19. Sitzung wurden zudem die Ergebnisse einer Studie zur Nutzung und Verwertung von Forschungsdaten beraten sowie der Stand einer Länderanalyse, der die Implikationen der EOSC-Gründung in den EU-Mitgliedsstaaten beschreibt. Weitere aktuelle Arbeiten im RfII betreffen Ausprägungen des Konzepts Datentreuhänderschaft und den Wandel von Sammlungen.

Der Rat für Informationsinfrastrukturen (RfII)

Der RfII wurde von der Gemeinsamen Wissenschaftskonferenz (GWK) eingerichtet. Die 24 Mitglieder des RfII sind durch die GWK als Vertretungen für Informationsinfrastruktureinrichtungen, für die Forschung, den Bereich Öffentliches Leben sowie von Bund und Ländern berufen worden. Aufgabe des Rates ist es, die Transparenz der Entwicklungen und Prozesse auf dem Gebiet der Informationsinfrastrukturen zu erhöhen sowie die Entwicklung und Vermittlung deutscher Positionen in europäischen und internationalen Debatten zu unterstützen.

Kontakt

Rat für Informationsinfrastrukturen – Geschäftsstelle

Dr. Barbara Ebert

Tel.: 0551-39-20951

E-Mail: barbara.ebert@rfii.de

Otto

E-Mail-Verschlüsselung mit X.509-Zertifikaten – Teil 5: Thunderbird, Notes und Mutt

Text und Kontakt:

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

In den ersten beiden Teilen wurde beschrieben, wie X.509-Zertifikate beantragt, gesichert und installiert werden. Im dritten Teil wurde mit den Microsoft Outlook E-Mail-Anwendungen und im vierten Teil für macOS Mail sowie iOS/iPad Mail gezeigt, wie E-Mails signiert und/oder verschlüsselt werden. In diesem Teil wird nun für Thunderbird, IBM Notes und zu guter Letzt Mutt beschrieben, wie mit diesen Anwendungen E-Mails signiert und/oder verschlüsselt werden.

THUNDERBIRD

Wie das Zertifikat in den Thunderbird-eigenen Zertifikatspeicher importiert werden kann, wurde im Teil 2 der Artikelserie in den GWDG-Nachrichten 1-2/2020 auf Seite 16 erläutert.

Um zum Menüeintrag „Konten-Einstellungen“ zu gelangen, gibt es zwei Möglichkeiten: erstens die Alt-Taste drücken und in der am oberen Fensterrand erscheinenden Menüleiste „Extras > Konten-Einstellungen“ anklicken oder zweitens oben rechts in der Ecke das sogenannte „Hamburger“-Menü-Symbol anklicken und in dem dann erscheinenden Menü „Einstellungen > Einstellungen“ anklicken (siehe Abbildungen 1 und 2) Wenn mehrere E-Mail-Konten eingerichtet sind, das entsprechend zu konfigurierende E-Mail-Konto in der Navigationsspalte links auswählen. In den dort aufgelisteten Einträgen zum ausgewählten Konto „S/MIME-Sicherheit“ anklicken.

Um nun für die digitale Unterschrift (Signierung) ein Zertifikat zu bestimmen, auf die Schaltfläche „Auswählen...“ in

der Gruppe „Digitale Unterschrift“ klicken (siehe Abbildung 3) und im angezeigten Dialog das vorher importierte Zertifikat auswählen. Wenn bisher kein Zertifikat für die Signierung/Verschlüsselung angegeben worden ist, bietet Thunderbird nun an, das Signaturzertifikat auch für die Verschlüsselung zu verwenden (siehe Abbildung 4) .An dieser Stelle auf „Ja“ klicken. Wenn gewünscht, noch die Auswahlmöglichkeit „Nachrichten digital unterschreiben (als Standard)“ auswählen, damit alle E-Mails von nun an signiert, also digital unterschrieben, versendet werden.

Nachdem alle Einstellungen für das ausgewählte E-Mail-Konto getroffen worden sind, sieht der fertig ausgefüllte Dialog wie in Abbildung 5 dargestellt aus. Zum Schließen und Speichern der Einstellungen den Dialog mit einem Klick auf „OK“ schließen.

Anmerkung: Soll ein anderes Zertifikat zur Verschlüsselung von E-Mails verwendet werden, in der Gruppe „Verschlüsselung“ auf die Schaltfläche „Auswählen...“ klicken und das entsprechende Zertifikat auswählen und auf „OK“ klicken.

Wenn auf eine E-Mail geantwortet oder gar eine neue E-Mail

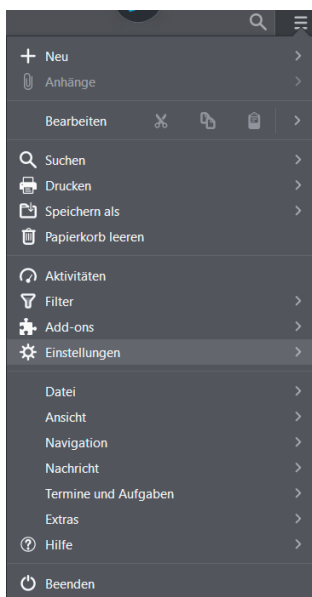


Abb. 1

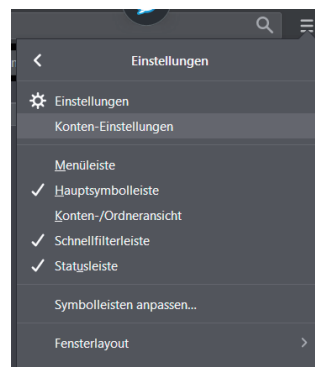


Abb. 2

E-Mail-Encryption Using X.509 Certificates – Part 5: Thunderbird, Notes and Mutt

The first two parts of this multipart article described how to request, secure and install X.509 certificates. The third part described the general process of e-mail signing and encryption and showed how e-mails are signed and/or encrypted for Microsoft Outlook e-mail applications. The fourth part explains for macOS Mail and iOS/iPadOS Mail how to sign and/or encrypt e-mails with these applications. In this part we will now describe for Thunderbird, IBM Notes and last but not least Mutt how to sign and/or encrypt e-mails with these three applications.

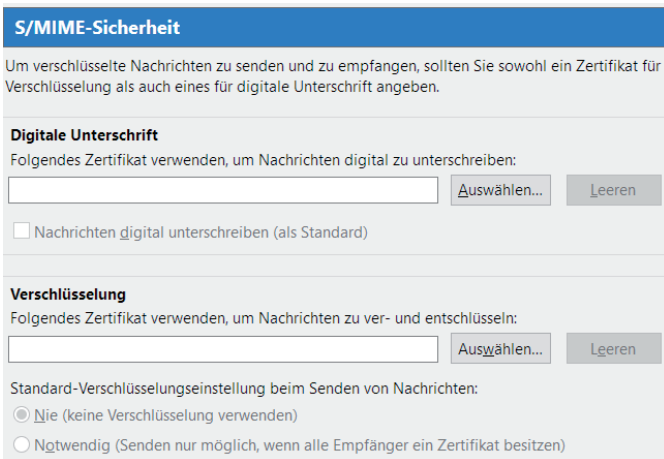


Abb. 3

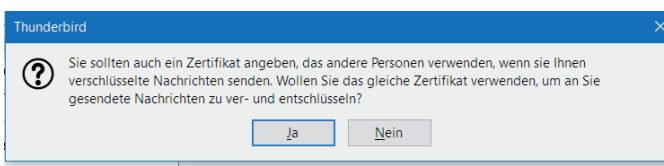


Abb. 4

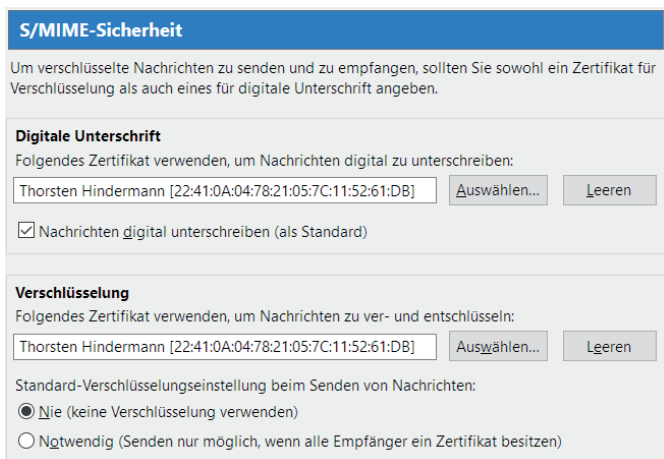


Abb. 5

verfasst wird, die folgenden Aktionen in Thunderbird ausführen.: In dem neu erscheinenden Bearbeitungsfenster für die E-Mail auf den Menüeintrag „Ansicht“ klicken. Ist ein Haken neben der „Kontakte-Sidebar“ sichtbar, ist dieser Bereich links neben dem Bearbeitungsfenster zu sehen. Falls der Haken noch nicht gesetzt ist, einfach auf „Kontakte-Sidebar“ klicken (siehe Abbildung 6).

Über den Menüeintrag „Optionen“ ist es nun möglich, die Auswahl zu treffen, ob die E-Mail signiert (digital unterschrieben) und noch zusätzlich verschlüsselt werden soll. Um diese Auswahlmöglichkeit ein und auszuschalten, in dem Optionen-Menü „Nachricht unterschreiben“ bzw. „Nachricht verschlüsseln“ anklicken. Ist die Auswahl aktiv, ist daneben ein Haken zu sehen, andernfalls ist die Möglichkeit deaktiviert. Optisch sind zusätzlich in der Statuszeile im Bearbeitungsfenster ggf. noch zwei Symbole zu sehen, je nach Auswahl (siehe Abbildungen 7 und 8).

Wird nun der öffentliche Schlüssel eines E-Mail-Empfängers für die Verschlüsselung einer E-Mail gebraucht, der diesen im DFN-LDAP-Verzeichnisdienst abgelegt hat, in der Kontakte-Seitenpalte unter „Adressbuch:“ den zuvor (im Teil 2 der Artikelseite in den GWDG-Nachrichten 1-2/2020 auf Seite 19) eingerichteten Eintrag „DFN-PKI“ auswählen. Wahlweise im Eingabefeld für

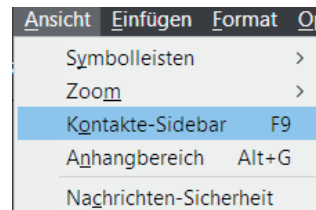


Abb. 6

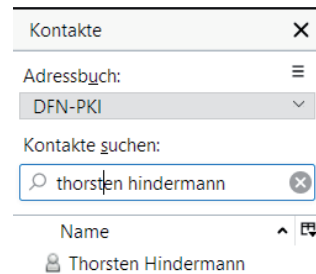


Abb. 9

„Kontakte suchen:“ den Namen oder die E-Mail-Adresse des E-Mail-Empfängers eingeben. Wird dieser angezeigt, den Eintrag anklicken und im unteren Bereich auswählen, ob der Eintrag für „An:“, „Kopie (CC):“ oder „Blindkopie (BCC):“ verwendet werden soll (siehe Abbildung 9).

IBM NOTES 9 UND 10

E-Mail signieren

Im eigenen Postfach eine neue E-Mail beginnen. Minimum: Empfänger-E-Mail-Adresse und einen Test eingeben. Im einfachsten Fall die vier magischen Buchstaben „Test“.

Über dem Eingabefeld für die Empfänger auf „Zustelloptionen...“ klicken (siehe Abbildung 10).



Abb. 10

In dem daraufhin erscheinenden Dialogfeld darauf achten, dass das Häkchen bei „Signieren“ angehakt ist (siehe Abbildung 11).

Weiterhin wird unter dem Eingabefeld für den Betreff auch noch der Hinweis in Grau in kleiner Schrift angezeigt (siehe Abbildung 12).



Abb. 12

Abb. 11

Aufgrund der Einstellungen unter den Sicherheitseinstellungen sollte dieser Haken nun im Standard für jede ausgehende E-Mail gesetzt sein. Nun nur noch auf „Senden“ klicken. Damit wird die Internet-E-Mail mit einer X.509-Signatur aus Lotus Notes an den Empfänger gesendet.

E-Mail verschlüsseln

Um eine E-Mail verschlüsselt an einen Empfänger zu senden, muss als erstes der öffentliche Schlüssel des Empfängers in das Notes-Adressbuch importiert werden.

Dazu im Arbeitsbereich auf das Adressbuch doppelt klicken

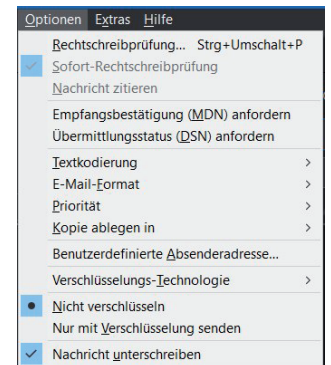


Abb. 7



Abb. 8

und dann auf den Registerreiter „Kontakt suchen...“ in der Adressbuchansicht „Meine Kontakte“ klicken (siehe Abbildung 13).

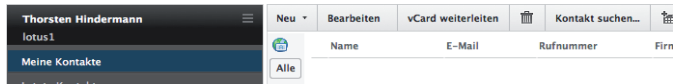


Abb. 13

Der „Personen suchen“-Dialog öffnet sich. In diesem dann in dem Auswahlfeld „Verzeichnis:“ das DFN-Verzeichnis mit der Bezeichnung DFN-PKI auswählen. Dies wurde ja im zweiten Teil des mehrteiligen Artikels beschrieben (siehe die GWDG-Nachrichten 1-2/2020, S. 16). Jetzt den Empfänger suchen, mit dem sensible Daten ausgetauscht werden sollen. Den Namen, Teile des Namens oder die E-Mail-Adresse eingeben und dann auf „Suchen“ klicken. Wenn der entsprechende Empfänger gefunden wurde, diesen per Klick auswählen (siehe Abbildung 14).

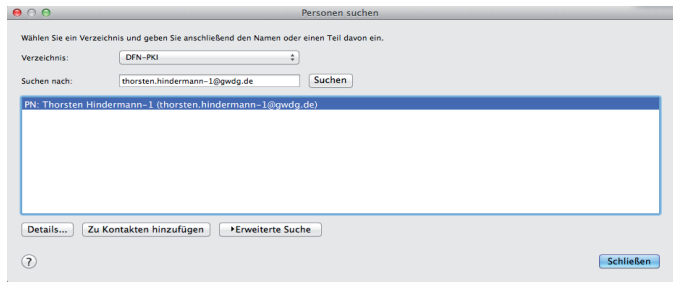


Abb. 14

Mit einem Klick auf „Details...“ können noch Details des Empfängers angesehen werden. Am Wichtigsten ist dabei zu überprüfen, ob das Internetzertifikat vorhanden ist. Dazu auf der/dem mehrfach geteilten Schaltfläche/Registerreiter „Zertifikate“ anklicken und das Vorhandensein des Internetzertifikats prüfen (siehe Abbildung 15).



Abb. 15

Wenn alles in Ordnung ist, diesen Dialog mit einem Klick auf „Schließen“ beenden und jetzt auf „Zu Kontakten hinzufügen“ klicken, damit der ausgewählte E-Mail-Empfänger zu den Kontakten hinzugefügt wird (siehe Abbildung 14).



Abb. 16

Diese Aktion wird mit dem in Abbildung 16 dargestellten Hinweis quittiert, der mit einem Klick auf „OK“ bestätigt wird.

Nun den „Personen suchen“-Dialog mit einem Klick auf „Schließen“ beenden (siehe Abbildung 14).

Das Adressbuch sieht in diesem Fall nun wie folgt aus (siehe Abbildung 17).

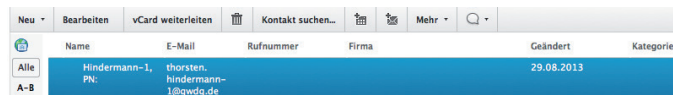


Abb. 17

Und die Detailsansicht zum ausgewählten Kontakt sieht entsprechend Abbildung 18 aus.



Abb. 18

Nun das Bearbeitungsfenster für eine neue E-Mail öffnen. Auf die/den anklickbare(n) Schaltfläche/Link „An:“ klicken. Der „Adresse auswählen“-Dialog öffnet sich. Den E-Mail-Empfänger auswählen und durch Klick auf „An >“, „Kopie >“ und „Blindkopie >“ entsprechend einordnen. Wenn der/alle Empfänger ausgewählt und eingeordnet ist/sind, den Dialog mit einem Klick auf „OK“ schließen (siehe Abbildung 19).

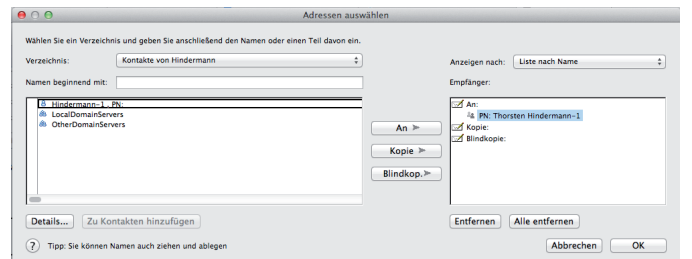


Abb. 19

Anmerkung: In der Auswahl „Verzeichnis:“ sollte das eigene Adressbuch ausgewählt sein, dass die Bezeichnung „Kontakte von <Nachname>“ hat.

Auf dem Registerreiter mit den Aktionen für die gerade in Arbeit befindliche E-Mail auf „Zustelloptionen...“ klicken. Daraufhin öffnet sich der „Zustelloptionen“-Dialog (siehe Abbildung 20).

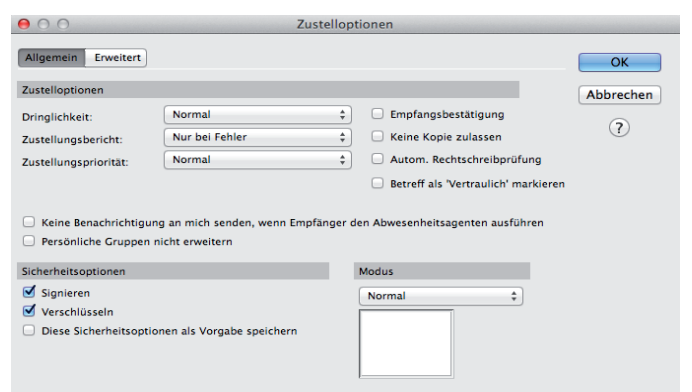


Abb. 20

In diesem Dialog nun noch unter der Gruppe „Sicherheitsoptionen“ die Auswahl „Verschlüsseln“ anhaken und den Dialog mit „OK“ schließen. Als Bestätigung sind die eingestellten Sicherheitsmöglichkeiten unter dem/den Empfänger- und Betreff-Eingabefeld(ern) noch einmal als Textausgabe sichtbar (siehe Abbildung 21).

Wenn jetzt alles eingestellt ist und die E-Mail vollständig geschrieben ist, wird sie mit einem Klick auf „Senden“ entsprechend den Einstellungen abgesichert zum Empfänger gesendet. Das Ergebnis dieser Aktion ist dann in der

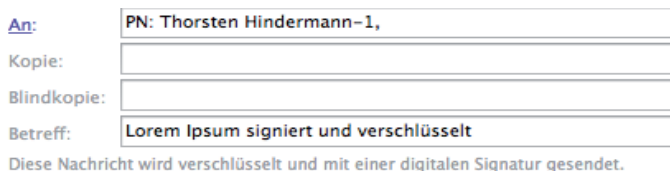


Abb. 21

Thorsten.Hindermann@lotus1.gwdg.de 29. August 2013 11:15
 An: Thorsten [TEST] Hindermann [Details ausblenden](#)
 Lorem Ipsum signiert und verschlüsselt
 Sicherheit: Verschlüsselt, Signiert (Thorsten Hindermann)

Lorem Ipsum
 Mit freundlichen Gruessen/With best regards/Med vänliga hälsningar

Abb. 22

Notes-E-Mail-Anwendung in Abbildung 22 dargestellt.

Erster Empfang einer signierten/verschlüsselten E-Mail

Wenn nun die erste mit einem X.509-Zertifikat signierte E-Mail empfangen wurde und diese angeklickt wird, ist es notwendig, für die erhaltenen Informationen mit der eigenen Notes-ID ein Gegenzertifikat auszustellen. Dazu einfach den angezeigten Dialog mit „Gegenzertifizieren“ bestätigen (siehe Abbildung 23).

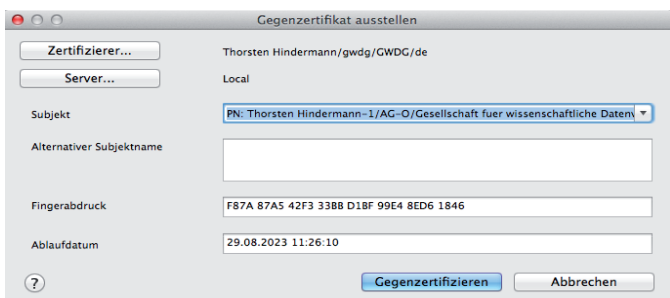


Abb. 23

In der Notes-Statuszeile wird dann ab sofort beim Klick auf signierte E-Mails in der Meldungszeile nur noch der in Abbildung 24 dargestellte Text angezeigt.

Signiert durch PN: Thorsten Hindermann-1 am 29.08.2013 11:24:47, gemäß Thorsten Hindermann/gwdg/GWDG

Abb. 24

In der empfangenen, signierten E-Mail wird in den Kopfzeilen die Information angezeigt, dass diese E-Mail signiert wurde (siehe Abbildung 25).

Re: Lorem Ipsum signiert und verschlüsselt
Hindermann, Thorsten [TEST] An: Thorsten.Hindermann@lotus1.gwdg.de
 Diese Nachricht ist digital signiert.

Abb. 25

Der Empfang einer signierten und/oder verschlüsselten E-Mail funktioniert genau so wie der Empfang einer nur signierten E-Mail. In den Details der Kopfinformationen zur empfangenen E-Mail kann entnommen werden, ob die E-Mail signiert und/oder verschlüsselt wurde. Dazu einfach rechts außen in den Kopfinformationen auf „Details anzeigen“ klicken. „Details anzeigen“ wandelt sich im Moment des Klicks um zu „Details verbergen“ (siehe Abbildung 26).

Re: Lorem Ipsum signiert und verschlüsselt
Hindermann, Thorsten [TEST] An: Thorsten.Hindermann@lotus1.gwdg.de
 Diese Nachricht ist digital signiert.
 Diese Nachricht ist verschlüsselt und digital signiert.

Abb. 26

MUTT

Zum Abschluss diese Artikels wird noch die Anwendung von X.509-Zertifikaten beim textbasierten E-Mail-Client Mutt beschrieben. Der Text stammt aus dem Artikel „Mutt in zwei Zügen“ aus den GWDG-Nachrichten 12/2013 und wird hier minimal verändert abgedruckt.

Damit Mutt erkennt, dass für das Versenden und Empfangen von E-Mails S/MIME-Zertifikate und -Schlüssel verwendet werden sollen, benötigt es entsprechende Einträge in der Konfigurationsdatei *muttrc*. Am einfachsten ist es, man kopiert dazu die Datei */usr/local/share/examples/mutt/smime.rc* in die Directory *~/mutt* und fügt am Ende der Datei *~/mutt/muttrc* die folgende Zeile ein:
`source ~/mutt/smime.rc`

Mehr ist zunächst einmal nicht zu tun. Zur Integration von Zertifikaten und Schlüsseln in die Mutt-spezifische Datenbank stellt Mutt ein eigenes Programm zur Verfügung, nämlich *smime_keys*. Dessen Funktionalität lässt sich wieder über *man smime_keys* ergründen. Als erstes muss einmalig die Mutt-spezifische Datenbank initialisiert werden, und zwar mit Hilfe des Kommandos *smime_keys init*.

Dies legt die Directory *~/smime* mit den Sub-Directories *certificates* und *keys* und zugehörige *.index*-Files an. In *certificates* werden später die öffentlichen und in *keys* die privaten Schlüssel abgelegt. Zum ersten Abschluss wird die Datei */usr/local/share/examples/mutt/ca-bundle.crt* in die Directory *~/smime* kopiert.

Import der eigenen Schlüssel

Damit eigene E-Mails vom Absender signiert werden können, müssen sowohl der eigene private als auch der eigene öffentliche Schlüssel in die Mutt-spezifische Datenbank integriert werden. Dazu extrahiert man seine Schlüssel beispielsweise in die Datei *meineschluesel.p12*. Dieses Verfahren ist ausführlich in den GWDG-Nachrichten 12/2019 im Artikel „E-Mail-Verschlüsselung mit X.509-Zertifikaten – Teil 1“ auf Seite 11 im Abschnitt „Sicherung von Zertifikaten aus dem Internet Explorer“ sowie für andere Browser in den GWDG-Nachrichten 4-5/2020 auf Seite 28 im Abschnitt „Abholung“ beschrieben.

Mit dem Kommando *smime_keys add_p12 meineschluesel.p12* werden die Schlüssel nun in die Mutt-spezifische Datenbank importiert. Sie benötigen dafür das für die Sicherung der Zertifikate verwendete Passwort. Für die weitere Verwendung kann ein neues Passwort festgelegt werden, das dann für die Signierung von E-Mails innerhalb von Mutt verwendet werden muss.

Bitte denken Sie daran: Das gewählte Passwort ist von niemandem wiederherstellbar, falls Sie es vergessen haben sollten!

Im Zuge des Dialogs nach dem obigen Kommando muss noch ein frei wählbarer Name für das Zertifikat festgelegt werden. Weiter wird der Hash-Wert für das importierte Zertifikat angezeigt. Er besteht aus einer hexadezimalen Zahl, gefolgt von *.0*, also beispielsweise *12345678.0*. Diese Zahl muss nun in die entsprechende Zeile der Datei *~/mutt/smime.rc* eingetragen werden:
`set smime_default_key="12345678.0"`

Damit steht der Signierung von eigenen E-Mails nichts mehr

im Wege. Standardmäßig wird nun jede zu versendende E-Mail signiert. Soll eine E-Mail ausnahmsweise nicht signiert oder zusätzlich verschlüsselt werden, so kann vor dem Absenden der E-Mail mit S (Shift-s) eine andere Auswahl getroffen werden.

Import von weiteren öffentlichen Schlüsseln

Um E-Mails zu verschlüsseln, damit sie nur vom Empfänger gelesen werden können, benötigt man den öffentlichen Schlüssel des Empfängers. Diesen kann man sich natürlich vom Empfänger zuschicken lassen, er lässt sich aber auch aus einer vom Empfänger signierten E-Mail extrahieren, da die Signatur auch den öffentlichen Schlüssel des Empfängers enthält.

Mutt stellt für die Extraktion und den Import des öffentlichen Schlüssels in die Mutt-spezifische Datenbank das Kommando `^K` (Strg-K) zur Verfügung. Da aber die Signatur in den meisten Fällen die komplette Zertifikatskette enthält, liefert das Kommando nicht immer ein erfolgreiches Ergebnis. In diesem Fall muss man folgendermaßen vorgehen:

Die E-Mail, aus deren Signatur der öffentliche Schlüssel extrahiert werden soll, wird mit dem Kommando `C` (Shift-c) beispielsweise in die Datei `email.msg` kopiert; dann werden die folgenden openssl-Kommandos ausgeführt:

```
openssl smime -verify -in email.msg -noverify -pk7out > email.pk7
openssl pkcs7 -print_certs -in email.pk7 > email.pem
```

Die Datei `email.pem` enthält nun alle aus der Signatur ermittelten Zertifikate in lesbarer Form. Mit Hilfe eines Editors kann nun das Zertifikat des Absenders beispielsweise in die Datei `email1.pem` separiert werden. Die anderen Zertifikate der Zertifikatskette sind in der Regel bereits durch den Import der eigenen Schlüssel in die Mutt-spezifische Datenbank integriert worden. Nun kann der neue Schlüssel importiert werden: `smime_keys add_cert email1.pem`

Auch hier wird wieder nach einem Namen für das Zertifikat gefragt. Damit ist der Import abgeschlossen.

Mit Hilfe dieses Zertifikats kann man nun auch dem Empfänger eine verschlüsselte E-Mail senden, die nur er mit Hilfe seines privaten Schlüssels lesen kann. ■



Mailinglisten

Mailversand leicht gemacht!

Ihre Anforderung

Sie möchten per E-Mail zu oder mit einer Gruppe ausgewählter Empfänger kommunizieren, auch außerhalb Ihres Instituts. Sie möchten selbstständig eine Mailingliste verwalten, z. B. Empfänger hinzufügen oder entfernen. Bei Bedarf sollen sich auch einzelne Personen in diese Mailingliste einschreiben dürfen.

Unser Angebot

Wir bieten Ihnen einen Listserver, der zuverlässig dafür sorgt, dass Ihre E-Mails an alle in die Mailingliste eingetragenen Mitglieder versendet werden. Die E-Mails werden automatisch archiviert. Das Archiv kann von allen Mitgliedern der Liste nach Schlagwörtern durchsucht werden. Die Anzahl Ihrer Mailinglisten ist unbegrenzt.

Ihre Vorteile

- > Leistungsfähiges ausfallsicheres System zum Versenden von vielen E-Mails
- > Sie senden Ihre E-Mail lediglich an eine Mailinglisten-Adresse, die Verteilung an die Mitglieder der Mailingliste übernimmt der Listserver.

- > Listenmitglieder können an diese E-Mail-Adresse antworten. Eine Moderationsfunktionalität ist verfügbar, mit der Sie die Verteilung einer E-Mail genehmigen können.
- > Voller administrativer Zugriff auf die Einstellungen der Mailingliste und der Listenmitglieder
- > Obsolete E-Mail-Adressen werden vom System erkannt und automatisch entfernt.
- > Wenn Ihre E-Mail-Domäne bei uns gehostet wird, können Sie auch die Adresse der Mailingliste über diese Domäne einrichten lassen.

Interessiert?

Für die Einrichtung einer Mailingliste gibt es zwei Möglichkeiten: Zum einen als registrierter Benutzer der GWGD im Selfservice über das Kundenportal der GWGD und zum anderen, indem Sie bitte eine entsprechende E-Mail an support@gwdg.de senden, die die Wunsch-E-Mail-Adresse der Liste sowie die E-Mail-Adresse der Person, die die Liste bei Ihnen administrieren soll, enthalten sollte. Die administrativen Aufgaben sind leicht zu erlernen.

>> www.gwdg.de/maillinglisten

Virtuell verbunden – 37. DV-Treffen der Max-Planck-Institute

Text und Kontakt:

Organisationsteam des DV-Treffens
dv-treffen@mpg.de

Rege Diskussionen bis in den Abend, neuer Teilnehmerrekord und keine Anreisemühen – das 37. DV-Treffen der Max-Planck-Institute vom 22. bis 24. September 2020 war auch virtuell eine gelungene Veranstaltung.

ERSTER TAG

Das 37. DV-Treffen wurde vom DV-Organisationsteam per BBB (BigBlueButton), das von der GWDG zur Verfügung gestellt wurde, online eröffnet. Zu Beginn konnten über 200 Online-Teilnehmer*innen gezählt werden, die der kompletten Veranstaltung nahezu konstant die Treue halten sollten. Viele beobachteten das DV-Treffen als Stream, und auch dies sollte sich als erfolgreich herausstellen.



DV-Organisationsteam

Die erste Keynote hielt selbstverständlich der Jubilar – die GWDG feierte schließlich in diesem Jahr ihren 50. Geburtstag! Ramin Yahyapour (Geschäftsführer der GWDG) wagte deshalb einen Blick von der Vergangenheit über die Gegenwart bis in die Zukunft. Die Themen waren vielfältig und reichten von HPC über Public Cloud und IT-Sicherheit bis hin zur Abhängigkeit von „Partnern“. Das Thema digitale Souveränität spielt dabei eine immer wichtigere Rolle. Dieser sehenswerte Vortrag war der ideale Einstieg zum ersten Online-DV-Treffen der MPI/MPG-Geschichte.

Die anschließende Keynote von Erwin Laure (Direktor der MPCDF) führte uns in die sich rasant ändernde HPC-Landschaft. Er vermittelte eine Idee zu ExaFlop-Maschinen und zeigte den Weg auf, den die MPCDF in den nächsten Jahren gehen will. Auch wenn Erwin Laure in einer Zeit des Social Distancings zur MPG gekommen ist, wird er mit dem zukünftigen Weg der MPCDF sicherlich für viel „Nähe“ sorgen.

Der IT-Sprecherkreis berichtete über aktuelle Themen aus der

Community und gab einen Ausblick auf das kommende IT-Verantwortlichen-Treffen. Auch dieses wird online auf der gleichen Basis wie das DV-Treffen stattfinden. Anschließend stellten sich dann die Kandidaten des neu zu wählenden IT-Sprecherkreises vor. Turnusmäßig wird hier alle zwei Jahre gewählt, dies erfolgt im Rahmen des DV-Treffens.

Der BAR-Vorsitzende Jochem Marotzke hielt die nächste Keynote zum Thema BAR (Beratender Ausschuss für EDV-Anlagen) und IT. Er ging auf die Kontroverse zu Storage-Systemen in Abhängigkeit zur Stärke der IT-Gruppen detailliert ein und betonte hierbei, dass dieses Thema sehr differenziert und durchaus umstritten im BAR betrachtet wird. Ein weiteres Thema war ein(e) evtl. kommende(r) CIO – dies wird im Präsidentenkreis zum Ende des Jahres abgestimmt. Im Hauptteil sprach er sich ganz klar für die Souveränität in der IT aus. Er führte aus, dass nur wir über unser eigenes IT-Verständnis und die dahinter liegende Arbeit zu entscheiden hätten. Dies führe unweigerlich zu einer digitalen Souveränität der MPG, die wir in Angriff nehmen müssten.

Bei den Lightning Talks ging es Schlag auf Schlag. Die Themen Zulip, EPICS, InIT-Plattform, VA-API, DNSSEC/DANE (DFN), Chromebooks, MPG-RSE und Seminarräume VC-fähig machen wurden kurz aber prägnant vorgetragen. Im Anschluss gab uns Irina Arndt (MPDL) eine aktuelle Übersicht über die Software-Grundversorgung der MPDL.

Auch in diesem Jahr folgte der Vizepräsident der CPTS (Chemisch-Physikalisch-Technische Sektion) unserer Einladung, seine Sicht auf die IT vorzutragen. Klaus Blaum betonte hierbei seine Forderung zu nachhaltiger IT. Besonders die Bestrebungen in

IT Meeting of the Max Planck Institutes

The 37th IT meeting of the Max Planck Institutes went virtual this year and attracted a record number of more than 200 participants. From 22th to 24th September over 30 talks and workshops were presented using the open source video conference and chat infrastructure of GWDG. One of the many highlights was Richard McElreath's talk on "Science as Amateur Software Development".

Richtung Digitalisierung der Institutsverwaltungen lag ihm hierbei am Herzen. Hier soll eine Kommission zur Digitalisierung ins Leben gerufen werden, die auch in Bezug auf Nachhaltigkeit ihre Arbeit aufnehmen soll.

Im Anschluss daran folgte eine Diskussionsrunde, die aus Ariane Rauschek (Hauptabteilungsleiterin BAU/IT), Klaus Blaum, Jochem Marotzke, Jörg Herrmann und Heinz Junkes (beide IT-Sprecherkreis) zusammengesetzt war – Bertin Klein (MPI AGE) moderierte diese Runde. Die besprochenen Themen waren vielfältig, hatten aber klare Schwerpunkte. Zum einen wurde das Thema IT-Personal angesprochen. Hier wurde quasi übereinstimmend bestätigt, dass nur mit qualifiziertem und ausreichendem IT-Personal die notwendige Unterstützung für die Wissenschaft erreicht werden kann. Ein weiteres Thema war der/die CIO. Hierbei wird der IT-Sprecherkreis die Interessen der IT-Community bei der Suche nach einer/m Kandidaten/in vertreten. Weitere Themen waren Open Source und geeignete offene Schnittstellen für die bessere Zusammenarbeit. Schließlich gab es Zustimmung unter den Teilnehmer*innen, dass eine digitale Souveränität der MPG ein wertvolles und notwendiges Ziel sein sollte.

Den offiziellen Schluss bildete die Verleihung des IT Community Awards, der jedes Jahr eine verdiente Person innerhalb der IT Community würdigt. Dieses Jahr ging der Preis nach Berlin an Heinz Junkes (FHI). Wir gratulieren ganz herzlich!



IT Community Award 2020



Preisübergabe

Mit einer anschließenden freien Diskussionsrunde, die bis weit in den Abend reichte, wurde der erste Tag des DV-Treffens beendet. Zahlreiche Themen wurden weiter diskutiert und ein Grundstein für gegenseitiges Verständnis gelegt.

ZWEITER TAG

Den zweiten Tag eröffnete Richard McElreath mit einem Highlight, das über den Tag hinaus für positiven Gesprächsstoff sorgte. Sein Thema „Science as Amateur Software Development“ ist absolut sehenswert und zum Glück auf YouTube unter https://www.youtube.com/watch?v=zwRdO9_GGhY zu sehen. Es lohnt sich!

Die Workshops – parallele Veranstaltungen zu speziellen Themenbereichen – fanden nach dem Mittag statt und haben viel Anklang gefunden. Ein Block beschäftigte sich mit Videokonferenzsystemen und beleuchtete dabei speziell Jitsi, BigBlueButton (BBB) und Cisco Webex. Auch hier konnte man gut sehen, dass sich die Open-Source-Systeme durchaus mit den Platzhirschen messen können. Auch wenn die eine oder andere Funktion noch nicht so komfortabel funktioniert wie bei den Riesen aus San José, kann man mit persönlichem Engagement und IT-Fachwissen eine stabile und umfassende Lösung für Videokonferenzen schaffen. Andere Workshops, zum Beispiel „Grundlagen Git & Gitlab“, beschäftigten sich mit grundlegenden IT-Themen, die bei der Arbeit immer wichtiger werden. Ein weiteres großes Thema waren wieder Datenbanken. Hier wurden Backend-Funktionen von MySQL und mit Django programmierte Frontends für wissenschaftliche Anwendungen vorgestellt. Auch hier konnte man beispielhaft sehen, wie IT-Wissen und Einsatzfreude wissenschaftliche Projekte positiv befeuern können.

Vor dem Abschluss des Tages haben wir uns dem Thema Mail-Server zugewandt. Auch wenn Chat-Programme und Instant Messaging heutzutage eine hohe Bedeutung haben, ist die Mail nicht wegzudenken. Die Relevanz und der Stellenwert dieses Mediums sind nahezu ungebrochen. Von daher ist es von strategischer Bedeutung, eine verlässliche und nachhaltige Kommunikation über Mail für die Zukunft sicherzustellen. Nach den Geschehnissen um und mit Microsoft in den letzten Jahren kommen immer stärkere Bedenken zu Exchange auf. Ein zukünftiger Zwang in die Cloud könnte für rechtliche Probleme sorgen. Um für einen solchen Fall gewappnet zu sein, hat sich die GWDG mit Open-Source-Alternativen zu Exchange beschäftigt. Mit Kopano, Zimbra und Open-Xchange wurden drei der bekanntesten Groupware-Server genau unter die Lupe genommen. Einige Institute betreiben bereits auf deren Basis seit einigen Jahren erfolgreich den lokalen Mail-Dienst.

Zum Abschluss des zweiten Tages hat Ariane Rauschek von ihren Erfahrungen in den letzten knapp zwei Jahren MPG zum Thema IT, IKT und Generalverwaltung berichtet. Themen wie neue Data Center, die Verwendung von Videokonferenzsystemen in der MPG und vorhandene Diskrepanzen zwischen Direktor*innen, den IT-Gruppen und der Generalverwaltung wurden hierbei angesprochen. Die Notwendigkeit einer guten Kommunikation als Basis für erfolgreiche Zusammenarbeit wurde von ihr herausgestellt. Themen, die für die Wissenschaftler*innen in den Instituten wichtig sind, können hierbei nur von der IT-Community aufgenommen und adressiert werden. Es müssen neue Wege gefunden werden, diese Themen gemeinsam zu bearbeiten. Ein erster Weg könnte

die gemeinsame Suche nach der CIO-Funktion sein. Die anschließende offene Diskussion war, wie die am ersten Tag, lebhaft und knüpfte an viele Themen des Tages an. Dieser freie Gedankenaustausch zeigte in vielen kleinen Aspekten, dass die Generalverwaltung und die IT-Gruppen der Institute gar nicht so weit voneinander entfernt sind. Ein solcher Austausch kann ein Schlüssel zu einer konstruktiven Zusammenarbeit sein. Damit gab diese Diskussion dem Tag ein erfolgreiches und rundes Ende.

DRITTER TAG

Der dritte Tag stand im Zeichen der IT-Community. Den Anfang machte der IT-Sprecherkreis, der während des Treffens neu gewählt wurde. Auch die Wahl fand diesmal elektronisch statt. Alle gewählten Kollegen haben die Wahl angenommen und sind bereit, die Kolleg*innen der lokalen IT-Gruppen zu unterstützen. Der Sprecherkreis wird auch in den nächsten zwei Jahren die Interessen der IT-Community innerhalb der verschiedenen Gremien der MPG (z. B. dem BAR) vertreten.

Verschiedene Projekte (BlueSky und Castellum & huscy) zeigten in detaillierter Form, wie die IT die Wissenschaft unterstützen kann. Hierbei wird, z. B. bei BlueSky, eine komplette Infrastruktur geschaffen, die den wissenschaftlichen Workflow von der Erfassung bis zur Datenhaltung komplett abbildet. Der gesamte Prozess basiert hierbei auf Open-Source-Software.

Ein Erfahrungsbericht aus Sicht der IT in Bezug auf die Corona-Pandemie rundete die Beiträge der IT-Community ab. Auch hier zeigte sich, dass die IT in dieser Ausnahmezeit eine wichtige und entscheidende Rolle eingenommen hat. Ohne das nötige Personal und das vorhandene IT-Know-how hätte in dieser Zeit der quasi „normale“ Institutsbetrieb niemals stattfinden können. Sei es die Bereitstellung von Videokonferenzen, die Umstellungen des virtuellen Verwaltungsarbeitsplatzes oder die Beseitigung von Anbindungsproblemen der Wissenschaftler*innen vom Homeoffice aus – unsere Kolleg*innen haben in dieser schwierigen Zeit einen klaren Job gemacht.

Selbstverständlich wurde auch ein Blick auf Softwareprodukte geworfen, die vielleicht noch nicht bei jedem bekannt sind. Hier kann man Zammad nennen, das Ticket- wie auch Helpdesk-Tool in einem ist. Dies wurde vom Firmengründer, Martin Edenhofer, persönlich präsentiert. Aufgrund seiner langjährigen Erfahrung – 2001 hatte er bereits das Produkt OTRS erfunden – kann dieses Tool bereits vieles „out of the box“ bereitstellen. Mit der ebenfalls offenen Softwarestruktur (Open Source) können aber auch eigene Prozesse abgebildet werden. Es lohnt sich, das Produkt einmal näher zu betrachten. Ein weiterer Blick hinter die Kulissen hat aufgrund unterschiedlicher Zeitzonen leider nicht funktioniert. Ein

Kollege von LibreOffice wollte hier aus England heraus präsentieren, leider zur falschen (lokalen) Zeit. Dies wird aber zu einem späteren Zeitpunkt als „TechTalk“ nachgeholt.

Den Abschlussvortrag hielt Eberhard Bodenschatz (Vorsitzender der Softwarekommission) zum Thema „Competing Demands und Open-Source-Software“. Er nannte die hohen Summen, die für Microsoft-Lizenzen ausgegeben werden, und sprach sich für eine Strategie aus, die Nutzung von Microsoft-Produkten in der MPG zu reduzieren. Er stellte aber auch die unterschiedlichen Interessen und Bedarfe von Nutzer*innen, IT und MPG da, die man bei einem Umstieg von proprietärer zu Open-Source-Software berücksichtigen muss.

Mit dem Abschlusswort des DV-Organisationsteams wurde das diesjährige DV-Treffen beendet. Unsere „lessons learned“ aus drei Tagen DV-Treffen kann man in etwa so zusammenfassen:

Dieses Online-Event hat bewiesen, dass Open-Source durchaus eine solide Basis ist, eine rundum gelungene Veranstaltung über drei Tage mit gut 200 Teilnehmer*innen zu realisieren. Da dieses Format wirklich gut funktioniert hat, werden wir (die IT-Community) die Idee der „TechTalks“ wieder ins Leben rufen – kurze Veranstaltungen (ca. 1 bis 2 Stunden) mit einem thematischen Schwerpunkt. Ideen hierfür und auch „Mitmacher*innen“ sind immer herzlich willkommen. Auch zeigten viele Diskussionen, dass Kommunikation auf respektvoller und konstruktiver Ebene mehr als wichtig ist. Oft sind wir alle gar nicht weit voneinander entfernt. Ein entscheidender Punkt wurde in einigen Vorträgen immer wieder betont: Ohne das notwendige Personal in den lokalen IT-Gruppen und das dazugehörige Know-how funktioniert die Wissenschaft nicht (mehr). IT ist mehr als die Bereitstellung einer virtuellen Maschine in einem entfernten Server. Und ohne lokales IT-Know-how können Anforderungen vor Ort nicht mehr getestet werden. Wo Projekte dann (nach erfolgreichen Tests) realisiert werden, kann dann von Fall zu Fall entschieden werden. Falls aber nicht genug Personal zur Verfügung steht oder dieses kein exzellentes IT-Know-how mehr aufweist, können keine Entscheidungen mehr vor Ort getroffen werden. Dies sollte jedem klar sein. Der letzte Punkt betrifft uns als MPG gemeinsam. Nur wenn die MPG insgesamt digital souverän wird und bleibt, können wir die Unterstützung für die Wissenschaft anbieten, um weiterhin Spitzenforschung zu gewährleisten. In diesem Sinne – Enabling Research! ■



MINE – a Text and Data Mining Platform

Text and Contact:
Dr. Sven Bingert
sven.bingert@gwdg.de
0551 201-2164

The project MINE is a cooperation between the Göttingen State and University Library and the GWDG with the aim to provide a text and data mining platform to work on the text resources at Göttingen Campus.

INTRODUCTION

The central library at Göttingen Campus, the Göttingen State and University Library (SUB), hosts a large collection of (digital) text resources. The traditional paradigm of reading books for analysing them (close reading) is more and more extended by using digital tools to apply new research questions (distant reading). With the new algorithms large data sets can be analysed. But at the current state it is difficult to get access to the resources or to apply a search on all the different resources in parallel. Many different APIs, access conditions, and licenses have to be regarded in order to make the text resources available for further analysis. And even if access to the resources is given, it will be challenging to select a suitable infrastructure to process those. To overcome these issues the MINE project was initiated.



THE IDEA

The aim of the project is to provide a platform to access, retrieve, and process selected texts from different resources. The platform should be easy to use and provide a standard set of text and data mining tools (TDM). These build in tools will help to overcome the access and license issues by allowing only to apply a specific set of operations on the resources. The platform will consist of a web interface and a Python library to be used on command line or Jupyter services utilizing High-Performance Computing (HPC) resources.

We set up a collaboration where we follow the Scrum methodology but within two institutions. E.g. the scrum master is member of the SUB, the product owner is member of the GWDG and the scrum team actually consist of members from both institutions. The project started at the end of 2019, is now in its second phase and aims to deliver the final platform in October 2021. Intermediate steps and data will be available beginning of 2021.

KNOWLEDGE GRAPH

To foster possible research with the given resources the platform should provide more than a simple search interface on the

available meta data. Therefore the MINE system will perform a set of TDM tools automatically on the available resources. Those tools will provide new insights or knowledge, e.g. entities or topics, that we map into a knowledge graph. An example for an extended graph is shown in Figure 1. Of course the complexity of the graph will grow with implementation of different TDM tools. Figure 2 shows a preliminary example graph based on the current state of the system. It shows already the connection of the two text files sharing the same subject.

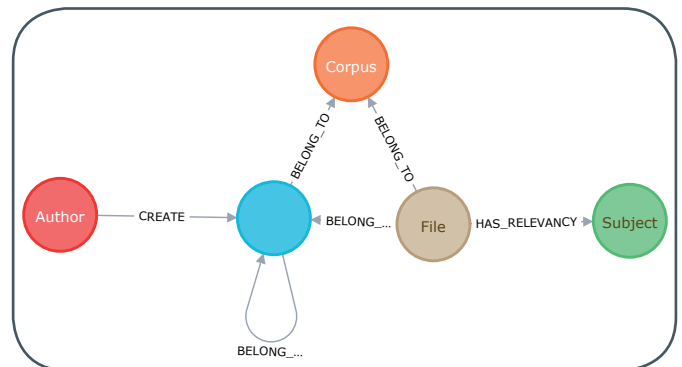


Figure 1: Example schema for the extended knowledge graph with the subjects found by an entity recognition tool. The other elements of the schema, e.g. author, are taken from the initial meta data provided by the text resource itself. The schema is subject of research itself and might change in different versions of the MINE system.

ARCHITECTURE

The focus of the architecture (see Figure 3) is to provide a flexible and fully automated back end to include new resources, update

Das Projekt MINE

Das Projekt MINE ist eine Kollaboration zwischen der Niedersächsischen Staats- und Universitätsbibliothek Göttingen und der GWDG mit dem Ziel, eine Text- und Data-Mining Plattform bereitzustellen. Diese soll den Wissenschaftler*innen erlauben, auf einfache Weise mit den Textquellen am Göttingen Campus zu arbeiten. Durch clevere technische Lösungen sollen Hürden wie Zugangbeschränkungen abgebaut werden. Dies wird z. B. erreicht, indem ausgewählte Algorithmen auf den Daten angewendet werden und dem/der Nutzer*in nur die Ergebnisse dargestellt werden.

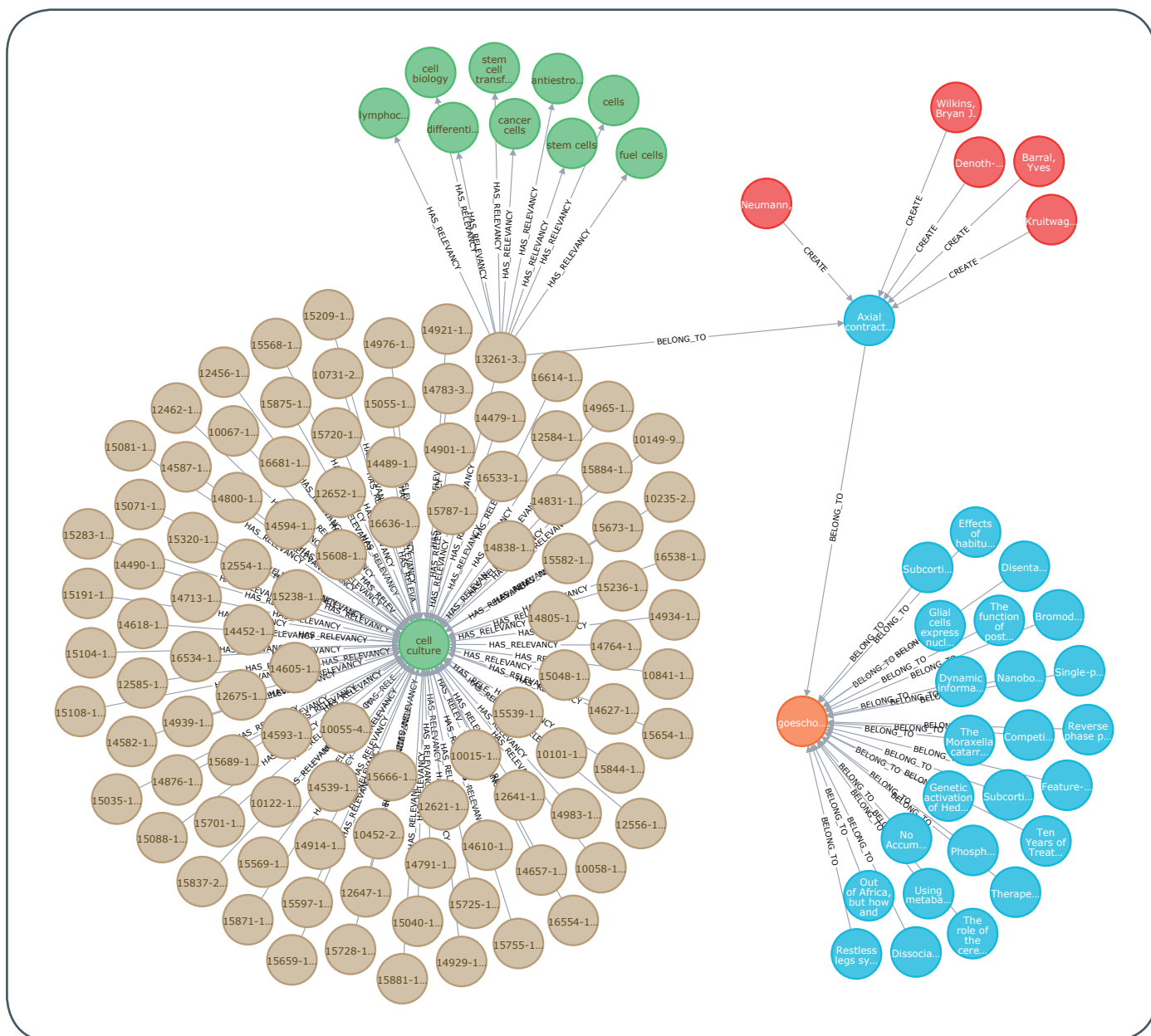


Figure 2: Example graph from the current MINE system. The orange node denotes a collection, the red nodes denote authors, and the brown nodes denote single files. In the example a specific subject (green node) found by the entity recognition tool was found in many different files. The blue nodes denote single publications.

existing resources and apply changes in meta data transformations or new tools for the knowledge graph. Therefore the architecture is separated in the following sub-services:

- **Connectors:** Services to connect to different resources. The connectors are specific for each resource, e.g. using different APIs. They allow to retrieve data and meta data. They also extract and transform selected meta data for the MINE system. Those meta data will be available in the search interface.
- **Knowledge extractor:** A set of tools to gain new information from the resources, e.g. using entity recognition. The new information will also be available in the search interface.
- **Web interface:** The website will allow to search on the meta data and the knowledge graph. It will present the available meta data and an excerpt of the knowledge graph.

- **Python library:** Allows to search and retrieve texts and to run tools on resources that are not directly accessible. Also, it allows to run own TDM tools on the selected texts using the MINE infrastructure.

As the mapping of the meta data or the schema of the knowledge graph is under research itself it is important to make the full process reproducible. All steps in the system are described in workflows and can be executed by a workflow engine either in a whole or separately.

CONCLUSION AND OUTLOOK

The first text resources are available in the MINE system and the entity recognition tool Annif (<http://annif.org>) is used to fill the knowledge graph. The selected resources are Goescholar (the collection of papers produced by Göttingen researchers), TextGrid (TEI encoded German literature works), Times Digital Archive (the OCR

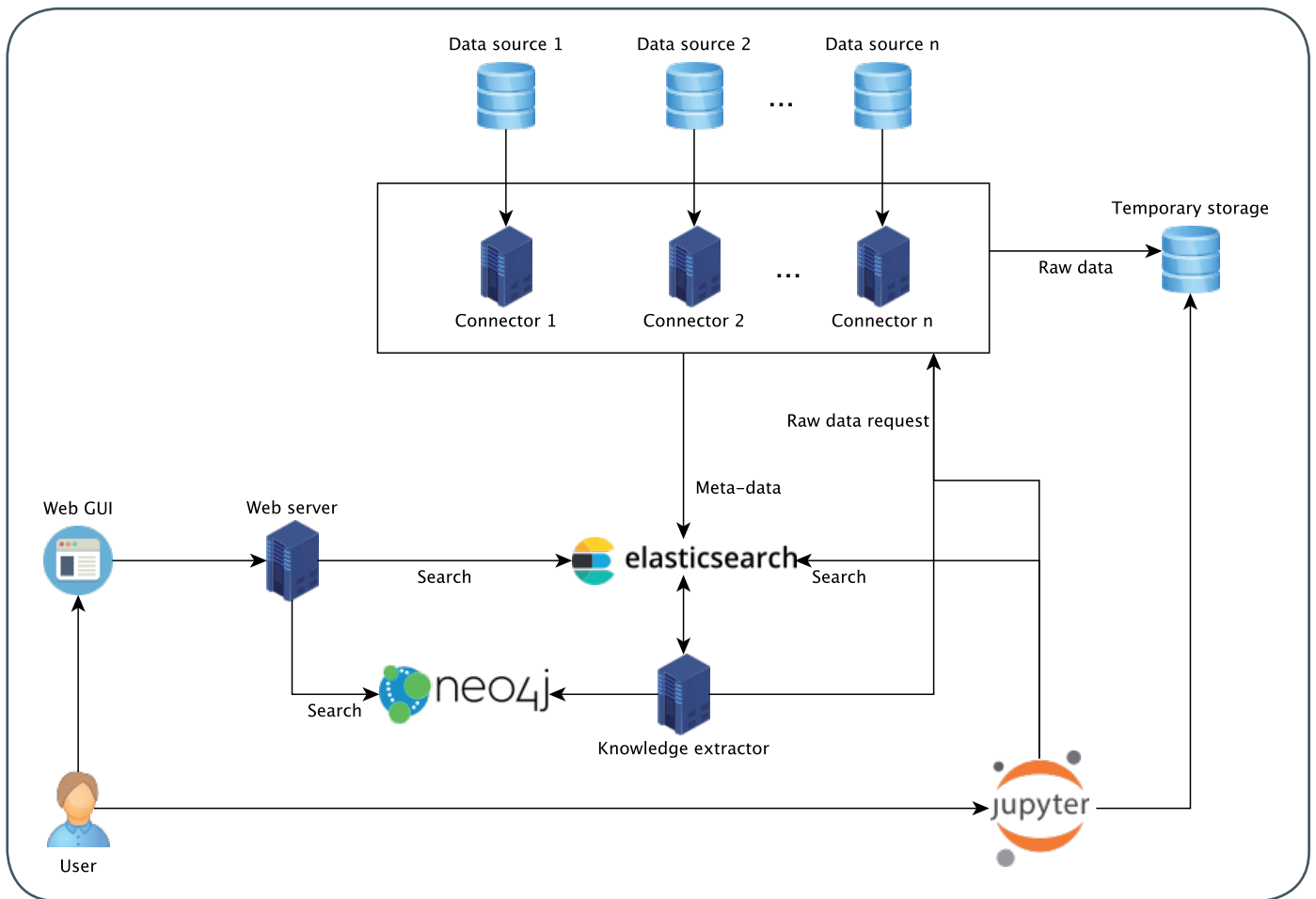


Figure 3: The graph depicts the current MINE architecture. The connectors are used to retrieve (meta-)data from the resources. We use common technologies for maintaining and serving the meta data.

processed pages of the famous English newspapers), and DigiZeitschriften (collection of German research journals). This selection shows already the complexity and the opportunities to investigate new research questions. The web interface will be made public soon. In the next course of the project we will include more text resources

from SUB but will also connect to other libraries and owners of large collections and corpora. The MINE system should be fully operational in October 2021. If you want to be early user to test our current system please contact at mine-team@gwdg.de. ■

BIG DATA ANALYTICS

INFORMATIONEN:
support@gwdg.de
0551 201-1523



Dezember 2020

Academy

KURS	DOZENT*IN	TERMIN	ANMELDEN BIS	AE
OUTLOOK – E-MAIL UND GROUPWARE	Helmvoigt	03.12.2020 9:15 – 12:00 und 13:00 – 16:00 Uhr	26.11.2020	4
ANGEWANDTE STATISTIK MIT SPSS FÜR NUTZER MIT VOR-KENNTNISSEN	Cordes	09.12. – 10.12.2020 9:00 – 12:00 und 13:00 – 15:30 Uhr	02.12.2020	8

Teilnehmerkreis

Das Angebot der GWDG Academy richtet sich an die Beschäftigten aller Einrichtungen der Universität Göttingen, der Max-Planck-Gesellschaft sowie aus wissenschaftlichen Einrichtungen, die zum erweiterten Kreis der Nutzer*innen der GWDG gehören. Studierende am Göttingen Campus zählen ebenfalls hierzu. Für manche Kurse werden spezielle Kenntnisse vorausgesetzt, die in den jeweiligen Kursbeschreibungen genannt werden.

Anmeldung

Für die Anmeldung zu einem Kurs müssen Sie sich zunächst mit Ihrem Benutzernamen und Passwort im Kundenportal der GWDG (<https://www.gwdg.de>) einloggen. Wenn Sie zum Kreis der berechtigten Nutzer*innen der GWDG gehören und noch keinen GWDG-Account besitzen, können Sie sich im Kundenportal unter dem URL <https://www.gwdg.de/registration> registrieren. Bei Online-Kursen kann das Anmeldeverfahren abweichen. Genauere Informationen dazu finden Sie in der jeweiligen Kursbeschreibung. Einige Online-Angebote stehen Ihnen jederzeit und ohne Anmeldung zur Verfügung.

Absage

Absagen können bis zu sieben Tagen vor Kursbeginn erfolgen. Bei kurzfristigeren Absagen werden allerdings die für den Kurs angesetzten Arbeitseinheiten (AE) vom AE-Kontingent der jeweiligen Einrichtung abgezogen.

Kursorte

Alle Kurse finden in geeigneten Räumlichkeiten am Göttingen Campus statt. Den genauen Ort entnehmen Sie bitte der jeweiligen Kursbeschreibung. Auf Wunsch und bei ausreichendem Interesse führen wir auch Kurse vor Ort in einem Institut durch, sofern dort ein geeigneter Raum mit entsprechender Ausstattung zur Verfügung gestellt wird.

Kosten bzw. Gebühren

Die Academy-Kurse sind – wie die meisten anderen Leistungen der GWDG – in das interne Kosten- und Leistungsrechnungssystem der GWDG einbezogen. Die den Kursen zugrundeliegenden AE werden vom AE-Kontingent der jeweiligen Einrichtung abgezogen. Für alle Einrichtungen der Universität Göttingen und der Max-Planck-Gesellschaft sowie die meisten der wissenschaftlichen Einrichtungen, die zum erweiterten Kreis der Nutzer*innen der GWDG gehören, erfolgt keine Abrechnung in EUR. Dies gilt auch für die Studierenden am Göttingen Campus.

Kontakt und Information

Wenn Sie Fragen zum aktuellen Academy-Kursangebot, zur Kursplanung oder Wünsche nach weiteren Kursthemen haben, schicken Sie bitte eine E-Mail an support@gwdg.de. Falls bei einer ausreichend großen Gruppe Interesse besteht, könnten u. U. auch Kurse angeboten werden, die nicht im aktuellen Kursprogramm enthalten sind.



Gesellschaft für wissenschaftliche
Datenverarbeitung mbH Göttingen