



IT-Landkarte der KIT-Software-Anwendungen
IT Map of KIT's Software Applications

DEEP Hybrid DataCloud
DEEP Hybrid DataCloud

MathSEE Modeling Week
MathSEE Modeling Week

Liebe Leserin, lieber Leser,

Landkarten aus Papier bieten zur Orientierung meist eine bessere Übersicht über ein großes Gebiet als digitale Karten-Apps auf Smartphones. Diese entfalten ihre Stärken erst mit Hilfe integrierter Informationsdienste und zeigen beispielsweise, wo ein geöffnetes Restaurant in der Nähe oder wie die Verkehrssituation einer Route ist oder wie man mit öffentlichen Verkehrsmitteln zum Zielort gelangt. Vergleichbar dazu sammelt und integriert eine neue interaktive IT-Landkarte (Seite 4) Informationen über verteilte Dienste und Anwendungen am KIT und stellt die zugehörigen Datenflüsse und Geschäftsobjekte in einen Gesamtzusammenhang. So entsteht nach und nach eine übergreifende Datenbasis für das Management von Diensten und Anwendungen am gesamten KIT.

Auch Forschende möchten sich im immer rascher wachsenden IT-Dickicht von wissenschaftlichen Werkzeugen und Diensten zurechtfinden. Hier schafft unter anderem das Projekt DEEP Hybrid DataCloud Abhilfe. Mit speziell zugeschnittenen Apps und Services sowie einer Vielzahl von Unterstützungsleistungen können Wissenschaftlerinnen und Wissenschaftler unterschiedlichste Fragestellungen mit Deep Learning Methoden adressieren. Diese sind innerhalb einer flexiblen Cloud-Infrastruktur – auch mit Hochleistungsrechenressourcen – verfügbar (Seite 14).

Eine andere Form der Orientierung im großen Themenfeld Modellierung und Simulation gibt die Projektgruppe CAMMP am SCC. Das KIT-Zentrum MathSEE richtete erstmals eine Modellierungswoche aus, um Studierende eines mathematisch-methodischen Faches in die mathematische Modellierung von realen Problemstellungen mit Hochleistungsrechnern einzuführen (Seite 21).

Zusammen mit den am SCC frisch Promovierten (siehe Titelseite) wünschen wir Ihnen eine gute Orientierung durch die neue Ausgabe der SCC-News.

Viel Freude beim Lesen!

Martin Frank, Bernhard Neumair, Martin Nußbaumer, Achim Streit

Dear reader,

for orientation, maps made of paper usually provide a better overview of a large area than digital map apps on smartphones. These develop their strengths only with the help of integrated information services and show nearby open restaurants, the traffic situation along the route or how to reach your destination by public transport. Likewise, a new interactive IT map (page 4) collects and integrates information about distributed services and applications at KIT, putting the associated data flows and objects into context. As a result, an overarching data base for the management of services and applications at KIT is gradually being created.

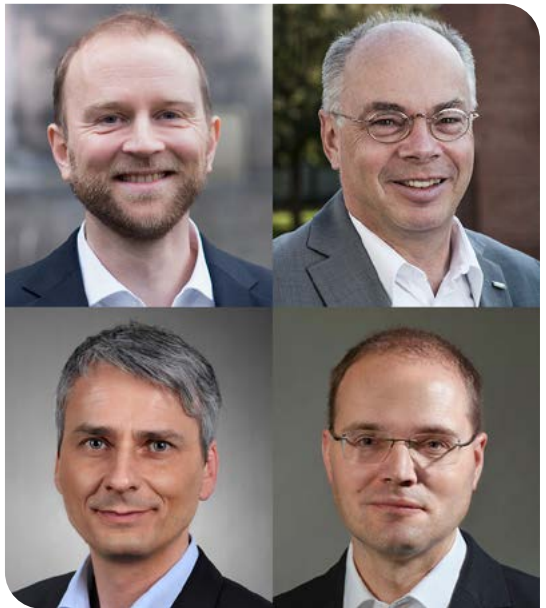
Similarly, researchers want to find their way around scientific tools and services in the ever-growing IT landscape. Among other things, the project DEEP Hybrid DataCloud provides a remedy. Specially tailored apps and services as well as a variety of supporting services enable scientists to address a wide variety of issues using deep learning methods. All of these services are available within a flexible cloud infrastructure, complemented even with high-performance computing resources (page 14).

Another form of orientation is provided by the project group CAMMP at the SCC in the large field of modeling and simulation. For the first time, the KIT Center MathSEE hosted a modeling week to introduce students from disciplines that use mathematical methods to the modeling of real problems using high-performance computers (page 21).

Together with the researchers who just obtained their doctoral degree at SCC (see title photo) we wish you a good orientation through this new edition of the SCC News.

Enjoy reading!

Martin Frank, Bernhard Neumair, Martin Nußbaumer, Achim Streit



Inhaltsverzeichnis

DIENTE UND INNOVATION

- 04 IT-Landkarte der KIT-Software-Anwendungen
- 07 bwSync&Share mit neuer Cloud-Software – ein Anwenderbericht
- 08 bwCloud-SCOPE – Cloud-Infrastruktur erweitert
- 10 VPN2VLAN mit IPv6 im Tunnel

FORSCHUNG UND PROJEKTE

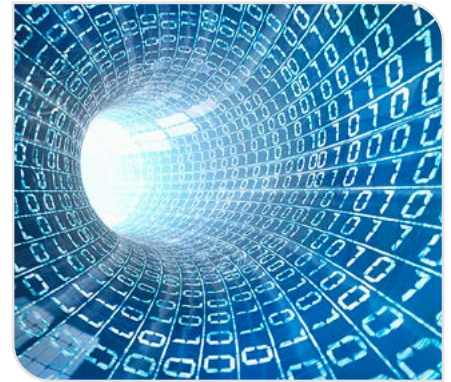
- 12 Erfolgreicher Abschluss der ersten Projektphase von V-FOR-WaTer
- 14 DEEP Hybrid DataCloud
- 16 Die EOSC jenseits der Technologie
- 18 Abgeschlossene Promotionsarbeiten
- 20 Innovationsteam des KIT gewinnt NCATS ASPIRE Design Challenge Preis

STUDIUM UND WISSENSVERMITTLUNG

- 21 MathSEE Modeling Week – Studierende präsentieren Ergebnisse
- 22 Aktionstag am KIT zum European Cyber Security Month 2019
- 23 17. Internationale GridKa School – The Art of Data
- 24 CAMMP in Mexiko

VERSCHIEDENES

- 25 Marie Weiel-Potyagaylo erhält den Erna-Scheffler-Förderpreis 2019
- 26 Bunte Nacht der Digitalisierung – Impressionen
- 27 Neues aus den SCC-Abteilungen
- 27 Impressum & Kontakte



Titelfoto:
Diana Gudu, Nico Schlitter, Ines Reinartz, Felix Bach (v. l. n. r.).

IT-Landkarte der KIT-Software-Anwendungen

Im SCC werden etwa 170 Dienste und rund 100 Fachanwendungen betrieben. Neben reinen Eigenentwicklungen sind darunter auch selbst- oder fremdgewartete Standard- und Spezialanwendungen zu finden. Die vom SCC entwickelte interaktive IT-Landkarte dokumentiert diese Software-Infrastruktur und deren Zusammenspiel, macht die Abhängigkeiten von IT Services und Geschäftsprozessen des KIT erfassbar und dient bei Bedarf als Hilfsmittel zur Optimierung.

Matthias Bonn, Ulrich Weiß

Eine der Aufgaben des SCC besteht darin, die sich permanent ändernden Anforderungen aus den fachlichen und administrativen Prozessen nachhaltig in einer leistungsfähigen IT abzubilden. Hierzu gehören sowohl die Spezifikation von Anwendungen, Diensten, Schnittstellen und Geschäftsobjekten¹ als auch deren Entwicklung, Betrieb, Wartung und Doku-

eigenentwickelte Webapplikationen und Portale oder auch die SAP-Systeme sowie Dutzende speziell angepasster Fachanwendungen aus praktisch allen Bereichen des KIT – Forschung, Lehre, Innovation und Administration. Während einerseits viele Anwendungen für Endnutzer sichtbar sind und von diesen genutzt werden, agieren andererseits manche

Antworten auf viele Fragen

Eines der Hauptziele der IT-Landkarte besteht darin, Anfragen zur heterogenen Anwendungslandschaft des KIT beantworten zu können. Sie durchdringt damit die gesamte IT-Architektur. Es wird erfasst und verwaltet, welche Systeme wie miteinander interagieren, welche Geschäfts-

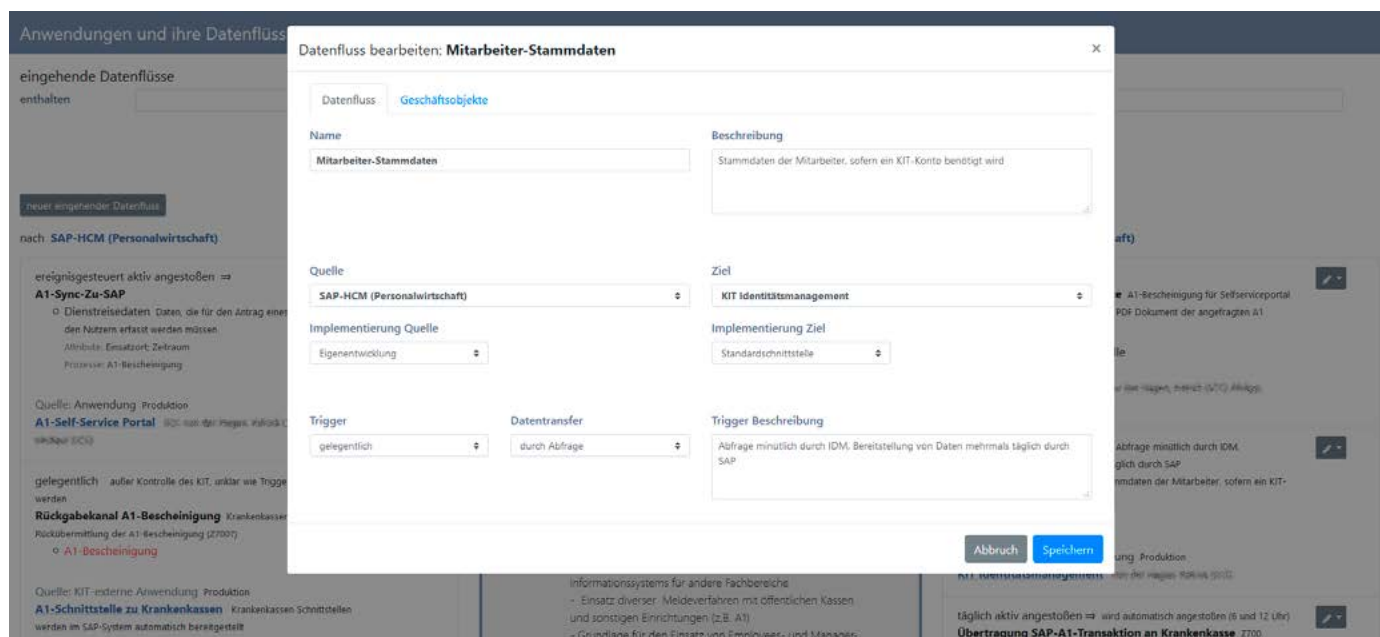


Abbildung 1: Datenfluss bearbeiten

mentation. Dies betrifft nicht nur organisationsübergreifende Dienste wie z. B. das Identitätsmanagement, Groupware, Speicher und Authentifizierungsdienste, sondern auch die Anpassung und den Betrieb von fachspezifischen Anwendungen. Beispiele solcher Anwendungen sind Bewerbungs- und Zulassungsmanagement für Studierende, Doktorandenverwaltung,

Systeme als unsichtbare Datendreh-scheiben im Hintergrund und versorgen eine Vielzahl angebundener Dienste mit Informationen. Weiterhin müssen Daten und Informationen mit Anwendungen ausgetauscht werden, die außerhalb der KIT-Organisationsgrenze angesiedelt sind, wie z. B. Systeme von Krankenkassen, Banken oder Finanzbehörden.

objekte ausgetauscht werden und welche Datenquellen autoritativ sind. Hier bietet sich ein großes Potential, die Qualität der Stammdaten und deren Synchronisation signifikant zu verbessern. Ebenso werden

¹ Ein Geschäftsobjekt bezeichnet allgemein die elektronische Repräsentation einer realen Person, eines Gegenstandes oder einer Verarbeitungslogik.

die Auslöser von Informationsflüssen dargestellt: Beispielsweise können Daten aufgrund einer Nutzeraktion oder durch Automatismen ausgetauscht werden. Dies kann als Grundlage für diverse Aspekte wie IT-Effizienz, IT-Sicherheit, Digitalisierung, IT-Compliance und Datenschutz dienen. Weiterhin können Inselfösungen und – im Falle eines Dienstausfalls oder einer Störung – potentielle Fehlerquellen leichter identifiziert werden. Mit Hilfe der IT-Landkarte wird u. a. auch ersichtlich, wo personenbezogene Daten verarbeitet und ausgetauscht werden, was Dokumentationsanforderungen hinsichtlich der IT-Sicherheit und des Datenschutzes vereinfacht. Schließlich lassen sich Zuständigkeiten wie verantwortliche Einrichtungen und Ansprechpartner erfassen, so dass Support und Wartungsfragen leichter geklärt und Betriebsprozesse insgesamt stabilisiert werden können.

Das SCC organisiert die Erstellung und Verwaltung der IT-Landkarte in Abstimmung mit den beteiligten Fachbereichen. Die Visualisierung wird dabei (teil-)

automatisiert auf Grundlage der erfassten Anwendungen, deren Geschäftsobjekten und Schnittstellen erzeugt. So stellt die IT-Landkarte ein wesentliches Hilfsmittel zur Ausbalancierung von Bedarf und Angebot im KIT dar.

Realisierung und Technik

Die im SCC seit vielen Jahren zur Verwaltung von Diensten, Anwendungen und Wartungsfenstern eingesetzte Configuration Management Database (CMDB) bildet die Datenbasis des IT-Service-Managements und stellt die autoritative Datenquelle dar. Die IT-Landkarte liest regelmäßig u. a. die in der CMDB erfassten Dienste, Anwendungen und Beziehungen zwischen konfigurierten IT-Elementen aus und reduziert die gelieferten Datenobjekte auf die für die Landkarte benötigten Attribute. Darüber hinaus stellt sie Funktionen bereit, die Anwendungen mit Datenflüssen, Schnittstellenbeschreibungen und Auslösern anzureichern und miteinander in einer Quelle/Ziel-Beziehung zu verknüpfen. Zusammen mit

diesen Verknüpfungen werden auch die Geschäftsobjekte definiert, die beschreiben, welche Art von Daten ausgetauscht werden und ob sie einen Personenbezug haben. Zudem lassen sich auch solche Anwendungen in der IT-Landkarte erfassen und verwalten, die weder technisch noch organisatorisch am KIT betrieben werden, so dass auch externe Schnittstellen und deren Beziehung mit KIT-internen Systemen abgebildet werden können.

Mit Hilfe von verschiedenen Filterfunktionen ist es möglich, die Anwendungen beispielsweise nach Zuständigkeiten oder Beziehungen zu untergeordneten Servicebausteinen zu filtern. Beispiele für solche Abfragen wären Anwendungen in der Zuständigkeit bestimmter Personen oder Organisationseinheiten, Anwendungen, die eine bestimmte Anbindungstechnologie nutzen oder Datenflüsse mit Personenbezug.

Um die an den Dokumentations- und Rechercheprozessen beteiligten Stellen bzw. Nutzerinnen und Nutzer zu unter-

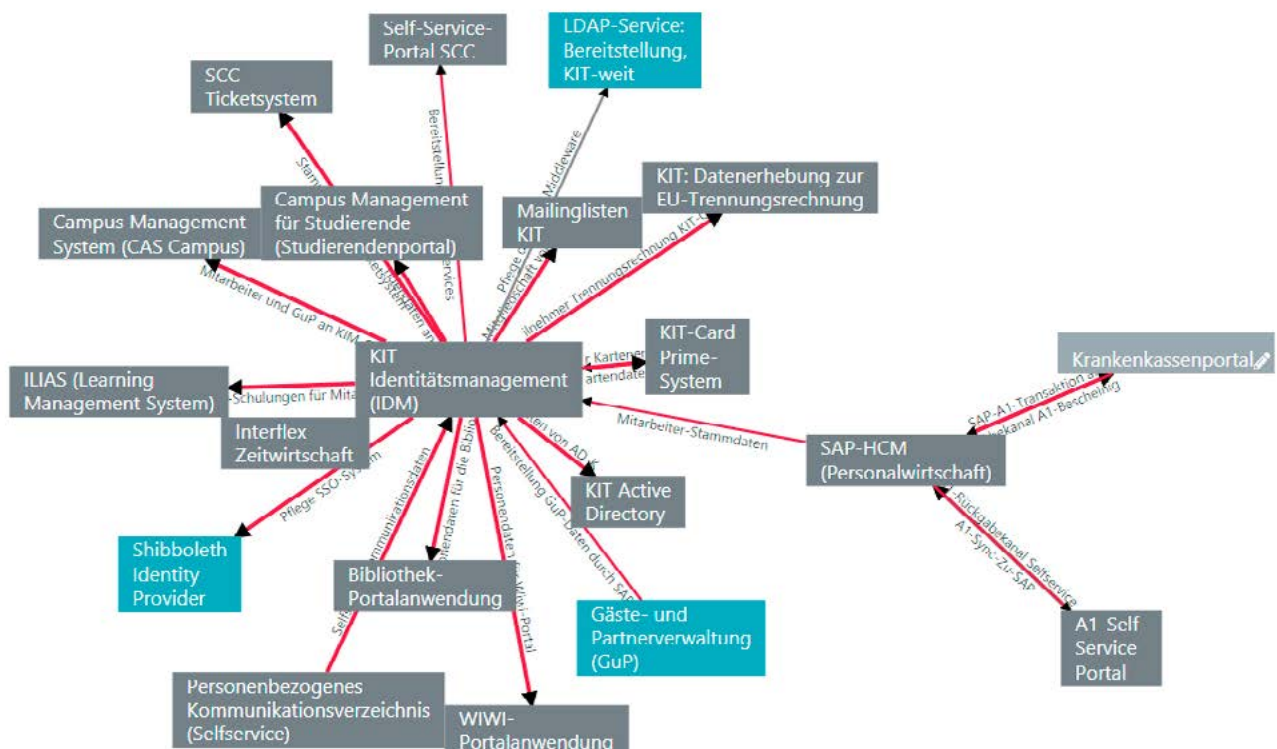


Abbildung 2: Grafischer Auszug aus der IT-Landkarte

stützen – in den Fachbereichen und am SCC durch das jeweilige Betreuungsteam der Anwendungen – existieren Schnittstellen, die eine bequeme, moderne und sichere KIT-weite Nutzung und konsequente Datenbestandspflege ermöglichen. Für den Endnutzer ist das Frontend als Javascript Single-Page-Anwendung realisiert und läuft vollständig im lokalen Browser. Anwender können je nach Nutzungsszenario zwischen verschiedenen Ansichten wählen, in denen der erfasste Datenbestand in Tabellen, Listen oder auch als Graph der Anwendungen und deren Schnittstellen visualisiert dargestellt und editiert werden kann. Abbildung 1 zeigt die Definition eines Datenflusses zwischen zwei Anwendungen, Abbildung 2 zeigt die graphische

Darstellung der Datenflüsse. Von allen Datensätzen, für die ein Pendant in der CMDB existiert, können Nutzerinnen und Nutzer – entsprechende Berechtigungen vorausgesetzt – per Klick direkt in die CMDB navigieren und die dort zusätzlich erfassten Daten sichten.

Technisch wird die Kommunikation mit der CMDB über eine SOAP-Schnittstelle abgewickelt, das IT-Landkarten-Backend lässt sich über eine REST-API ansprechen, welches das Anlegen und Editieren der Daten sowie deren Abfrage über Filterfunktionen erlaubt. Zudem übernimmt sie das Umsetzen der Benutzerberechtigungen und ermöglicht darüber hinaus eine Integration der Landkartendaten in weitere Dienste. Die besagte REST-

Schnittstelle ist hierzu über signierte JSON-Webtokens (JWTs) abgesichert, das Frontend authentifiziert Anwenderinnen und Anwender mit OpenID Connect (OIDC) über Shibboleth. Die Zugriffssteuerung geschieht ausschließlich über OAuth-Rollen im JWT, wobei die Rollen eines Nutzers beim Shibboleth-Login direkt aus KIT-Gruppenmitgliedschaften abgeleitet werden. Hiermit wird u. a. festgelegt, welche Daten die jeweiligen Nutzerinnen und Nutzer sehen. Schreib-, Lese- und administrative Berechtigungen werden ebenfalls über diesen Mechanismus durchgesetzt. Abbildung 3 zeigt die technischen Zusammenhänge der beteiligten Komponenten sowie einige Implementierungs- und Betriebsaspekte.

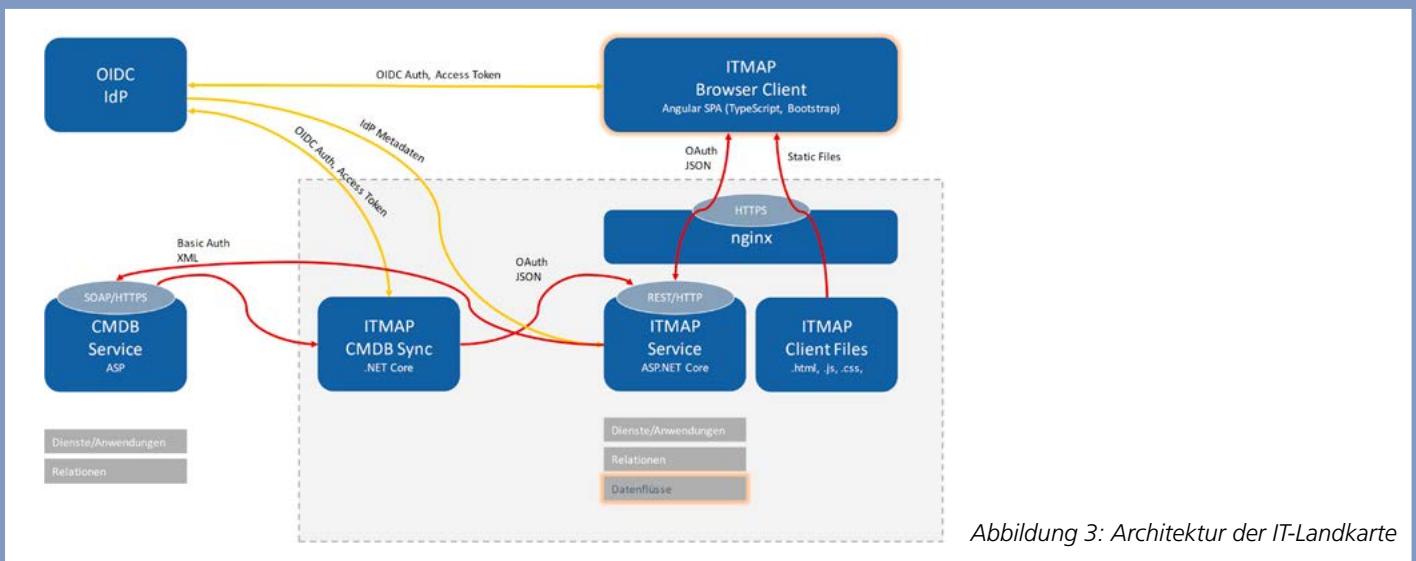


Abbildung 3: Architektur der IT-Landkarte

IT Map of KIT's Software Applications

The SCC operates approximately 170 productive services and around 100 specialized applications. In addition to purely in-house developments, these also include locally or externally maintained standard and special-purpose applications. The interactive IT map developed by the SCC documents this software infrastructure and its interaction, makes the dependencies of KIT's IT services and business processes comprehensible and, if required, serves as a tool for optimization.

bwSync&Share mit neuer Cloud-Software – ein Anwenderbericht

Das SCC betreibt seit sechs Jahren einen föderierten Cloudspeicher am KIT, mit dem die landesweite Zusammenarbeit und das Synchronisieren von Desktopdaten auf unterschiedlichen Endgeräten möglich sind. Dieser Dienst wird im Frühjahr 2020 auf die neue Softwarebasis Nextcloud umgestellt. Anwender konnten die neue Software in der Pilotphase ausgiebig testen.

Achim Grindler

In meiner täglichen Arbeit ist es notwendig, Dokumente und Bilder für andere Institutionen und Personen – intern wie extern – bereitzustellen, um diese bei Bedarf gemeinsam bearbeiten und ergänzen zu können. Hierfür stellt das SCC den landesweiten Cloud-Dienst bwSync&Share sowie – KIT-intern – Microsoft SharePoint Teamseiten zur Verfügung. Das Synchronisieren von Daten auf meinem Endgerät steht nicht im Vordergrund, sondern die plattform- und standortunabhängige Nutzung des Dienstes über einen beliebigen Webbrowser, auch an fremden Geräten.

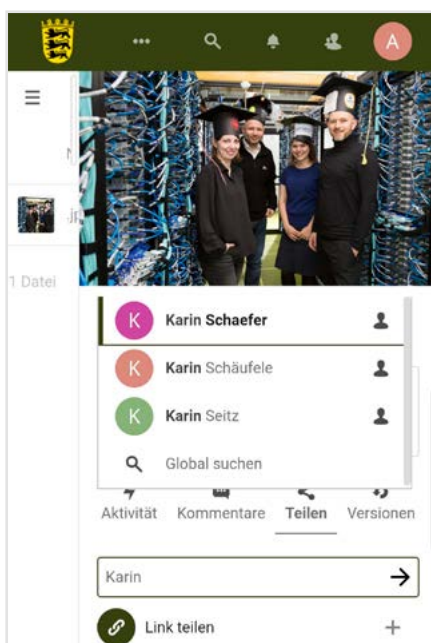


Bild 1

Die bisherige Weblösung von bwSync&Share wies einige Schwächen auf und konnte trotz der Integration der Online-Editing-Software OnlyOffice nicht annähernd an Cloud Services von Google oder Microsoft heranreichen.

Im Frühjahr 2020 wird der Online-Speicherdienst bwSync&Share auf eine neue Softwareplattform umgestellt. Mit ihr kommt eine verbesserte, anwenderfreundlichere Weboberfläche mit modernem, intuitivem Design. Mit dem vorhandenen Funktionsumfang deckt die Lösung meine Anwendungsfälle vollständig ab und ermöglicht ebenfalls das gemeinsame Editieren von Dateien (z.B. Word- oder Excel-Dokumente).

Neben der einfacheren Bedienung hat das neue Webinterface gegenüber dem alten weitere Vorteile: Sehr komfortabel ist das Teilen und Freigeben von Dateiordnern für Personen, die sich schon an der Cloud angemeldet und somit ein Profil erstellt haben (Bild 1). Diese Personen sind in einer durchsuchbaren Liste schnell auswählbar. Alle registrierten Personen können ihr Profil mit weiteren Daten ergänzen, was bei der Zusammenarbeit mit externen Projektpartnern von Vorteil sein kann. Ebenfalls sehr einfach kann ein Freigabelink erstellt und beliebig versendet werden. Durch ein Symbol am Ordner ist sofort erkennbar, ob dessen Inhalt über einen Link freigegeben oder geteilt wurde. Für Freigabe-Links kann zusätzlich ein Ablaufdatum angegeben werden.

Auch die Dateiansichten sind verbessert, so z.B. in der Listenansicht, wo nun bei Bilddateien auch jeweils ein Vorschaubild sowie bei Office- und PDF-Dateien ein Standard-Dateisymbol vorhanden ist. Zudem können nun Dateien mehrfach markiert und Aktionen auf diesen markierten Elementen ausgeführt werden. In der Galerie-Ansicht können Dateien auch

gelöscht werden, was vorher ebenfalls nicht möglich war.

Für Dateien wie beispielsweise Fotos ist es möglich, Schlagworte, sogenannte Tags, zu vergeben. Dies erleichtert so das Auffinden von Bildern (Bild 2).

Alle Tests habe ich mit dem Smartphone (s. Bilder) und dem Desktop-Browser durchgeführt. Hier gibt es keine Unterschiede in der Funktion, Bedienung und Performanz. Bei der bisherigen Cloud-Lösung war das Webinterface nicht brauchbar auf dem Smartphone zu nutzen. Aus meiner Sicht ist die Anwendungsfreundlichkeit des Dienstes bwSync&Share mit dieser neuen Softwareplattform jetzt durchaus mit der der vielgenutzten Online-Speicher-Dienste von Google oder Microsoft vergleichbar.

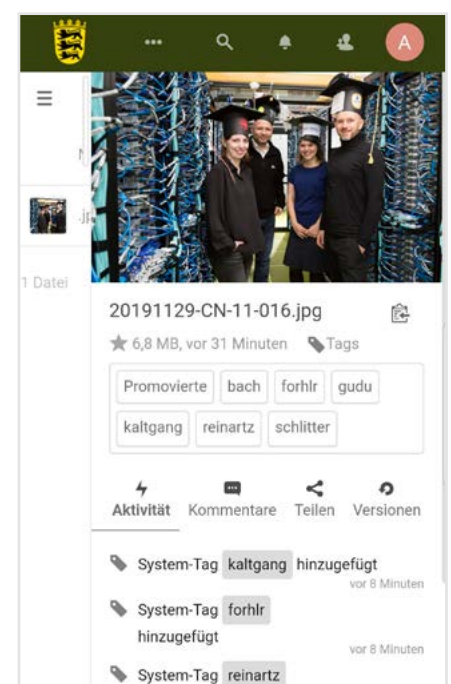


Bild 2



bwCloud-SCOPE – Cloud-Infrastruktur erweitert

Um eigene Projekte oder Arbeitsgruppen individuell mit virtualisierter Infrastruktur erweiterter Leistung und Funktion auszustatten, wurde am KIT die Hardware von bwCloud-SCOPE, der föderierte Cloud-Service für Universitäten und Hochschulen im Land Baden-Württemberg, ausgebaut.

Durch die Erweiterung auf Netzwerkebene können Beschäftigte und Studierende des KIT mit dieser auf Open-Stack basierenden Virtualisierungsumgebung auch Server betreiben, die auf KIT-interne Dienste zugreifen. Den möglichen Nutzungsumfang regeln die zuständigen ITB weiterhin durch zwei verschiedene Berechtigungsstufen.

Matthias Knoll, Klaus Scheibenberger

Ausbau der Cloud-Ressourcen

Über das Betriebsmodell, die Funktionen sowie die Verwaltung der bwCloud-SCOPE berichteten ausführlich die SCC-News 2/2018. Aktuell hat das Betriebsteam des SCC die lokale Installation der bwCloud-SCOPE auf Basis von OpenStack um acht weitere Compute-Server erweitert. So laufen nun am SCC 32 Compute-Hosts für virtuelle Maschinen (VM) mit insgesamt 10 Terabyte Arbeitsspeicher und rund 32 Terabyte NVMe-Datencache. Die Storage-Einheit (JBODs) wurde zudem um 36 weitere Festplatten maximal bestückt, so dass die Lösung insgesamt 936 Terabyte Speicherplatz für virtuelle Server und deren Daten bereitstellt. Damit ist bwCloud-SCOPE für lokale Projekte sowie für den Ausbau zu einer PaaS/

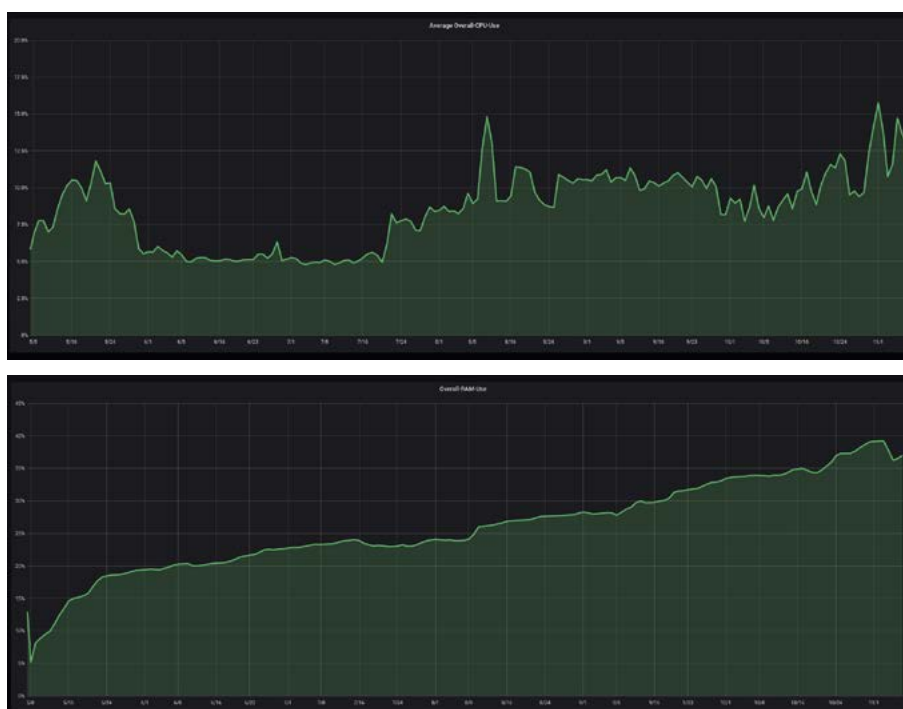


Abbildung 2: CPU- (oben) und RAM-Nutzung (unten) im letzten halben Jahr am Standort Karlsruhe (über alle Compute-Hosts)

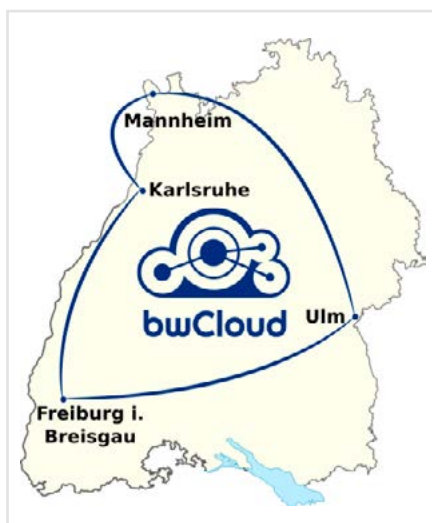


Abbildung 1: Die vier Betriebsstandorte der bwCloud-SCOPE (links)

SaaS¹-Lösung gerüstet. Mit der größeren RAM-Ausstattung lassen sich zudem virtuelle Server mit mehr Arbeitsspeicher und größerer Systemplatte starten. Das ermöglicht Nutzern nun das Starten und Betreiben von VMs mit Windows-Betriebssystemen.

bwCloud-Erweiterungen für KIT-Nutzer

Ab Mitte 2019 äußerten immer mehr Nutzerinnen und Nutzer den Bedarf die Ressourcen der Cloud zusammen mit internen Diensten zu nutzen z.B. den Zu-

griff auf den KIT-Lizenzserver oder interne Backup-Dienste. Dies ist jedoch mit dem bisherigen Betrieb der VMs im externen BelWü-Netz mit öffentlichen IP-Adressen im Internet nicht möglich, da Firewalls den Zugriff auf interne Ressourcen verhindern. Daher wird das bisherige Netzwerkschema der bwCloud in Karlsruhe um interne Netzwerke erweitert.

¹ PaaS = Platform as a Service
SaaS = Software as a Service

| bwCloud-Standort | Anzahl Rechnerknoten | Speicherplatz | CPUs / Kerne |
|-----------------------------|----------------------|---------------|--------------|
| Karlsruhe (KIT) | 32 | 936 TByte* | 32 / 512 |
| Ulm (kiz) | 36 | 1.200 TByte* | 36 / 576 |
| Freiburg (Rechenzentrum) | 27 | 504 TByte* | 27 / 432 |
| Mannheim (Rechenzentrum) | 27 | 504 TByte* | 27 / 432 |
| Gesamt für Produktivbetrieb | 122 | 3.144 TByte* | 122 / 1952 |

Tabelle 1: Die aktuelle Ausstattung an den vier Betriebsstandorten

* = Brutto

Interne Netze

Interne Netze erlauben Nutzerinnen und Nutzern von KIT-Instituten virtuelle Server der bwCloud auch in einem internen VLAN zu starten, das dem jeweiligen Institut zugeordnet ist. Der Zugriff auf das interne VM-Netz muss per Ticket bei der lokalen Betriebsgruppe angefragt werden.

Interne Dienste

Mit solchen intern betriebenen virtuellen Servern können die Anwender oder Anwendergruppen² beispielsweise den Backup³- oder Archiv⁴- Server über das interne Netz nutzen. Ebenso sind weitere Self-Services wie MySQL-Datenbanken⁵ oder die Freischaltung von Software über interne Lizenzserver im KIT-Netz erreichbar.

Windows-Betrieb/-Lizenzen

Beschäftigte am KIT können für die internen VMs über den KMS-Lizenzserver auch Windows-Server-Lizenzen erhalten und somit Windows-Server betreiben. Diese sind damit KIT-Systeme hinter der KIT-Firewall und können bei Bedarf von der Abteilung NET für den Zugriff von außen freigeschaltet werden. Für diese gelten die gleichen IT-Sicherheitsbedingungen wie für den Betrieb von institutseigenen Servern mit Zugriff von außen. Deshalb ist es wichtig, sich je nach Anwendungsfall vor der Inbetriebnahme eigener VMs für den internen oder externen Betrieb im BelWü-Netz zu entscheiden. Technisch

lassen sich so alle aktuellen Windows-Server-Varianten in der bwCloud betreiben. Derzeit stellt das Betriebsteam ein eigenes, manuell vorbereitetes Standard-Windows-Image bereit. Die bwCloud-Betriebsgruppe unterstützt Interessierte bei der Erstellung und dem Betrieb von eigenen Windows-VMs. Auch VMs mit anderen Betriebssystemen mit der Architektur x86-64 sind so lauffähig. Dies

erfordert, dass der Nutzer ein passendes Image manuell vorbereitet⁶ und dann in die Cloud lädt und startet.

Weitere technische Details:

www.bw-cloud.org

Informationen zu den aktuellen

Entwicklungen der bwCloud:

www.scc.kit.edu/dienste/bwcloud.php.

| # | Flavour | Freiburg | Mannheim | Ulm | Karlsruhe |
|--------------|------------|------------|------------|------------|------------|
| 1 | m1.nano | 36 | 33 | 12 | 23 |
| 2 | m1.tiny | 204 | 38 | 36 | 95 |
| 3 | m1.small | 212 | 55 | 86 | 113 |
| 4 | m1.medium | 174 | 121 | 126 | 152 |
| 5 | m1.large | 140 | 160 | 162 | 179 |
| 6 | m1.xlarge | 26 | 17 | 42 | 33 |
| 7 | m1.xxlarge | 13 | 41 | 4 | 41 |
| SUMME | | 805 | 465 | 468 | 636 |

Tabelle 2: Anzahl der aktiven virtuellen Server je Betriebsstandort

| bwCloud Netze | vom Internet aus erreichbar |
|---|-------------------------------|
| BelWü (öffentliche IP-Adressen für VMs) | Ja (extern) |
| SCC-bwcloud-1 (interne IP-Adressen) | Nein (intern, nur im KIT-LAN) |
| SCC-bwcloud-2 (interne IP-Adressen) | Nein (intern, nur im KIT-LAN) |
| <OE-Kürzel>-bwcloud-x (interne IP-Adressen) | Nein (intern, nur im KIT-LAN) |

Tabelle 3: Das externe und die neuen, intern verfügbaren Netze der bwCloud

bwCloud-SCOPE – Cloud infrastructure expanded

In order to individually equip own projects or working groups with virtualized infrastructure of extended performance and function, the hardware basis of bw-Cloud-SCOPE, the federated cloud service for universities and colleges in the state of Baden-Württemberg, has been expanded at the KIT.

The extension at network level enables KIT employees and students to use this OpenStack-based virtualization environment to run servers that access KIT-internal services. The responsible IT officers in the institutes continue to regulate the possible scope of use through two different authorization levels.

² www.bw-cloud.org/de/faq/gruppen

³ www.scc.kit.edu/scc/sw/backup/tsm/index.php

⁴ www.rda.kit.edu/

⁵ www.scc.kit.edu/dienste/5926.php

⁶ docs.openstack.org/image-guide/create-images-manually.html

VPN2VLAN mit IPv6 im Tunnel

Im allgemeinen VPN-Zugang schon länger vorhanden, fehlte die IPv6-Unterstützung im Tunnel bei VPN2VLAN jedoch immer noch. Da es hierfür keine Standardlösung gab, musste eine eigene Lösung gefunden werden. Nach langer Zeit der Überlegungen und Tests wurde im Oktober 2019 die IPv6-Konnektivität im Tunnel bei VPN2VLAN realisiert.

Klara Mall

Für den Remote-Zugriff auf nur aus dem KIT-Netz erreichbare Ressourcen steht der VPN-Dienst des SCC zur Verfügung. Der allgemeine VPN-Zugang des KIT ist mit OpenVPN realisiert und stellt ein geroutetes VPN dar. Die VPN-Clients bekommen zwar alle IP-Adressen aus demselben Subnetz, können sich aber untereinander nicht „sehen“. Bei diesem Konzept lässt sich OpenVPN leicht so konfigurieren, dass zusätzlich zu dem IPv4-Subnetz ein IPv6-Subnetz zur Verfügung steht, aus dem jeder Client eine IPv6-Adresse erhält. Somit hat der Client sowohl eine IPv6- als auch eine IPv4-Adresse, und die Konnektivität ist im Tunnel sowohl mit IPv6 als auch mit IPv4 gegeben. Dies ist bei dem allgemeinen VPN-Zugang des KIT seit Mitte 2017 der Fall.

Bei den am KIT darüber hinaus zur Verfügung gestellten VPN-Spezialzugängen – VPN2VLAN genannt – handelt es sich um ein gebridgtes OpenVPN, bei dem jeder Client abhängig von seinem Realm – das ist der mit @ an den Benutzernamen

angehängte Bereich – eine direkte Verbindung in das gewünschte Ziel-VLAN erhält. Die Verbindung verhält sich dabei genauso wie wenn man sein Endgerät im Büro an die LAN-Dose anschließen würde, auf der dieses VLAN konfiguriert ist. Der VPN-Nutzer erhält mit VPN2VLAN dieselbe Netzwerkumgebung wie vor Ort am KIT. Ebenso können auf diese Art und Weise Freischaltungen in bestimmte Netze für einen ausgewählten Personenkreis ermöglicht werden.

Wie bekommt der Client eine IPv6-Adress-Konfiguration im VPN2VLAN?

Bei VPN2VLAN gibt es keine Standardlösung für die IP-Adressvergabe mit IPv6. Je Bereich muss die IP-Adresse für den Client aus dem passenden Subnetz vergeben werden. Zudem ist der Client mit dem VLAN auf Layer 2 verbunden und erhält so Router Advertisements aus dem VLAN. Zunächst war die Überlegung, ob die Router Advertisements ausreichen. Anhand dieser konfiguriert sich der Client wie im LAN eine IP-Adresse. Dies würde

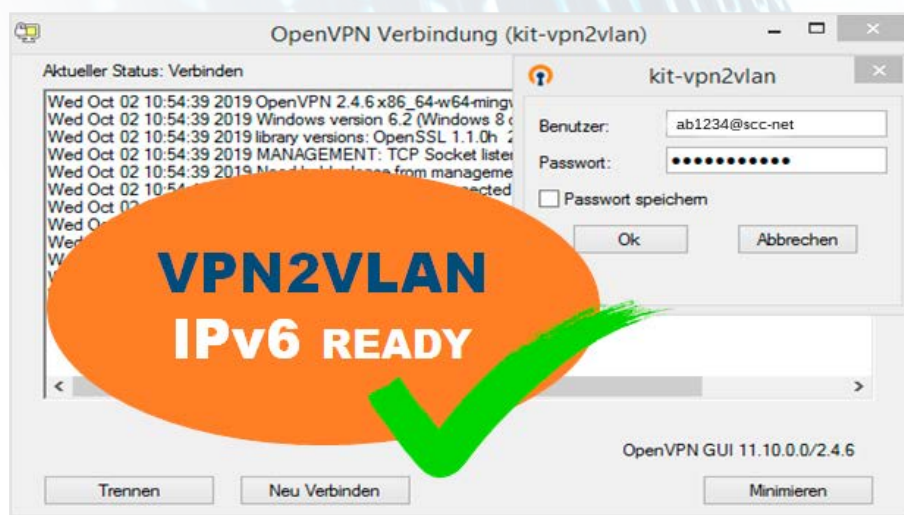
prinzipiell funktionieren, das Problem ist aber die Routing-Tabelle. Für die darunterliegende Verbindung hat der Client bereits eine Default Route. Nun erhält er durch das Router Advertisement eine weitere, die allerdings wirkungslos bleibt.

Die nächste Überlegung war, die Router Advertisements aus dem VLAN zu blockieren und mit der Open Source Software radvd (router advertisement daemon) selbst Router Advertisements zu schicken, so dass der Client sich mit Hilfe dieser seine IP-Adresse und Routen konfiguriert. Hier war die Idee, die Route Information Option im Router Advertisement (RFC 4191) zu nutzen. Nach einigen Tests schief dies aber wieder ein, da dies auch der Client unterstützen muss, was nicht bei allen Clients der Fall war.

Wie im Artikel "IPv6 in der Fläche" (SCC-News 1/2019) geschildert, treibt das SCC inzwischen die flächendeckende Einführung von IPv6 voran. Sobald Institute anfangen, ihre eigenen Dienste mit IPv6 zu betreiben und VPN2VLAN nutzen, wird das Manko offensichtlich, dass auch 2019 die IPv6-Unterstützung im Tunnel bei VPN2VLAN noch immer nicht gegeben war.

Lösung: Vergabe zufälliger IP-Adressen durch den OpenVPN-Server

Mit dem Projekt bwIPv6, an dem alle Universitäten in Baden-Württemberg teilnehmen, rückte das Thema erneut in den Fokus und folgende Idee wurde geboren: warum machen wir es bei IPv6 nicht analog zu IPv4? Bei IPv4 wird die IP-Adresse für den Client durch den OpenVPN-Server per DHCP bezogen. Diese IP-Adresse wird dann vom OpenVPN-Server an den Client vergeben. Genauso könnte man bei IPv6



eine IP-Adresse beziehen oder generieren und diese an den Client vergeben. Anstatt dass nun der OpenVPN-Server eine IP-Adresse per DHCPv6 bezieht, kann er selbst eine IP-Adresse für den Client generieren. Diese IP-Adresse bleibt über einen Tag hinweg stabil. Dahinter steckt derselbe Gedanke wie bei der IPv6 Stateless Address Autoconfiguration. Der Adressraum ist hier so groß, dass eine Kollision extrem unwahrscheinlich ist. Allerdings – käme es zum Fehlschlagen der Duplicate Address Detection durch den Client nach der Zuweisung der IPv6-Adresse – gäbe es keinen Mechanismus in OpenVPN, um eine neue Adresse zu beziehen. Sollte dieser unwahrscheinliche Fall eintreten, hätte der Client in der Folge im Tunnel keine IPv6-Adresse. Dies

ist aber aufgrund der geringen Wahrscheinlichkeit und der Tolerierbarkeit des Fehlers zu verkraften. Auch bei dieser Lösung werden die Router Advertisements auf dem Weg zum VPN-Client blockiert. Der Client erhält also analog zu IPv4 die Informationen für seine IP-Konfiguration ausschließlich vom VPN-Server.

Zur Umsetzung dieser Idee wurde in den Programmcode, der auf dem OpenVPN-Server läuft und die IPv4-Adresse für den Client per DHCP bezieht, eine Erweiterung eingebaut. Diese generiert für den Client nun zusätzlich eine IPv6-Adresse, sofern ein IPv6-Subnetz auf diesem VLAN vorhanden ist. Nach einem kurzen Testbetrieb mit ausgewählten Nutzern ging die Implementierung im Oktober 2019 in Produktion. Mit der vollständigen IPv6-

Fähigkeit des VPN-Dienstes am KIT haben wir einen weiteren Meilenstein bei der Einführung von IPv6 erreicht.

Ausblick

Für den OpenVPN-Dienst am KIT allgemein sei hier noch die Information gegeben, dass dieser in den nächsten Monaten mit der kommenden OpenVPN-Version 2.5 komplett neu aufgesetzt werden wird. Aufgrund geänderter Optionen wird ein Austausch der Konfigurationsdatei notwendig werden, auf den aber während einiger Monate Parallelbetrieb der alten und neuen VPN-Server ausgiebig hingewiesen werden wird.

VPN2VLAN with IPv6 inside the tunnel

The general VPN service at KIT supported IPv6 in the tunnel since 2017. However, with VPN2VLAN you received no IPv6 address from the VPN server. Nevertheless, you received router advertisements, which could lead to either a functioning IPv6 connection or unclear routing tables with a malfunctioning IPv6 connection. We implemented a solution which blocks all router advertisements and enables the VPN server to generate an IPv6 address and to push the address and the routes to the VPN client.



Microsoft
Windows 7

Achtung Windows 7 - Notfälle in die „Quarantäne-Station“

„Windows Update Service – Es sind wichtige Updates verfügbar. Rufen Sie die PC-Einstellungen auf, um diese zu installieren.“ Diese Ankündigung, die Windows-Anwenderinnen und -Anwender beim Hochfahren des Systems regelmäßig zu sehen bekommen, verstummt **ab dem 14. Januar 2020** für

PCs mit Windows 7 endgültig. Dann stellt Microsoft jede technische Unterstützung sowie die Software-Updates ein. Das SCC hat hierzu umfangreiche Vorbereitungen zur Unterstützung getroffen und die IT-Beauftragten in den Organisationseinheiten (OE) über die notwendigen Maßnahmen informiert, um die IT-Infrastruktur des KIT zu schützen; denn es besteht ein hohes Risiko für nicht aktuelle Systeme, kompromittiert zu werden.

Der IT-Sicherheitsrichtlinie des KIT folgend, müssen „sowohl Client- als auch Serversysteme regelmäßig mit Updates der Betriebssystem- und Anwendungssoftware versorgt werden. Die Verwaltung der Sicherheitsupdates ist [...] Bestandteil der IT-Administration, um Sicherheitsrisiken zu minimieren. [...] Für den sicheren Betrieb ist es unerlässlich, diese Updates zeitnah einzuspielen.“

Somit müssen alle PCs mit Windows 7, die im KIT-Netz betrieben werden, entweder abgeschaltet oder auf eine vom Hersteller gewartete Version aktualisiert werden – und das bis Mitte Januar 2020! Die vom SCC zentral verwalteten PCs sind bereits auf einen aktuellen Stand migriert bzw. durch aktuelle Systeme ersetzt worden. Für Systeme, die unbedingt Windows 7 voraussetzen, bspw. in Mess-, Steuer- oder Laboranlagen, hilft das SCC entsprechende Quarantänenetze einzurichten, so dass die Rechner dorthin umgezogen werden können. Hier sind sie isoliert vom Internet und dem Rest des Datennetzes bzw. den SCC-Diensten und damit auch von den übrigen, geschützten PCs im KIT-Netz.

Für Mess-, Steuer- und Laborsysteme, die nicht in Quarantänenetzen betrieben werden können, besteht ein letzter Notnagel: Der Erwerb der kostenpflichtigen "Extended Security Updates" (ESU).

Ansprechpartner: Andreas Lorenz

Erfolgreicher Abschluss der ersten Projektphase von V-FOR-WaTer

V-FOR-WaTer ist eine Virtuelle Forschungsumgebung zur Prozessierung von Wasser- und terrestrischen Umweltdaten. Sie ermöglicht Wissenschaftlern den einfachen Zugriff auf Daten, schnellere Analysen durch automatische Präprozessierung und die Reproduzierbarkeit von Analysen.

Marcus Strobl, Sibylle Haßler¹



Ziel von V-FOR-WaTer ist es, über ein Webportal den Zugang zu heterogenen hydrologischen Daten und den entsprechenden Analysewerkzeugen zu ermöglichen und somit deren Präprozessierung,

Skalierung und wissenschaftliche Analyse zu beschleunigen. Nicht nur Daten, sondern auch die zugehörigen Werkzeuge und Arbeitsabläufe sollen öffentlich zugänglich sein, sodass Datenanalysen reproduzierbar werden. Die Daten im Webportal stammen sowohl aus Universitätsprojekten als auch aus dem Basismonitoring von Landesämtern wie der LUBW (Landesamt für Umwelt Baden-Württemberg). Als Nutzer des Portals stehen Wissenschaftlerinnen und Wissenschaftler von Universitäten und Landesämtern im Fokus. Im Portal sind umfangreiche Anforderungen bezüglich Datenzugriff und -publikation technisch

umgesetzt: Durch eine feingranulare Zugriffsverwaltung wird der Schutz von sensiblen Daten gewährleistet. Eine Embargoperiode ermöglicht das Abschließen von Arbeiten und Publikationen, bevor hochgeladene Daten für alle Nutzerinnen und Nutzer öffentlich zugänglich werden. Durch den Anschluss des Portals an bestehende Datenrepositorien wird zudem die Möglichkeit gegeben, hochgeladene Daten über die bestehenden Repositorien zu publizieren. Somit können die Wissenschaftlerinnen und Wissenschaftler auf einfache Weise die Anforderungen von Zeitschriften und Geldgebern erfüllen, wissenschaftliche Daten öffentlich zugänglich zu publizieren.

In den verstrichenen drei Jahren wurde das Projekt V-FOR-WaTer vom Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg (MWK) finanziert und ein Prototyp der virtuellen

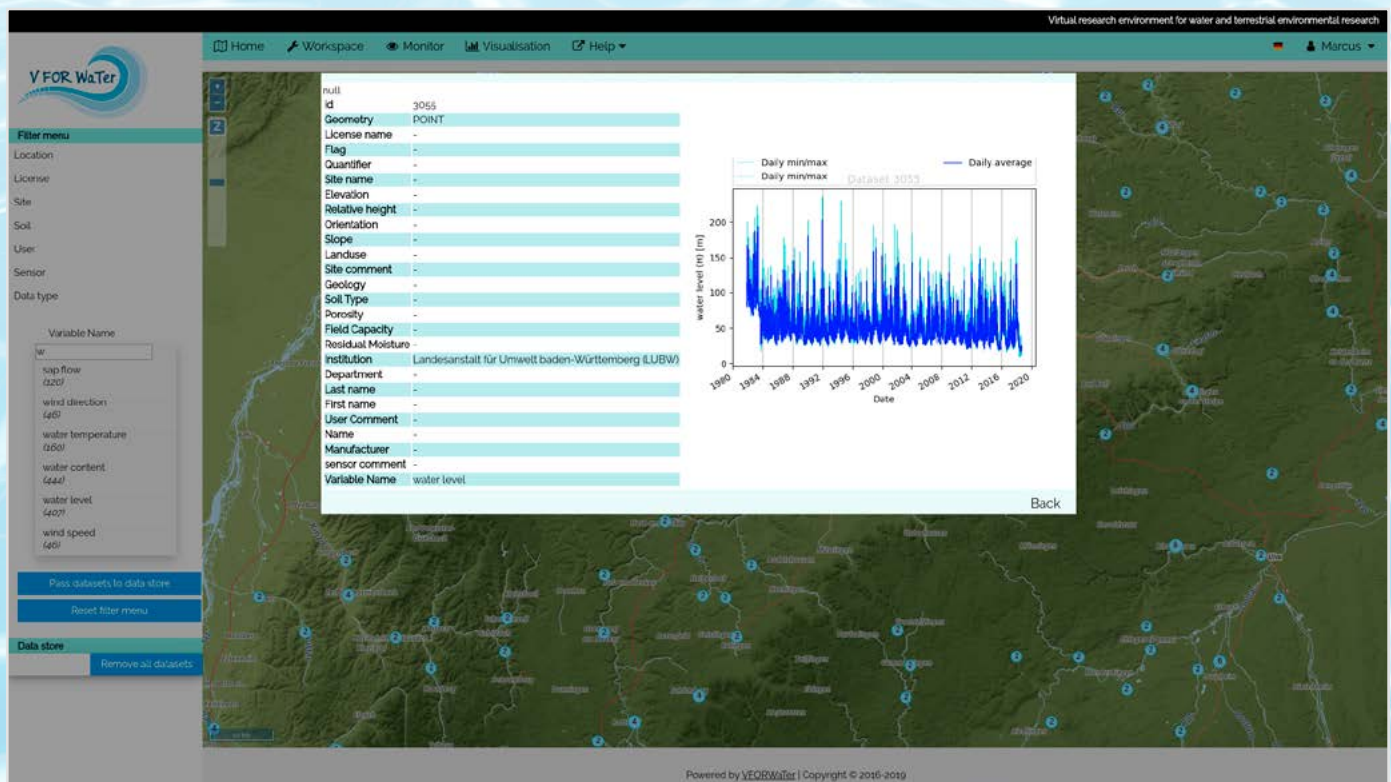


Abb. 1: Screenshot der Startseite des V-FOR-WaTer Webportal mit Vorschau auf den Datensatz einer Zeitreihe.

¹ Institut für Wasser und Gewässerentwicklung (IWG)

Forschungsumgebung entwickelt. Aktuell umfasst das System eine Datenbank (PostGIS) mit Punkt- und Zeitreihendaten mit einem ausführlichen Metadatenmodell, das auf Wasser- und Umweltdaten angepasst und mit internationalen Standards (INSPIRE, ISO19115) konform ist. Die für den Prototypaufbau verwendeten Datensätze im Portal stammen aus Forschungsprojekten sowie aus dem Monitoringnetzwerk der LUBW. Um Zugriff auf eine noch breitere Auswahl an Daten zu ermöglichen und den Weg für Datenpublikationen aus dem Portal zu ebnen, ist geplant, das Forschungsdatenrepositorium für geowissenschaftliche Daten, GFZ Data Service, an das Portal anzubinden. Zu diesem Zweck wird derzeit das Metadatenmodell überarbeitet. Der Zugang zu den Daten im Portal erfolgt über ein umfangreiches Filtermenü, in dem die Datensätze über vorgegebene Metadaten sowie graphisch über die Auswahl einer Lokation auf der Karte ausgewählt werden können (Abb. 1). Im Arbeitsbereich kann auf Werkzeuge zur Präprozessierung für typische hydrologische Anwendungen und zur Geostatistik zugegriffen werden. Da die Werkzeuge als Web Processing Services (WPS des OGC-Standards) implementiert sind, kann das System sehr einfach erweitert werden. Die Auswahl der Werkzeuge orientiert sich stark an den enthaltenen Daten und den daraus abgeleiteten Fragestellungen und wird entsprechend des Bedarfes der Wissenschaftlerinnen und Wissenschaftler im Fachgebiet Hydrologie

erweitert.

Um eine möglichst große Akzeptanz zu erfahren, ist neben der einfachen Bedienbarkeit des Portals eine gute Auffindbarkeit und eine passende Datengrundlage nötig. Dies wird durch eine enge Zusammenarbeit zwischen dem SCC und dem Institut für Wasser und Gewässerentwicklung (IWG) gewährleistet. Zur Evaluation der unterschiedlichsten Anforderungen aus den Wasserwissenschaften und der Erhöhung der Sichtbarkeit präsentiert sich V-FOR-WaTer regelmäßig auf Konferenzen und fragt Bedarfe und Wünsche über Workshops in der lokalen Fachcommunity ab. Aktuell wird das Portal am SCC im Data Life Cycle Lab „Climatology“ des Helmholtz-Programms „Supercomputing and Big Data“ weiterentwickelt, und am IWG werden zusätzliche Werkzeuge für das Portal erarbeitet. Zudem ist V-FOR-WaTer über das Postdoc-Projekt „BRIDGET“, das vom IWG eingeworben wurde, in der Helmholtz-Initiative „Digital Earth“ involviert. In BRIDGET werden Daten aus Verdunstungsmessungen (Evapotranspiration) und daran angepasste Skalierungswerkzeuge in einem Paket exemplarisch über V-FOR-WaTer zur Verfügung gestellt.

Weitere Informationen sind unter www.vforwater.de zu finden.

Der Programmcode des Portals ist unter der MIT-Lizenz auf GitHub publiziert:

github.com/VForWaTer/vforwater-portal

Successful completion of the first project phase of V-FOR-WaTer

V-FOR-WaTer is a virtual research environment for accessing and processing of water and terrestrial environmental data. By giving scientists from universities and state offices open access to data, appropriate pre-processing, analysis tools and workflows, we accelerate scientific work and facilitate the reproducibility of analyses. Data in the web portal originates from university projects and state offices (e.g. LUBW). V-FOR-WaTer will be connected to established repositories and thus simplifies the process both of accessing publicly available datasets and publishing the portal user's data, which is increasingly demanded by journals and funding organisations.

Within the project V-FOR-WaTer, funded by the Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg (MWK) for a duration of three years, a prototype of the virtual research environment has been developed. Today it consists of a database with a detailed metadata scheme that is adapted to water and terrestrial environmental data and compliant with international standards (INSPIRE, ISO19115). The time series data we used for development originates from research projects and the LUBW. A filter menu that is based on the metadata and a graphical selection on the map gives access to the data (Fig. 1). A work area so far provides tools for data pre-processing, geostatistics and common hydrological applications. The tools are implemented as Web Processing Services (WPS); hence, the system can easily be extended according to the needs of the scientists using the portal.

DEEP Hybrid DataCloud

Das Projekt DEEP Hybrid DataCloud untersucht, wie rechenintensive Anwendungen, die Anforderungen an Hochleistungsrechner (HPC) und Grafikprozessoren (GPUs) haben, diese Ressourcen als Cloud Dienst nutzen können. Ein hybrider Cloud-Ansatz soll Wissenschaftlerinnen und Wissenschaftlern einen einfachen Zugang zu den Cloud-Ressourcen ermöglichen.

Marcus Hardt

Was ist DEEP?

Das Konzept des europäischen Forschungsvorhabens DEEP Hybrid DataCloud (DEEP) ist es, Computertechniken zu unterstützen, die spezielle Hardware wie Grafikkarten (Graphic Processing Units, GPUs) oder Low-Latency-Interconnects (wie z.B. Infiniband) erfordern, um sehr große Datensätze zu untersuchen. Ein hybrider Cloud-Ansatz ermöglicht dabei den Zugang zu solchen Ressourcen, die von wissenschaftlichen Rechenzentren als Projektpartner bereitgestellt werden. Ganz im Sinne der European Open Science Cloud (EOSC) wird es dabei auch Forschenden einfacher gemacht, Zugang zu diesen Ressourcen zu bekommen.

Unter dem gemeinsamen Label "DEEP as a Service" entwickelt das Projekt eine Reihe von Bausteinen, mit deren Hilfe einfache Anwendungen zusammengebaut werden können. Dabei fokussiert sich das Projekt auf Anwendungen, die folgende Techniken erfordern: Machine Learning / Deep Learning mit Deep Learning Neural Networks, parallele Nachbearbeitung sehr großer Datenmengen und die Analyse von gewaltigen Online-Datenströmen. Diese Infrastrukturen und Dienste werden in einem Projekt-Testbett bereitgestellt und in die europäische Forschungs-Cloud EOSC integriert, wo sie in Zukunft weiter ausgebaut werden können. Die am Projekt mitwirkenden Wissenschaftlerinnen und Wissenschaftler haben hier die Möglichkeit, die neuen Methoden anhand von Pilotanwendungen auszuprobieren.

Die Aufgaben von DEEP

Das Projekt gliedert sich in verschiedene Aufgabenbereiche: (1) Eine Vernetzungstätigkeit, die sich auf drei Pilotanwendungen mit rechenintensiven Techniken konzentriert. Dazu kommt (2) eine Serviceaktivität, die ein Testbett mit einigen Cloud- und einzelnen HPC-Ressourcen, einschließlich GPUs der neuesten Generation und Infiniband, zur Verfügung stellt, um die Leistung und Skalierbarkeit der Lösungen zu bewerten. Eine weitere (3) Serviceaktivität implementiert einen DevOps-Ansatz, um dieses moderne Softwareentwicklungskonzept auch in die



Wissenschaft einzubringen¹. Dies stellt einerseits sicher, dass die Qualitätsmerkmale der Software im Projekt eingehalten werden, und dient andererseits dazu, möglichst schnell und agil neue Dienste bereitzustellen, die während der gesamten Projektlaufzeit erneuert und auch von Forschungsanwendungen verwendet werden.

Am SCC werden die Aufgaben von Dr. Valentin Kozlov organisiert und bearbeitet. Hier geht es zum einen um Infrastruktur, wo wir gezeigt haben, dass auch existierende Infrastrukturen, wie z.B. das HPC-System ForHLR II, prototypisch als Projektressourcen integriert werden können. Zum anderen geht es um

Anwendungen, bei denen Herr Kozlov die teilnehmenden Anwendungsbetreuer koordiniert und selbst Integrationslösungen entwickelt. Hierbei wird er tatkräftig von den studentischen Hilfskräften Adrian Grupp und Silke Donayre unterstützt.

Ergebnisse und Erfolge

Die Software von DEEP wird zusätzlich zum DevOps-Ansatz auch versioniert freigegeben. Im Januar wurde Version 1 "Genesis" veröffentlicht. Nach dem im September abgehaltenen Vorbereitungsmeeting im polnischen Poznan (Bild 1) arbeitet das Team daran, die Version 2 noch vor Jahresende fertig zu stellen.

Das Projekt DEEP läuft seit November 2017 und endet im April 2020. Die Begutachtung zur Hälfte der Laufzeit fand bereits Ende März 2019 in Luxemburg statt. Hier konnte Valentin Kozlov bereits einen großen Teil der Pilotanwendungen vorführen und mit den Gutachtern diskutieren. Dementsprechend positiv verlief die Begutachtung. Im Juni 2019 organisierte das SCC zwei Seminarveranstaltungen zu DEEP. Valentin Kozlov stellte das Projekt vor, und Mario David vom LIP (Lissabon) erläuterte das Werkzeug uDocker.

Lösungen auf Basis von DEEP werden auf dem Marketplace² veröffentlicht. Nutzer können dort fertige Module für Maschinelles Lernen und Deep Learning finden, teilen, speichern und herunterladen.

¹ siehe z.B. GridKa School 2019
indico.scc.kit.edu/event/460/contributions/5776/

² marketplace.deep-hybrid-datacloud.eu



Bild 1: Teilnehmerinnen und Teilnehmer des DEEP All Hands Meeting am Poznan Supercomputing and Networking Center – PSNC

Dies beinhaltet funktionierende und einsatzbereite Anwendungen, wie z.B. Bildklassifikationstools, Erkennungsmaschinen für Netzwerkanomalien etc. sowie allgemeinere und generische

Modelle, z.B. für die Bildsegmentierung, oder Auflösungsverbesserungswerkzeuge. Außerdem umfasst der Marktplatz zusätzliche Komponenten wie z.B. Anwendungstopologien für die Daten-

aufnahme. Mit einer speziellen "DEEP as a Service Training Facility"³ – bereits über den EOSC-Marktplatz verfügbar – können schnell Erfahrungen mit den DEEP Lösungen gesammelt werden.

Deep Hybrid DataCloud

The key concept proposed in the DEEP Hybrid DataCloud project is to support intensive computing techniques that require specialized hardware, like GPUs or Infiniband, to explore very large datasets. A Hybrid Cloud approach enables the access to such resources that are not easily reachable by the researchers at the scale needed in the current EU such as EGI, EUDAT or WLCG. We also propose to deploy under the common label „DEEP as a Service“ a set of building blocks that enables the easy development of applications: e.g. DEEPaaS API to access machine learning models, Open Catalog for sharing developed applications, Orchestration of resources or DevOps approach. The user applications feature deep learning techniques for e.g. image classification, satellite imagery or analysis of massive online data streams. The developed services are offered to research communities of the project and will be integrated under the EOSC framework, where they can be further scaled up in the future. Solutions based on DEEP are published in the marketplace² where users can browse, share, store and download ready to use machine learning and deep learning modules. This includes working and ready to use applications (such as image classification tools, network anomaly detection engines etc.) as well as more general purpose and generic models (such as image segmenters or image super resolution tools). Moreover, the marketplace also comprises additional components, like complex application topologies (for example for data ingestion), and model and application associated metadata. The EOSC Marketplace already offers a service "DEEPaaS training facility" provided by DEEP³.

³ marketplace.eosc-portal.eu/services/deepaas-training-facility

Die EOSC jenseits der Technologie

Dieser Artikel ist Teil einer Reihe, die sich mit Arbeiten des SCC im Zusammenhang mit der Entwicklung der europäischen Dateninfrastruktur, der European Open Science Cloud (EOSC), befasst. Diesmal steht das Projekt EOSCsecretariat.eu im Mittelpunkt. Im EOSCsecretariat.eu¹ ist das SCC verantwortlich für die Kontaktpflege und Vernetzung in der EOSC und weiteren europäischen Forschungsinfrastrukturen wie z. B. den ESFRI² Projekten.

Jos van Wezel

Die Vision einer Dateninfrastruktur, die einen europaweiten Datenzugriff für die Forschung ermöglicht, beinhaltet nicht nur technische Herausforderungen, sondern erfordert die Zusammenarbeit bei der Überwindung von Unterschiedlichkeiten im Management und der Förderung von Forschung. Europa ist ein Kontinent mit einem breiten sozialen, politischen und rechtlichen Spektrum, was sich besonders beim Überschreiten von Ländergrenzen bei dem Aufbau einer länderübergreifenden Dateninfrastruktur bemerkbar macht. Denken Sie daran, dass der europäische Binnenmarkt auch nicht über Nacht etabliert wurde. EOSCsecretariat.eu hilft dabei, Grenzen der europäischen Forschungslandschaft zu überschreiten, europäische Wissenschaftsgemeinschaften zusammen zu bringen und damit eine nachhaltige und weltweit führende Dateninfrastruktur für alle Forschungsgemeinschaften in Europa aufzubauen.

Aufbau der technischen Infrastruktur der EOSC

In der jüngeren Vergangenheit haben viele EU-Projekte in großem Maßstab Werkzeuge und Dienste zur Unterstützung des Datenmanagements bereitgestellt. Eine Leistung, die gleichermaßen von Experten aus Rechenzentren, wie dem SCC, und von Wissenschaftsgemeinschaften erbracht wurde. In diesem Zusammenhang wäre die vom CERN geleitete internationale Zusammenarbeit in der Teilchenphysik zu nennen, die zu einer weltweiten Rechner- und Dateninfrastruktur für das LHC³-Computing führte. Forschungseinrichtungen aus den



Bereichen Biowissenschaften, Klimatologie, Astronomie und andere bauten ähnliche, thematisch orientierte Strukturen auf, um den ständig wachsenden Anforderungen ihrer Mitglieder an die Datenverarbeitung und eine nachhaltige Datenablage gerecht zu werden. Damit bestätigen diese internationalen Dateninfrastrukturen den datenorientierten sowie den internationalen Charakter der modernen Wissenschaft.

Von der Vision bis zur Verwirklichung der EOSC zu kommen, bedeutet, Dateninseln und Netzwerke zusammenzulegen und die gängigen Praktiken zahlreicher Wissenschaftsgemeinschaften und Anbieter von IT-Ressourcen zu verstehen. Damit das Projekt EOSC erfolgreich sein kann, müssen Dienste bereitgestellt werden, mit denen potenziell 1,7 Millionen Forschende in Europa ihre Forschungstätigkeiten effektiver durchführen können. Diese Aufgabe hat im Januar 2018 das EOSC-Hub-Projekt⁴ übernommen, das nun

schnell auf die erste technische Realisierung der EOSC zusteuert. Dazu gehören auch Schulung und Dokumentation, die durch die Zusammenarbeit mit dem OpenAire-Projekt⁵ eingebracht wurden.

Auf dem Weg zu einer nachhaltigen EOSC

Die Zukunft und Nachhaltigkeit der EOSC hängt auch von soliden Finanz-, Rechts- und Governance-Modellen ab. Das EOSCsecretariat.eu Unterstützungsprojekt initiiert hierbei gezielt Aktivitäten, die die technische Entwicklung der EOSC unterstützen. Seit Beginn dieses Jahres war das Projekt maßgeblich an

¹ www.eoscsecretariat.eu/

² European Strategy Forum on Research Infrastructures (ESFRI)

³ Am CERN betriebene Teilchenbeschleuniger-Anlage Large Hadron Collider (LHC)

⁴ The EOSC-hub project (www.eosc-hub.eu) contributed the catalogue of services

⁵ www.openaire.eu/

der Einbindung von Forschungsförderorganisationen, Forschungsinfrastrukturen und Forschungsgemeinschaften beteiligt; darunter auch alle ESFRI-Projekte⁶ mit dem Ziel, die organisatorische Entwicklung und den Zusammenhang innerhalb der EOSC gemeinsam zu diskutieren und zu planen. Die zweite Handlungslinie des Projekts ist die Organisation und Unterstützung verschiedener Gremien und damit verbundener Arbeitsgruppen.

EOSC-Gremien und Vertreter

Nach den Empfehlungen des Projekts EOSCpilot wurden Ende 2018 drei Gremien eingerichtet, die die Entwicklung der EOSC auf der politischen und organisatorischen Ebene treiben (Abbildung 1). Die ‚Governance‘-Struktur der EOSC wird ab 2021 darauf basieren.

Das **EOSC Governance Board (GB)** setzt sich aus Vertretern der Mitgliedstaaten und der Europäischen Kommission zusammen und bemüht sich um eine wirksame Aufsicht der EOSC-Implementierung. Vorsitzender des GB ist Dr. Hans-Josef Linkens vom Bundesministerium für Bildung und Forschung (BMBF). Das GB setzte sogenannte „Sub-groups“ ein, um Bereiche von besonderem Interesse zu untersuchen. Im **EOSC Executive Board (EB)** tragen Vertreter der EOSC-Stakeholder dazu bei, die ordnungsgemäße Umsetzung und Rechenschaftspflicht der EOSC sicherzustellen. Mitglieder des EB leiten wiederum sechs „Working Groups“⁷, die sich jeweils auf ein Thema spezialisieren, z. B. die technische Architektur, die Implementierung der FAIR-Prinzipien für Daten, die Erstellung einer Landkarte von Daten-

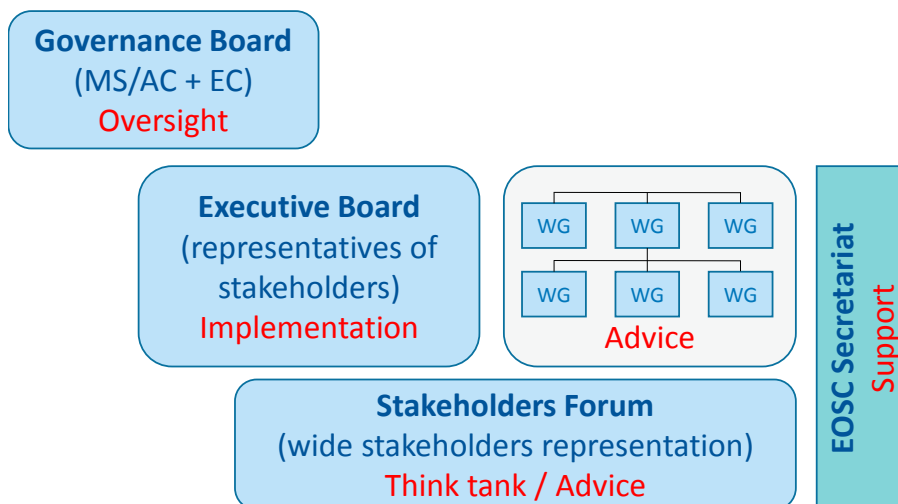


Abbildung 1: Die drei Ebenen in der Governance-Struktur der (pre)EOSC

infrastrukturen in Europa und vieles mehr. Eine vollständige Liste der Arbeitsgruppen und deren Mitglieder findet man auf den Webseiten EOSCsecretariat.eu. Schließlich soll das **EOSC-Stakeholder-Forum** Beiträge von zahlreichen Akteuren aus der Forschung, von Anbietern von IT-Ressourcen und von öffentlichen Organisationen einbringen. Eine gewisse Überschneidung zwischen Mitgliedern des GB, EB und dem Stakeholder-Forum verbessert den Informationsaustausch. Das EB trifft sich jeden Monat sowie regelmäßig mit dem GB.

Ideen, die es zu finanzieren gilt

Eine weitere Aktivität des Projekts besteht darin, kleinere Projekte für die Erforschung von Themen zu finanzieren oder Themen zu lenken, die für die Vorstände oder Arbeitsgruppen von unmittelbarem Interesse sind. Die Fähigkeit, schnell und möglichst agil zu bleiben und auf Entwicklungen zu reagieren, ist durch

den im EOSCsecretariat gelebten „Co-Creation“-Ansatz gegeben. Es stellt den Stakeholdern ein Budget zur Verfügung um Aktionen zur Unterstützung der EOSC-Governance und Studien zum Aufbau der EOSC durchzuführen. Für Ideen, die mithelfen, die EOSC zu verstetigen, kann heute schon eine Finanzierung online beantragt werden⁸. In diesem Sinne ermöglicht das EOSCsecretariat auch Studien und andere kleinere Projekte. Informationen zu den Anforderungen sind auf den Co-Creation-Seiten des Projekts zu finden.

Zusammenfassung

EOSCsecretariat.eu hilft, die Interessengruppen der European Open Science Cloud (EOSC) und ihre Vertreter zusammenzubringen und fördert das Studieren und Diskutieren, um die Zukunft der EOSC zu entwickeln. Ungeachtet der Unterschiede in der Verwaltung und Finanzierung der Forschung in den europäischen Mitgliedstaaten, muss die Arbeit der EOSC-Gremien und -Projekte zu einem weithin akzeptierten Ansatz für die Umsetzung und zu einer länderübergreifenden Organisation der EOSC führen. Damit spielt das EOSCsecretariat.eu eine wichtige unterstützende Rolle bei der Verwirklichung der zukünftigen EOSC.

The EOSC beyond technology

EOSCsecretariat.eu helps to bring EOSC stakeholders and their representatives together and promotes studying and discussing in order to develop the future of EOSC. Irrespective of differences in management and funding of research in the European member states, the work of the EOSC boards and projects must result in a widely accepted approach for the implementation and the cross country organisation of the EOSC. The EOSCsecretariat.eu has a key supporting role in the realisation of EOSC.

⁶ www.esfri.eu/

⁷ www.eoscsecretariat.eu/eosc-working-groups

⁸ www.eoscsecretariat.eu/funding-opportunities

Abgeschlossene Promotionsarbeiten

Zahlreiche Menschen, die am SCC in verschiedenen Wissenschaftsdisziplinen forschen, streben eine Promotion an. Diana Gudu, Nico Schlitter, Ines Reinartz und Felix Bach – auf dem Titelfoto von links nach rechts zu sehen – konnten bis Redaktionsschluss dieser Ausgabe ihre Dissertation in den Disziplinen Informatik und Physik erfolgreich abschließen. Eine weitere Verteidigung steht kurz vor Weihnachten an.

Simulation and Analysis of Protein-Fluorophore Systems for Comparison with Fluorescence Spectroscopy Data

Dr. Ines Reinartz, Fakultät für Physik
am 03.05.2019

Proteine sind die Grundbausteine des Lebens und wichtig für viele Funktionen wie den Transport von Molekülen, Zellbewegungen oder die Katalyse von chemischen Reaktionen. So transportiert das Protein Hämoglobin beispielsweise den Sauerstoff im Blut. Störungen der Proteinfunktionen können schwere degenerative Krankheiten verursachen, darunter die Parkinson-, Huntington- oder Alzheimer-Krankheit. Die Funktionen, Struktur und Dynamik von Proteinen zu verstehen, ist daher ein wichtiges Forschungsgebiet.

Da Proteine für die Lichtmikroskopie zu klein sind, verwendet man stattdessen indirekte Methoden.

Eine dieser Methoden nutzt den Förster-Resonanzenergietransfer (FRET), um Proteindynamik und andere molekulare Prozesse *in vitro* und *in vivo* zu untersuchen. Sie wird auch in Biosensoren zur Konzentrationsmessung von Molekülen wie z. B. Glukose eingesetzt. Die dabei verwendeten Systeme aus Proteinen und Fluorophoren unterliegen physikalischen Prozessen wie Molekulardynamik und Photophysik, die man nicht direkt beobachten kann. Daher ist die Funktion vieler Systeme noch nicht vollständig verstanden.

Computersimulationen können experimentelle Messungen ergänzen und einen Einblick in molekulare Systeme und ihre Funktion geben. Die bisherigen Modellie-

rungsmethoden für Protein-Fluorophor-Systeme sind meist Näherungen, die nur für spezielle Anwendungen verwendbar sind oder, bei Modellierung aller relevanten Bewegungen, zu rechenaufwändig.

Vorgestellt wird eine neue Methode zur Simulation der Dynamik in Protein-Fluorophor-Systemen. Sie basiert auf recheneffizienten vereinfachten Molekulardynamiksimulationen. Mit nur wenigen Parametern bietet sie eine Beschreibung des Systems, die quantitativ mit Experimenten übereinstimmt.

Der Vergleich von Simulationen mit Experimentdaten verbessert die Experimentplanung und -interpretation. Das präsentierte systematische Simulationsprotokoll zur Modellierung von Protein-Fluorophor-Systemen *in silico* kann zur Erforschung vieler biologisch bedeutsamer Anwendungen verwendet werden.

Algorithm Selection in Auction-based Allocation of Cloud Computing Resources

Dr. Diana Gudu, Fakultät für Informatik
am 17.07.2019

Die Merkmale von Cloud Computing wie Pay-as-you-go, On-Demand-Provisioning und klar definierte Service Level Agreements schaffen einen starken Bezug zwischen Nutzung und Kosten. Eine marktorientierte Ressourcenallokation einzuführen, wäre somit konsequent und könnte zudem den Umsatz und die Ressourcennutzung von Anbietern steigern.

Aufgezeigt werden Grenzen aktueller Ansätze: Zum einen bei schwankendem Angebot und Nachfrage, zum ande-

ren in Bezug auf die Bereitstellung von Steuerungsmechanismen für Kunden. Ziel ist es, eine flexiblere und effizientere Ressourcenallokation zu ermöglichen. Als Lösung wird das Allokationsproblem als doppelt-kombinatorische Auktion modelliert. Unter der Annahme, dass es in einem beschränkten Maß akzeptabel ist, Optimalität zugunsten der Geschwindigkeit zu opfern, untersucht diese Arbeit Methoden mit approximativen, heuristischen Algorithmen, die diese Auktionen praktikabel machen.

In der Arbeit werden zwei meta-heuristische Ansätze vorgestellt, welche für eine Eingabe den am besten geeigneten Algorithmus auswählen. Diese Verfahren basieren auf einem Portfolio einer Vielzahl heuristischer Algorithmen aus dem Feld der Auktionstheorie, als auch auf generischen Optimierungsmethoden, die für das Modell angepasst wurden. Es werden systematisch und umfassend vorhandene Arbeiten durch eine gemeinsame Problemformulierung harmonisiert und, nach Anwendung auf eine gemeinsame Basis an Testdaten, verglichen. Eine entscheidende Erkenntnis dabei ist, dass in allen Testfällen kein Algorithmus alleine alle anderen übertrifft. Da Cloud-spezifische Auktionsdaten nicht öffentlich verfügbar sind und bestehende Methoden zur Generierung von Auktionsdaten für die modellierte Auktionsvariante nicht anwendbar sind, wurde ein neuer Ansatz zur Erzeugung von realistischen Inputdaten vorgeschlagen.

Die beiden Ansätze zur Algorithmenauswahl unterscheiden sich von verwandten Arbeiten dadurch, dass sie auf heuristische an Stelle von optimalen Algorithmen

abzielen und bei der Auswahl des besten Algorithmus implizit zwei widersprüchliche Ziele berücksichtigen: sowohl hohe Lösungsqualität als auch schnelle Ausführung. Hierzu wird ein Kostenmodell vorgeschlagen, das diesen Kompromiss quantifiziert und den Algorithmenvergleich im multikriteriellen Raum erleichtert. Die beiden Ansätze ergänzen sich und haben jeweils eigene Vorzüge. Während der erste domänenunabhängig und damit breiter einsetzbar ist, ist der zweite schneller in der Auswahl einer Lösung. Schließlich wird in der Arbeit gezeigt, dass beide Verfahren den für das Problem besten Algorithmus hinreichend genau vorhersagen können und dass sie darüber hinaus besser abschneiden als jeder einzelne Algorithmus alleine.

Neues Konzept zur skalierbaren, explorativen Analyse großer Zeitreihendaten mit Anwendung auf umfangreiche Stromnetz-Messdaten

Dr. Felix Bach, Fakultät für Informatik
am 25.07.2019

Die Arbeit beschäftigt sich mit der Entwicklung und Anwendung eines neuen Konzepts zur skalierbaren explorativen Analyse großer Zeitreihendaten. Die Untersuchung zahlreicher datenintensiver Methoden aus dem Data-Mining und der Zeitreihenanalyse bezüglich deren Skalierbarkeit mit wachsendem Datenvolumen ist ein Augenmerk der Arbeit. Weiterhin stellt sie neue Verfahren und Datenrepräsentationen vor, die eine Exploration sehr großer Zeitreihendaten erlauben. Mit diesen sind die Daten effizienter auswertbar als mit herkömmlichen Methoden.

Methoden zur Verwaltung und Visualisierung großer multivariater Zeitreihen werden mit Methoden zur Detektion seltener und häufiger Muster – sog. Discords und Motifs – kombiniert und zu einem leistungsfähigen Explorationssystem namens

ViAT (von engl. Visual Analysis of Time series) zusammengefasst. Um auch Zeitreihendaten analysieren zu können, deren Datenvolumen Hunderte von Terabyte und mehr umfasst, wurde eine datenparallele, verteilte Verarbeitung auf Basis von Apache Hadoop entwickelt. Sie erlaubt die Ableitung datenreduzierter Metadaten, welche statistische Eigenschaften und neuartige Strukturbeschreibungen der Zeitreihen enthalten. Auf dieser Basis sind neue Anfragen und Auswertungen sowie Suchen nach bekannten und zuvor unbekanntem Mustern in den Daten möglich.

Zudem vorgestellt wird das Design der neuentwickelten Methoden und deren Integration zu einem Gesamtsystem namens FraScaTi (von engl. Framework for Scalable management and analysis of Time series data). Das System wird evaluiert und im Anwendungsfeld der Stromnetzanalyse erprobt, welches von der Skalierbarkeit und den neuartigen Analysemöglichkeiten profitiert. Hierzu wird eine explorative Analyse hochfrequenter Stromnetz-Messdaten durchgeführt, deren Ergebnisse im Kontext des Anwendungsbereichs präsentiert und diskutiert werden.

Analyse verteilter Daten unter Berücksichtigung von Korrektheit, Skalierbarkeit und Vertraulichkeit

Dr. Nico Schlitter, Fakultät für Informatik
am 05.11.2019

Big Data Datenanalysen sind mit großen Herausforderungen verbunden. Die Daten sind großvolumig und können aus unterschiedlichen Quellen stammen, sind aber gleichzeitig strategisch oder wirtschaftlich wertvoll und sollen daher nicht öffentlich bekannt werden. Diese Arbeit fokussiert deshalb auf Skalierbarkeits- und Vertraulichkeitsaspekte der Analyse großer Datenmengen.

Zum einen wurde ein Verfahren entwickelt, das eine vertraulichkeitserhaltende Analyse von horizontal partitionierten Datensätzen mithilfe eines verteilt trainierten Neuronalen Backpropagation-Netzwerks ermöglicht. Grundlage ist eine lokale Berechnung der Gewichtsänderungsmatrizen und deren anschließende Integration in ein Gesamtmodell.

Zum anderen wird eine neue, vertraulichkeitserhaltende Methode zur Erzeugung von Klassifikationsmodellen auf vertikal partitionierten Datensätzen vorgestellt. Der GenPAC-Ansatz beruht auf einer durch lokales Clustering erzielten Maskierung der vertraulichen Daten und deren anschließender Klassifikation. Das lokal einzusetzende Clusteringverfahren und das global anzuwendende Klassifikationsverfahren kann frei gewählt werden. Das untersuchte Verfahren ist skalierbar und zeigt mit Benchmark-Datensätzen eine hohe Klassifikationsgüte.

Die Arbeit erweitert zudem einen Algorithmus, der ursprünglich zur Entdeckung von Communities in sozialen Netzwerken entwickelt wurde und stellt damit neue Verfahren vor, die das bis dahin inhärente Problem der Bestimmung geeigneter Dichteparameter lösen. Während DenGraph ein flaches Clustering liefert, dessen Charakteristik stark von der Parameterwahl abhängt, berechnet DenGraph-HO eine Hierarchie von Clustern. Eine bei inkrementellen Graphänderungen effiziente Anpassung des hierarchischen Clustermodells erlaubt DenGraph-HIO und trägt somit zur performanten dichte-basierten, hierarchischen Clusteranalyse für Graphen bei. Die Variante DenGraph-PHO parallelisiert die Algorithmen DenGraph-HO und DenGraph-HIO und sorgt so für die erwünschte vertikale Skalierbarkeit bei großen Graphen.

Innovationsteam des KIT gewinnt NCATS ASPIRE Design Challenge Preis

Bei einem amerikanischen Wettbewerb zu Chemiedatenbanken hat ein Innovationsteam des KIT mit der Idee für ein elektronisches Synthetische-Chemie-Portal (eSCP) als Informationssystem zum Speichern, Sammeln, Teilen, Suchen und Analysieren von Chemiedaten einen Preis gewonnen.

Felix Bach

Das Innovationsteam des KIT, bestehend aus Nicole Jung, Stefan Bräse, Pierre Tremouilhac, Felix Bach und Dr. Ravindra Peravali konnte in einem internationalen Wettbewerb der National Institutes of Health (NIH), einer Behörde des US-amerikanischen Gesundheitsministeriums, überzeugen und gewann einen Preis der Kategorie "Challenge 2". In der Ausschreibung mit dem Titel "A Specialized Platform for Innovative Research Exploration (ASPIRE)" wurden Konzepte gesucht, die Möglichkeiten zur Etablierung von moderner Forschungsinfrastruktur aufzeigen. In der Kategorie 2 wurden Vorschläge zur Ausarbeitung eines elektronischen Syntheseportals gesucht.

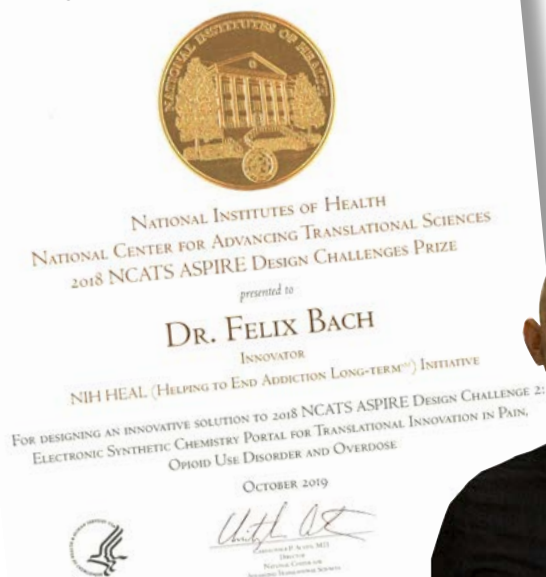
Das Ziel der Ausschreibung des US National Center for Advancing Translational Sciences (NCATS), Teil des NIH, ist es, innovative Ansätze zur Lösung der Opioidkrise in den USA durch die Entwicklung von (1) neuartigen chemischen Verbindungen, (2) Data Mining, Analysewerkzeugen und -technologien und (3) biologischen Assays zu belohnen und voranzutreiben. Hierdurch soll die Entwicklung von präklinischen Tests sicherer gemacht und die Entdeckung von Schmerzmittel der nächsten Generation, die keine Abhängigkeit auslösen, gefördert werden. Damit sollen neue Therapiemöglichkeiten bei Opioidabhängigkeit und Überdosierung identifiziert werden.

Das Team des KIT wurde für die Einreichung "Electronic Synthetic Chemistry Portal for Translational Innovation in Pain, Opioid Use Disorder and Overdose" prämiert. Das vom Innovationsteam vorgeschlagene elektronisch Portal für synthetische Chemie (eSCP) ist ein Informationssystem zum Speichern, Sammeln, Teilen, Suchen und Analysieren von Daten – speziell für synthe-

tische und medizinische Chemiker und Biologen. Das Portal ist auf Basis der Labor-Anwendungssoftware Chemotion konzipiert, welches ein Elektronisches Laborbuch (ELN) und ein Repository für Chemiker beinhaltet. Chemotion ist ein gemeinsames DFG-gefördertes Projekt des SCC, der Bibliothek und des Instituts für Organische Chemie des KIT. Es umfasst Funktionen zur Dokumentation der Arbeit im Labor, zur Suche nach chemischen Reaktionen und Molekülen sowie zur Verarbeitung und Offenlegung der gesammelten Daten nach den FAIR-Data Prinzipien. Die Idee ist es – auf den existierenden Arbeiten aufbauend – eine verteilte Struktur aus einzelnen ELN zu entwickeln, die alle interoperabel mit einem zentralen eSCP verbunden sind. So wird das eSCP die synthetische Forschung beschleunigen, indem es eine Infrastruktur zur Speicherung und gemeinsamen Nutzung chemischer Reaktionen bereitstellt und langfristig zu einer breit aufgestellten Datenbank mit Reaktionen und Analysen führen wird. Mittels Machine Learning (ML) können aus diesen Daten Modelle für Retrosynthese und Reaktionsprognose gewonnen werden, welche die synthetische Planung erleichtern. Das eSCP wird die Entdeckung bislang unbekannter chemischer Strukturen unterstützen, die zur Behandlung von Schmerzen, Opioidmissbrauch und Überdosierungen einsetzbar sind, indem es eine Infrastruktur mit verschiedenen Algorithmen und Vorhersagemodellen für Verbindungseigenschaften bietet. Auf diese Weise werden Innovationen durch die interdisziplinäre Arbeit in Chemie und Biologie in einer Plattform vereint und zudem unterstützt durch einen barrierefreien Austausch von Forschungsdaten. Das Konzept soll auf andere fachspezifische oder generische Projekte übertragbar sein.

KIT Innovation Team Wins NCATS ASPIRE Design Challenge Prize

In an American competition on chemical databases, an innovation team from KIT has won a prize with the idea of an electronic synthetic chemistry portal (eSCP) as an information system for storing, collecting, sharing, searching and analyzing chemical data.



Weitere Informationen

ncats.nih.gov/aspire/2018ChallengeWinners

MathSEE Modeling Week – Studierende präsentieren Ergebnisse

Studierende eines mathematisch-methodischen Faches und der „SEE-Fächer“ – Sciences, Engineering and Economics – nahmen vom 18. bis 23. August 2019 an der MathSEE Modeling Week in Bad Liebenzell teil. Das KIT-Zentrum „MathSEE“ bündelt seit Oktober 2018 die interdisziplinäre mathematische Forschung am KIT und richtet erstmals eine Modeling Week für Studierende aus. In der öffentlichen Abschlussveranstaltung am 23. August 2019 am SCC verfolgten wissenschaftliche Mitarbeiter/innen mit großem Interesse die Präsentationen und Erkenntnisse der Studierenden.

Kirsten Wohak

Am 18. August 2019 sammelten sich das erste Mal Studierende verschiedener Studiengänge in Karlsruhe, um an der MathSEE Modeling Week 2019 teilzunehmen (Bild 1). Gemeinsam fuhren alle mit einem Reisebus zur Burg Bad Liebenzell, wo die nächsten Tage mit der Erforschung spannender Problemstellungen aus der realen Forschung des KIT-Zentrums MathSEE verbracht werden sollten.

Nach einem ausgiebigen ersten Kennenlernabend ging die Arbeit in Gruppen (Bild 2) am Montagmorgen los. Die betreuenden wissenschaftlichen Mitarbeitenden Kevin Daiß (FAST), Philipp Hamberger, Pia Stammer und Tianbai Xiao (SCC) führten in die Problemstellungen ein. Die Studierenden beschäftigten sich mit chemischen Prozessen beim Säurenebel, dem Einfluss der Erdkrümmung auf riesigen Ackerflächen, der optimalen Formfüllung eines Bauteils und der automatischen Verfolgung der Höhe der Tropopause, eine Größe, die für Mischungsprozesse in der Atmosphäre von Bedeutung ist. Die Problemstellungen sind sehr breit gefächert, haben jedoch alle eine Gemeinsamkeit: Sie be-

schreiben allesamt ungelöste Fragen aus der aktuellen Forschung. Während der Woche haben die Studierenden die mathematische Modellierung kennengelernt und ausgiebig angewendet. So erhielten sie einen Einblick in den Forschungsalltag an einer Universität und mussten auftretende Probleme selbst lösen.

Am 23. August 2019 präsentierten die Studierenden ihre beeindruckenden Ergebnisse in einem öffentlichen Vortrag am SCC vor Problemstellern und Problemstellerinnen und weiteren interessierten Personen des KIT. Rudolf Lohner, Experte für energieeffiziente Rechenzentren, führte die Studierenden zum Abschluss durch den Hochleistungsrechner ForHLR.

Allen, die bei dieser Woche mitgewirkt haben, indem sie ein Problem gestellt, die Werbetrommel gerührt oder bei den Vorbereitungen geholfen haben, sei an dieser Stelle noch einmal ein riesengroßes Dankeschön ausgesprochen!



Bild 1



Bild 2

Students present their results from the MathSEE Modeling Week

Students of a mathematical-methodical subject and the „SEE subjects“ – Sciences, Engineering and Economics – participated in the MathSEE Modeling Week in Bad Liebenzell from August 18th – 23rd 2019. Since October 2018, the KIT Center "MathSEE" has been bundling interdisciplinary mathematical research at KIT and organizing a Modeling Week for students for the first time. With the help of scientific assistants, the students dealt with chemical processes in acid mist, the influence of the earth's curvature on huge farmland, the optimal filling of a component's shape and the automatic tracking of tropopause, a boundary layer in the atmosphere that is important for climate observation. At the closing event on August 23rd 2019 at Campus North academic staff followed the findings and results of the students with great interest.

Aktionstag am KIT zum European Cyber Security Month 2019

Das KIT beteiligte sich am European Cyber Security Month 2019 mit einem Info-Stand und aktuellen Materialien. Der Fokus lag auf Praxistipps zur IT-Sicherheit und der Erkennung von betrügerischen Nachrichten sowie dem sicheren Umgang mit diesen.

Melanie Volkamer¹, Milan Burgdorf², Andreas Lorenz



Das KIT beteiligte sich am 23. Oktober 2019 am diesjährigen European Cybersecurity Month 2019. Dieser findet seit 2012 jährlich statt und ist eine Aktion der EU-Mitgliedstaaten unter Federführung der ENISA³, um das Bewusstsein für Cybersicherheit zu fördern.

Der Fokus am KIT lag in diesem Jahr auf Praxistipps zur IT-Sicherheit und der Erkennung von betrügerischen Nachrichten sowie dem sicheren Umgang mit diesen. Die Beschäftigten und Studierenden wurden mit einer Rundmail sowie zwischen 11 und 14 Uhr mit einem Info-Stand im Casino des Campus Nord auf das Thema „Betrügerische Nachrichten“ und entsprechende IT-Sicherheitstipps aufmerksam gemacht.

Die Mitarbeiterinnen und Mitarbeiter der beteiligten Einrichtungen und Forschungsgruppen des KIT (AIFB, KASTEL, SECUSO, DO, SCC und ZML) erläuterten in vielen Gesprächen die sogenannten betrügerischen Nachrichten, wozu auch die klassischen Phishing E-Mails gehören. Diese sind mittlerweile immer schwerer zu erkennen, weil Cyberbetrüger sehr trickreich vorgehen. Nach dem Aktionstag anlässlich des „Safer Internet Day“⁴ am Campus Süd im Februar 2019 standen betrügerische Nachrichten und der sichere Umgang mit ihnen auch an diesem Tag im Vordergrund.

Mit Praxisbeispielen in Flyern, Postern, Infokarten und zwei Videos konnte am Info-Stand anschaulich gezeigt werden, wie betrügerische Nachrichten zu erkennen sind und welche Meldeverfahren es am KIT dafür gibt. Außerdem wurde die freiwillige Online-Schulung „NoPhish“ zur Vertiefung der Thematik für die Beschäftigten und Studierenden überarbeitet und in einer neuen Version auf der Lernplattform des KIT bereitgestellt.

Des Weiteren beantwortete das Aktionsteam (s. Bild oben) allgemeine Fragen zur IT-Sicherheit, hielt den Flyer „Praxistipps IT-Sicherheit am KIT“ bereit und informierten über die Meldepflicht von IT-Sicherheitsvorfällen im Rahmen des IT-Sicherheitskonzepts⁵. Alle Dokumente und Materialien sind über die Aktionswebseite⁶ oder den Servicedesk des SCC verfügbar.

Aufgrund der durchweg positiven Rückmeldungen auf diesen Aktionstag planen die beteiligten Einrichtungen und Forschungsgruppen, sich erneut am internationalen Safer Internet Day am 11. Februar 2020 mit Aktionen am Campus Süd des KIT zu beteiligen.

¹ Institut für Angewandte Informatik und Beschreibungsverfahren (AIFB), Forschungsgruppe Security-Usability-Society (SECUSO)

² Digital Office (DO), Informationssicherheitsbeauftragter des KIT

³ Europäische Agentur für Netz- und Informationssicherheit

⁴ www.klicksafe.de/ueber-klicksafe/safer-internet-day

⁵ s.kit.edu/it-sicherheitskonzept

⁶ s.kit.edu/ecsm19



Teilnehmerinnen und Teilnehmer der 17. GridKa School 2019

17. Internationale GridKa School – The Art of Data

„The Art of Data“, die Kunst, optimal mit komplexen und großen wissenschaftlichen Datensätzen umzugehen, stand im Mittelpunkt der 17. internationalen GridKa School, die vom 26. bis 30. August am KIT stattfand. Insgesamt nahmen mehr als 130 Wissenschaftlerinnen und Wissenschaftler aus 11 Nationen an 13 Stunden Plenarvorträgen und bis zu 81 Stunden Workshops teil.

René Caspart

Auch in diesem Jahr blickt das SCC auf eine erfolgreiche GridKa School zurück. Sie fand vom 26. bis 30. August am Karlsruher Institut für Technologie unter dem Thema „The Art of Data“ statt. Die Agenda der Schule umfasste 13 Stunden Plenarvorträge und bis zu 81 Stunden Workshops, in denen die Teilnehmenden Wissen aus den Bereichen Open Science, moderne Programmiersprachen und -techniken, GPU-Programmierung und Machine Learning gewinnen und vertiefen konnten. Insgesamt nahmen in diesem Jahr mehr als 130 Wissenschaftlerinnen und Wissenschaftler aus 11 Nationen teil. Die Agenda der GridKa School ist unter <http://gridka-school.scc.kit.edu/2019/> verfügbar.

Bei den Workshops wurden in diesem Jahr auch diverse Computing-Ressourcen genutzt. Diese umfassten neben einer dedizierten, am SCC gehosteten OpenStack-Instanz für die GridKa School auch HPC-Ressourcen am SCC und dem Jülich Supercomputing Centre (JSC). Mit diesen Ressourcen konnten die Teilnehmenden die Workshop-Inhalte direkt selbst anwenden und eigene Erfahrungen sammeln.

Neben den Plenarvorträgen und Tutorien bot die GridKa School den Teilnehmenden dieses Jahr wieder viele Gelegenheiten für Diskussionen, interdisziplinären Austausch und Networking.

Unter anderem fand der traditionelle Flammkuchenabend mit Führungen durch die Anlagen des Hochleistungsrechners ForHLR II und des Grid Computing Centre Karlsruhe (GridKa) großen Anklang.

Wie in den vergangenen Jahren war auch in diesem Jahr der als offenes SCC-Kolloquium organisierte Abendvortrag ein besonderes Highlight. Für den Vortrag zu dem Thema „Why the future of weather and climate prediction will depend on supercomputing, big data handling and artificial intelligence“ konnte Dr. Peter Bauer vom European Centre for Medium-Range Weather Forecasts (ECMWF) gewonnen werden. Er stellte in seinem Vortrag die Herausforderungen für eine zuverlässige Wettervorhersage und die dabei aktuell verwendeten Systeme vor.

Er zeigte sehr anschaulich auf, dass für Wettervorhersagen in Zukunft nicht nur leistungsfähigere HPC-Systeme notwendig sein werden, sondern auch verbesserte Technologien für den Umgang mit großen Datenmengen sowie mit künstlicher Intelligenz ausgestattete Softwaresysteme.

Die nächste GridKa School wird vom 31. August bis 4. September 2020 am Karlsruher Institut für Technologie stattfinden.

17th International GridKa School – The Art of Data

"The Art of Data", the art of dealing optimally with complex and large scientific data sets, was the focus of the 17th international GridKa School, which took place from 26 to 30 August at KIT. In total, more than 130 scientists from 11 nations participated in 13 hours of plenary lectures and up to 81 hours of workshops.



Die GridKa School bot den Teilnehmerinnen und Teilnehmern ein breites Programm an Plenarvorträgen und Tutorien

CAMMP in Mexiko

Vom 17. bis 26.10.2019 führten Vertreter/innen der Projekte CAMMP und KOMMS Schülerinnen und Schüler sowie Studierende der Universidad Autónoma de Baja California in Tijuana/Mexiko in die Problemlösung mit Hilfe mathematischer Modellierung ein.

Kirsten Wohak

Vom 17. bis zum 26. Oktober waren Martin Bracke (TU Kaiserslautern), Maren Hattebuhr (KIT), Sarah Schönbrodt (KIT), Janna Tinnes (RWTH Aachen) und Kirsten Wohak (KIT) als Vertreter/innen der Projekte CAMMP¹ (Computational And Mathematical Modeling Program) und KOMMS² (Kompetenzzentrum für mathematische Modellierung in MINT-Projekten in der Schule) an der Universidad Autónoma de Baja California (UABC³) in Tijuana zu Besuch. Dort erhielten mexikanische Teilnehmende in Form von mehreren Modellierungstagen und einer parallel dazu laufenden Modellierungswoche Einblicke in das Problemlösen mit Hilfe mathematischer Modellierung.

Zuerst fand eine Einführung in die Idee und Ziele von CAMMP und KOMMS für die ausshelfenden wissenschaftlichen Mitarbeitenden von vor Ort statt. Mit Beginn der neuen Woche starteten wir mit ihrer Unterstützung in eine Reihe erfolgreicher eintägiger Workshops und eine Modellierungswoche.



Die ca. 160 Teilnehmenden waren motiviert, begeistert und interessiert dabei. Zudem nahmen sie aus diesen neuen Erfahrungen viel Wissenswertes mit und konnten ihre Modellierungskompetenzen ausbauen. Auch wir haben viele lehrreiche und spannende Erfahrungen gesammelt, während wir mit Menschen einer anderen Kultur zusammenarbeiten durften.

Ein großes Dankeschön gilt dabei dem Organisator Luis Ramon Siero Gonzalez (UABC) und Martin Frank, wissenschaftlicher Leiter des Projekts CAMMP am SCC. Es war für uns eine erfolgreiche und sehr erlebnisreiche Zeit.



Bild 1: v. l. n. r.: Sarah Schönbrodt (KIT), Martin Bracke (TU Kaiserslautern), Kirsten Wohak (KIT), Marcel Marnitz, Luis Ramon Siero Gonzalez (UABC), Maren Hattebuhr (KIT) und Janna Tinnes (RWTH Aachen).

CAMMP in Mexico

From 17 to 26 October 2019, representatives of the CAMMP and KOMMS projects introduced students of the Universidad Autónoma de Baja California in Tijuana/Mexico to problem solving using mathematical modelling.



Bild 2: Martin Bracke (TU Kaiserslautern) mit mexikanischen Betreuern für die Modellierungswoche

¹ www.scc.kit.edu/forschung/CAMMP.php

² kommms.uni-kl.de

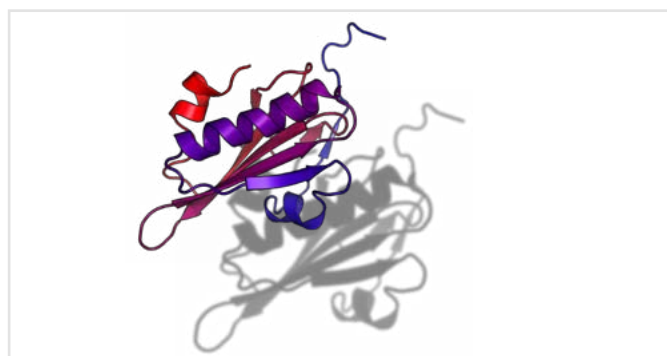
³ www.uabc.mx

Marie Weiel-Potyagaylo erhält den Erna-Scheffler-Förderpreis 2019

Den Erna-Scheffler-Förderpreis 2019 für eine Masterarbeit erhält die Physikerin Marie Weiel-Potyagaylo für ihre am SCC verfasste Arbeit mit dem Titel „Integration of SAXS Data into Biomolecular Simulations“. In dieser Arbeit zeigt sie wie Daten aus Experimenten zur Analyse der dynamischen Struktur von Proteinen durch die Biosimulation so ergänzt werden, dass sie als eine Art virtuelles Mikroskop hochaufgelöste Einblicke in atomare Details molekularer Systeme bieten.

Achim Grindler

Mit Small-Angle X-Ray Scattering (SAXS) oder Kleinwinkel-Röntgenstreuung lassen sich dynamische Strukturen von Makromolekülen wie zum Beispiel Proteinen analysieren. Diese Biomoleküle fungieren im Körper als molekulare Werkzeuge und erfüllen, abhängig von ihrer individuellen Struktur, unterschiedlichste Aufgaben. Sie ermöglichen beispielsweise Zellbewegungen oder beeinflussen den Zellstoffwechsel. „In einem SAXS-Experiment bestrahlt Röntgenlicht in Wasser gelöste Proteine. Dabei registriert ein Detektor die Streustrahlung. Die so gemessene Intensitätsverteilung liefert Informationen über die durchschnittliche Größe und Gestalt der enthaltenen Moleküle“, sagt Marie Weiel-Potyagaylo. „Mithilfe von Computersimulationen können wir daraus dann ein dreidimensionales Molekülmodell ableiten. Solche Simulationen bieten als eine Art virtuelles Mikroskop hochaufgelöste Einblicke in atomare Details molekularer Systeme.“



In einem SAXS-Experiment bestrahlt Röntgenlicht in Wasser gelöste Proteine. Ähnlich einem Schatten liefert die gemessene Intensitätsverteilung der Streustrahlung Informationen über die durchschnittliche Größe und Gestalt der Moleküle.

Marie Weiel-Potyagaylo hat in ihrer Masterarbeit experimentelle Daten aus SAXS-Messungen in Simulationen eingebunden. Dies ermöglicht die strukturelle Interpretation von SAXS-Daten innerhalb der Simulationen und liefert ein detailliertes Bild der Dynamik und damit der Funktion der untersuchten Biomoleküle. Marie Weiel-Potyagaylo forscht in der Junior Research Group

Multiscale Biomolecular Simulation des SCC. Zum zweiten Mal erhält nun eine Wissenschaftlerin der Forschungsgruppe um Alexander Schug die alle zwei Jahre verliehene Auszeichnung für eine Masterarbeit. 2015 ging der Preis an Ines Reinartz, die in diesem Jahr erfolgreich ihre Doktorprüfung abgelegt hat.



Die Preisträgerinnen, daneben Vertreter des KIT und von Soroptimist International

Marie Weiel-Potyagaylo receives the Erna Scheffler Award 2019

Marie Weiel-Potyagaylo received the Erna Scheffler Award 2019 for her master thesis "Integration of SAXS Data into Biomolecular Simulations" at SCC. In this paper she shows how data from experiments to analyze the dynamic structure of proteins can be augmented by biosimulation to provide high-resolution insights into atomic details of molecular systems as a kind of virtual microscope.

Dieser Text enthält Auszüge aus der KIT-Presseninformation:
www.kit.edu/kit/pi_2019_131_erna-scheffler-foerderpreis-fuer-forscherinnen-des-kit.php
 Pressebericht mit Bild von Preisverleihung:
ESFP_2019_Bericht_SI_KA_Homepage_mit_Bild.pdf

Bunte Nacht der Digitalisierung – Impressionen

Bürgerinnen und Bürgern die mannigfaltigen Aspekte der Digitalisierung erleb- und erfahrbar zu machen, war das Ziel der Bunten Nacht der Digitalisierung. Sie fand am 11. Oktober in Karlsruhe statt. Mit Wissenschaftsvorträgen (Bilder 1 und 2) und Führungen durch die Forschungsinfrastruktur (Bilder 3 und 4) sowie Informationsständen zu Projekten (Bilder 5 – 7), stellte das SCC einen Ausschnitt aus seinem Aufgabenspektrum vor.



Bild 1



Bild 2



Bild 3



Bild 4



Bild 5



Bild 6

Neues aus den SCC-Abteilungen

Abteilung Scientific Computing & Simulation (SCS)



Jennifer Schröter

Zum 1. November übernahm Dr. Jennifer Schröter die Leitung der Abteilung Scientific Computing & Simulation des SCC.

Jennifer Schröter hat Physik an der Rheinischen Friedrich-Wilhelms-Universität Bonn studiert. An der Uniklinik Bonn hat sie im Forschungsfeld der Neurophysik erste Erfahrungen im Bereich der numerischen Simulationen und des Betriebs von HPC-Systemen. 2013 kam Jennifer Schröter an das KIT, um an der Fakultät für Physik zu promovieren. 2017 schloss sie die Promotion erfolgreich ab und arbeitete bis Oktober 2019 am Institut für Meteorologie und Klimaforschung.

Sie gehört zu den Hauptentwicklern des Wetter- und Klimamodells ICON(-ART), welches das KIT gemeinsam mit Wissenschaftlerinnen und Wissenschaftlern des Deutschen Wetterdienstes (DWD), des Max-Planck Instituts für Meteorologie (MPI-M) und des Deutschen Klimarechenzentrums (DKRZ) erstellt und fortentwickelt. Zuletzt war sie für die technische Umsetzung und Koordination der kampagnenbegleitenden Simulationen verantwortlich.

Dr. Olaf Schneider, der bisherige Leiter der Abteilung SCS, nimmt zum 1. Dezember 2019 eine neue Herausforderung als Leiter des Universitätsrechenzentrums der Friedrich-Schiller-Universität Jena an. Im Zuge der Neubesetzung der Abteilungsleitung SCS wurde der Zuschnitt der Abteilung angepasst. Die Abteilung SCS umfasst weiterhin den HPC-Betrieb, das Performance Engineering und die Betreuung wissenschaftlicher Software. Die Aktivitäten der Simulation Labs und der Research Groups Computational Science and Mathematical Methods werden künftig in einer neuen Abteilung Scientific Computing & Mathematics (SCM) zusammengefasst.



Bild 7

IMPRESSUM

SCC news

Magazin des Steinbuch Centre for Computing

Herausgeber

Präsident Professor Dr.-Ing. Holger Hanselka
Karlsruher Institut für Technologie (KIT)
Kaiserstraße 12
76131 Karlsruhe

Anschrift

Steinbuch Centre for Computing (SCC)
Karlsruher Institut für Technologie (KIT)
Redaktion SCC-News
Zirkel 2
76131 Karlsruhe
oder:
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen
Fax: +49 721 608-24972

Redaktion

Achim Grindler (verantwortlich),
Karin Schäufele, Andreas Ley
E-Mail: redaktion@scc.kit.edu

Gestaltung, Satz und Layout

Hella Grolmus
AServ – Crossmedia – Grafik
Karlsruher Institut für Technologie (KIT)
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen

Titelfoto

Patrick Langer

Fotos

Achim Grindler, Fotolia, Christoph Heidecker, Tanja Meißner, Friederike Pasquay, Radek Januszewski, Anastasiya Sultanova

Druck

Systemedia GmbH, 75449 Wurmberg

Erscheinungstermin dieser Ausgabe

Dezember 2019

www.scc.kit.edu/publikationen/scc-news

Der Nachdruck und die elektronische Weiterverwendung sowie die Weitergabe von Texten und Bildern, auch von Teilen, sind nur mit ausdrücklicher Genehmigung der Redaktion gestattet.



Karlsruher Institut für Technologie (KIT)
Steinbuch Centre for Computing (SCC)

ISSN: 1866-4954

www.scc.kit.edu
contact@scc.kit.edu