

Datenschutz und Datenhoheit in Forschung und Lehre

Gerhard Schneider
Albert-Ludwig-Universität Freiburg

21.10.2014 DINI Bonn

Albert-Ludwigs-Universität Freiburg



UNI
FREIBURG

Es war einmal...

- Das Internet war eine aufstrebende Technologie, die
 - nur wenige verstanden und die
 - nur wenige mehr nutzten
 - die alle toll fanden und deshalb die Ressource verteidigten
- Verteidigen einer Ressource?
 - Noch 1993 hieß es beim DFN: **2 Mbit/s** brauchen nur wenige große Unis, für die anderen reichen 64 kbit/s
 - Angst vor den Kanzlern, die das Geld für schnelle Netze nicht herausgaben (mangels Verständnis)
- Weltweite Wild-West-Methoden
 - Nutzer, der sich daneben benahm, wurde vom Admin gesperrt
 - Eigene „Gerichtsbarkeit“!
 - SPAM war kein Thema – wer es versuchte, war „draußen“

22.10.2014

Datenschutz und Datenhoheit

2

Eigene Regierung



- Gelebte Sozialutopie
 - ICANN, RIPE, DeNIC, etc funktionierten, ohne dass sich „die klassischen Verdächtigen“ einmischten
 - Einst hatte .de ganze 800 Domänen.
 - „technische“ Gesetze auf Basis von RFCs
 - Man diskutierte solange, bis der „Standard“ stand
 - Erschöpfung, Frust der Andersdenkenden, geniale Lösung, schiere Notwendigkeit, etc.
 - „technisch vernünftig“ geht leichter als „sozial gerecht“
 - Man hielt sich daran – weil sonst nichts funktionierte
 - Und die Außenstehenden keine Ahnung hatten
 - Gerichte, Anwälte, Lobbyisten, Regierungen, Besserwisser,
- Neuartige Geschäftsmodelle
 - *Best effort* = keine einklagbare Zusage möglich

22.10.2014

Datenschutz und Datenhoheit

3

Im Kielwasser...



- Das Gedankengut des Internet (im Sinne des „Betriebs“ eines Datennetzes) hatte Folgen.
- Wenn man der Allgemeinheit „Netzkapazität“ zur Verfügung stellt, warum nicht auch Programmcode?
 - Frage nach den Rechten und Pflichten
 - Wer verantwortet Fehler → Reparaturpflicht?
 - Idee des „best effort“ → *public domain*
 - Umgekehrtes Problem: wie verhindert man, dass die „klassischen Verdächtigen“ sich daraus bedienen?
 - Copyright, copyleft, Lizenzbedingungen, etc wurden entwickelt
 - Und andere damit vor Gericht gezerrt – mit Erfolg
- Entwicklungen in einem von Ahnungslosigkeit geschützten Reservat

22.10.2014

Datenschutz und Datenhoheit

4

Im Kielwasser...

- Plötzlich: Firmen zahlten OpenSource-Programmierer
 - Um sich selbst auch aus der Haftung zu nehmen
 - Um Kosten zu sparen (falls nicht *mission critical*)
 - Um die Komplexität der Entwicklung zu umgehen
- Die Idee der „Allmende“ gewann unerwartete Freunde
- Plötzlich galt der zweite bayerische Verwaltungsgrundsatz: *dös hamma scho imma so gmacht*
 - *Ich darf witzeln – ich habe die bayerische Staatsangehörigkeit*
- Dann wuchs das Internet schneller, als man mit der „Erziehung“ der Nutzer hinterher kam.
 - Und dann war die schöne Zeit vorbei ☹ ☹
 - Der Rechtsstaat versuchte, Einzug zu halten – wäre auch schon viel früher möglich gewesen, wenn...

22.10.2014

Datenschutz und Datenhoheit

5

Folgen

- Andere (die Mächtigen) versuchten, zu erklären, wie das Internet zu funktionieren hat
 - Sperrverfügungen
 - Abmahnwellen
 - Aber: Wasser fließt nicht bergauf, auch wenn der König dies fordert.
- Mittlerweile ist die Kompetenz an vielen Stellen (leicht?) gestiegen
 - Das hilft, denn es geht nun um echte Fragen, die nicht leicht zu lösen sind.
 - Beispiel:
Ein U.S. District Court hat in einem Urteil vom 25. April 2014 den Datenschutz für ausländische Kunden amerikanischer IT-Firmen infrage gestellt. Demnach müssen US-Internet-Firmen ihren Behörden auf Verlangen auch dann vollen Zugriff auf Daten ihrer Kunden verschaffen, wenn sich die betroffenen Rechenzentren außerhalb der Vereinigten Staaten befinden

22.10.2014

Datenschutz und Datenhoheit

6

??????

- Das Thema:
wenn das so ist und uns unser eigener Rechtsstaat nicht schützen kann, wie schützen wir unsere Daten und wie behalten wir darüber die Kontrolle?
 - Nicht im kommerziellen Sinne (Urheberrecht, Verwertungsrecht)
 - Sondern im technischen Sinne
 - Was einmal gedacht wurde, kann nicht mehr zurückgenommen werden (Dürrenmatt, Die Physiker)
 - Was der andere weiß (erfahren hat – wie und warum auch immer), wird er nicht mehr vergessen
- Ausgestattet mit der alten Sozialutopie stolpern wir von einem Problem zum nächsten

22.10.2014

Datenschutz und Datenhoheit

7

Alltagssituationen

- Vorweg:
- Ich bin kein Jurist!!!
 - Ich habe lediglich stellenweise ein gefährliches (?) Amateurwissen
- Es folgen verschiedene Fallbeispiele
 - Zur Sensibilisierung
 - Kein Lösungsvorschlag ist wirklich rechtlich belastbar
 - Außer, falls offensichtlich
- Zeigt lediglich: es eröffnet sich nun ein Minenfeld.

22.10.2014

Datenschutz und Datenhoheit

8

Datenschutz



- Einfache Fälle:
 - Unsere Nutzer vor anderen Neugierigen schützen
 - MatNr und Klausurergebnisse...
 - Warum? Zeugenschutzprogramm / Ehrenmord / Annäherungsverbot (Gewalt in der Familie, Stalking)
 - Unis sind groß und deshalb gibt es solche Fälle rein statistisch auch an Unis
 - Diese Menschen sind zu schützen.
 - Daneben aber noch viele einfachere Fälle
 - Abmahnschreiben? (Tauschbörsen)
 - Unsere Nutzer...
 - Wie sieht es damit an den Fakultäten und Lehrstühlen aus?

22.10.2014

Datenschutz und Datenhoheit

9

Logdateien



- Logfiles:
 - Einst zur Fehlersuche: wann/was/wo/warum?
 - Hilft bei der Jagd nach Hackern → „cuckoo's egg“
 - Inzwischen: Selbstverteidigung des RZ gegen Nutzer
 - Nachweis: „Sie“ haben es gemacht, nicht das RZ.
 - Nebenwirkungen: Logfiles können beschlagnahmt werden
 - Sind Logfiles angesichts der resultierenden Überwachung der Nutzer noch zeitgemäß?
 - „Ich weiß, was Du letzten Sommer gemacht hast“.
- Was müssen wir eigentlich aufheben?
 - Wie weit geht die Nachweispflicht, auch bei Pressehatz?

22.10.2014

Datenschutz und Datenhoheit

10

Plagiatserkennung – Sündenfall?



- **Misstrauensgesellschaft**
 - Zumindest, wenn alles automatisch geprüft wird.
- **Marktmacht:**
 - Können Sie es verantworten, dass Ihre Uni keine automatische Prüfung macht?
- **Ursprung der Software:**
 - US-Unis – wenn 500 Studis in einem Fach jede Woche einen essay abliefern müssen, und das das ganze Semester lang
 - Kohorten-Fahndung in den Schreibwissenschaften
- **Klappt nicht so einfach bei normalen Masterarbeiten**
 - Darf eine Masterarbeit einfach so der Plagiatsprüfungssoftware einer ausländischen Firma übergeben werden?
 - Die Firma will ihre Fahndungsdatenbank damit befüllen!

22.10.2014

Datenschutz und Datenhoheit

11

Datenhoheit - Vorbetrachtung



- **Dürfen wir die Informationen überhaupt aus unseren Daten extrahieren?**
 - Um einen Text zu lesen, brauchen Sie heute:
 - Windows-Lizenz, Office-Lizenz, Adobe-Lizenz
 - Sonst ist die Lektüre illegal!
- **Nicht besser:**
 - elektronische Laborbücher
 - Virtuelle Forschungsumgebungen
 - Datenbanken
 - E-book-Reader
- **Nutzbarkeit der Datenformate wird viel zu oft übersehen**
 - Der Besitz der Daten im eigenen Speicher = Repositorium ist nur die halbe Wahrheit

22.10.2014

Datenschutz und Datenhoheit

12

Datenschutz = Datensicherheit



- Datenschutz missverstanden als Behinderung. Dabei:
 - Wer hat physikalischen Zugriff auf die Daten?
 - Welche Zugriffsrechte wurden vergeben?
- Offene Speicher in den Labors öffnen der Verfälschung bzw. Manipulation durch Dritte Tür und Tor
 - Und gefährden die wissenschaftliche Integrität
 - Plagiat nur untergeschoben?
 - Möglicherweise erlauben Laborgeräte (und ihre proprietären Formate) keinen Datenschutz ?!
- Die Sicherheitskette ist schon am Anfang unterbrochen ☹
- Systematischer Schutz ist leider mühsam
 - Administrativ = Arbeit = keine Ressourcen vorgesehen

22.10.2014

Datenschutz und Datenhoheit

13

Datenbesitz = Datenhoheit



- Wo liegen die Daten in der Wissenschaft?
 - In der Schublade?
- Prozessketten in der Uni halten einer Überprüfung nicht stand
 - Verantwortung liegt bei der Wissenschaft
 - Welche die technischen Probleme nicht kennt
- Forschungsdatenmanagement
 - Overload: nicht das Management von Daten über Forschung – sondern Handhabung von PB und EB
- Wo werden die Daten professionell gespeichert?
 - Prognose: nicht mehr in der eigenen Uni
 - Weder an der UB noch im RZ
- Spezialisierung!
 - RZs beginnen damit, die UBs sollten dies auch (jenseits der SSG)

22.10.2014

Datenschutz und Datenhoheit

14

Open ...



- Repositorien sind auch keine Lösung
 - Die Welt besteht aus Daten, nicht nur aus Publikationen
 - Dürfen wissenschaftliche Publikationen auf universitären Repositorien weltweit im Zugriff sein, wenn darin enthaltene Informationen Persönlichkeitsrechte verletzen?
- Open Data, Open Access: verschlimmern das Problem nur!
 - Allgemeiner Zugriff heißt nicht: unkontrollierte Bereitstellung, unkontrollierter Zugriff
 - Nötig: qualitätssichernde Prozeduren vor der Bereitstellung

22.10.2014

Datenschutz und Datenhoheit

15

Verstöße vs. Bedürfnis



- Dropbox ist halt notwendig
 - Warum? Haben Sie mal überprüft, wie einfach damit die Bedürfnisse befriedigt werden
 - Nach internationalen Arbeitsgruppen
 - Nach adhoc-Gruppen zum Dokumententausch?
 - Vorbereitung von Vorlesungsfolien
 - Funktioniert auch offline im Bahntunnel
- Angebote der RZ?
 - Feste administrative Gruppenrichtlinien
 - NFS-Server (SMB nur mit Gewalt)
 - Viel Spass im Tunnel!
- Mit nicht zeitgemäßen Angeboten treibt man die Nutzer den anderen zu
 - Wie schnell adaptieren sich RZ, UB, ... ???

22.10.2014

Datenschutz und Datenhoheit

16

Verstöße vs. Bedürfnis



- Wir versuchen gerade eine Regelung, in der die Verantwortung des Mitarbeiters abstrakt formuliert wird:
 - Wer kritische Daten aus dem Gebäude herausträgt, muss sorgfältig sein ...
 - Aber wer weiß, dass die email auf dem Handy in Teilen abgespeichert wird?
 - Und nicht nur angeschaut wird?
- Inwieweit dürfen Cloud-Lösungen überhaupt in einer wissenschaftlichen Umgebung eingesetzt werden?
 - seit Snowden ist das Bewusstsein zumindest verstärkt.
 - Verbot bringt nichts, denn manche Daten können problemlos in die Dropbox
 - Dropbox sammelt Bewegungsprofile? Anderes Thema...

22.10.2014

Datenschutz und Datenhoheit

17

Datenhoheit



- Wer besitzt die wissenschaftlichen Daten?
- Möglicherweise der Nutzer, der sie erfasst ???
 - Offenbar „seltsame“ Gerichtsurteile in jüngster Zeit
 - Also der Mitarbeiter, nicht der Prof oder die Uni?
 - Zumindest, wenn er/sie sie selbständig erfasst
 - Auswirkungen auf Forschungsprojekte!
 - Chef unterschreibt für das Projekt, hat aber hinterher nichts in der Hand ☹
 - Müssen Arbeitsverträge von Forschungsmitarbeiter/inne/n in Zukunft anders formuliert werden?

22.10.2014

Datenschutz und Datenhoheit

18

Datenhoheit



- BYOD
 - war letztes Jahr der große Hype – spart Investitionen
- Versuchen Sie, Ihre Daten von einem Mitarbeiter zurückzuerhalten, wenn die Daten auf dessen Privatgerät liegen!
 - Es gibt Streitfälle!
 - Vorgeschmack: Zugriff auf die dienstliche Mailbox, falls Privatnutzung nicht ausdrücklich ausgeschlossen
 - Das kann Firmen in die Insolvenz treiben
 - Viel Spass bei der Diskussion!
- Sicherheit von Privatgeräten
 - ...

22.10.2014

Datenschutz und Datenhoheit

19

Datenhoheit



- Sicherheit von Privatgeräten
 - IT hat keinen Zugriff, kann keine Vorschriften erlassen
 - Wenn die Privatgeräte mit dem internen Netz verbunden werden – was geschieht dann?
 - Viren, Trojaner, Sicherheitsüberwachung, ...
- Dual-SIM-Handy
 - Telefonieren ist nicht alles...
 - Wie schützt man den dienstlichen Bereich gegen den privaten Bereich?
 - Virtualisierung überfordert die Handys (noch)
- Einziger Ansatz: nur dienstliche Geräte zulassen
 - Bei Homeoffice wie beim mobilen Nutzer
 - Ist leider etwas unbequem und kostet €
 - 2 Handys, 2 Laptops,... Ausnahmen regeln

22.10.2014

Datenschutz und Datenhoheit

20

Achtung: Minenfeld

- Die Fragestellungen sind viel zu wichtig als dass man sie den IT-Mitarbeitern einfach so überlassen dürfte
 - Die sammeln zur eigenen Absicherung Daten und rücken sie ggfls auch heraus.
 - Die Verantwortung aber haben die „Chefs“
 - Und wenn die keine Ahnung von IT haben (wollen)?
 - Selbstdefiniertes Ahnungs-Umfeld reicht nicht!
 - Ich sage nur: Buchmesse 2014
- Lösungsstrategien tun manchmal weh
 - Und erfordern neues Denken
 - Vor 100 Jahren musste auch die Kavallerie umlernen

22.10.2014

Datenschutz und Datenhoheit

21

Ruf nach dem Staat?

- Die Suche nach einer optimalen Balance der unterschiedlichen Ansprüche ist ein andauernder Prozess.
- Nationale Eigenheiten werden gerne hervorgehoben, haben aber im internationalen Kontext keine Chance.
 - Das Internet ist grenzenlos.
 - Protektionismus ist auch keine Lösung
- Deshalb: wie in den guten alten Zeiten: selber bestimmen, wo es lang gehen soll.
 - Schneller sein, nicht warten.
- Natürlich muss man nationale Regeln beachten
 - Auftragsdatenverarbeitung, Dienstvereinbarung,...
 - Auch wenn dies zu komischen Auswüchsen führt

22.10.2014

Datenschutz und Datenhoheit

22

Lösungsmöglichkeiten

- Wir verletzen den Datenschutz und verlieren die Datenhoheit, weil es so bequem ist
 - Niemand will die damit zusammenhängenden Probleme verstehen
 - IT ist noch viel zu oft: sorgen Sie dafür, dass es funktioniert.
- Starre Regeln (*Industrie*) sind der Tod der Wissenschaft
 - Wir brauchen Freiheit – sonst geht die Forschung unter
 - Bestes Beispiel: Firma SUN
 - Wir müssen die Freiheit aber richtig nutzen
 - Nutzer damit oft überfordert / leichtsinnig / fahrlässig
- Angebote müssen stimmen
 - Was gibt die zentrale Einrichtung auf, um sich den neuen Anforderungen zu widmen?
 - Dropbox verbieten hilft nicht

22.10.2014

Datenschutz und Datenhoheit

23

Lösungsansätze in Ba-Wü

- Projekte in Baden-Württemberg
 - Wir versuchen es zumindest
- **bwIDM** als zentrales, landesweites Identitäts-Clearing
 - Begonnen als HPC-Zugang, Shibboleth
 - Inzwischen eine Basisfunktionalität
 - Jeder muss seine eigene Einrichtung im Griff haben
 - Gar nicht so einfach – workflows müssen bis tief in die Verwaltungsprozesse analysiert werden
 - Zusätzliche Köder (z.B. in Freiburg)
 - Eduroam funktioniert nur mit dieser Kennung
 - UB-Zugang erfordert diese Kennung

22.10.2014

Datenschutz und Datenhoheit

24

Lösungsansätze in Ba-Wü



- **bwSync&Share** als Serviceangebote der zentralen Einrichtungen
 - Dropbox kann man nur mit Alternativen „bekämpfen“
 - bwIDM als zwingende Grundlage
 - Dann aber ohne weitere Anträge ☺
 - Noch ungeklärt: wer liest das alles mit?
wie sammelt man wieder ein, wer Zugriff hat?
Wer wird informiert, wer alles Zugriff hat?
Naja, Dropbox kann das auch nicht
- **bwFDM:**
 - An allen Universitäten: welche Daten haben / produzieren / sammeln die Wissenschaftler/innen eigentlich; wie werden sie aufbewahrt?
 - Warum erst jetzt???
 - Dann: wie sehen die Konzepte eigentlich aus?

22.10.2014

Datenschutz und Datenhoheit

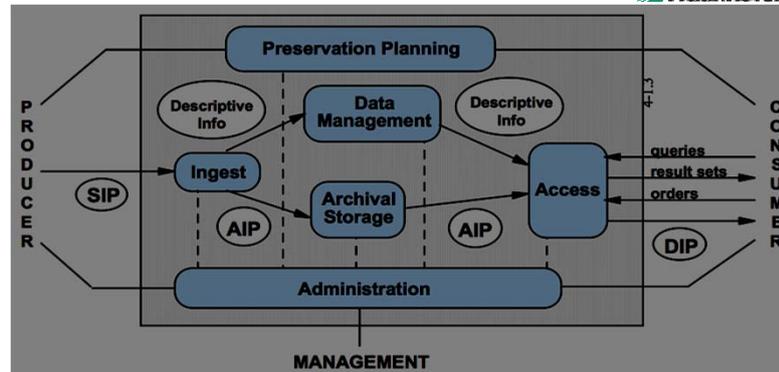
25

Wiedererlangung der Datenhoheit



- Projekt **bwFLA**
 - Ursachenanalyse spätestens seit bwFDM bekannt: An Universitäten wird mit Daten recht sorglos umgegangen
 - Nach erfolgreichem Experiment werden nur die zielführenden Daten aufbewahrt – und wie?
 - Andere müssen in anderem Kontext die Experimente wiederholen – teuer
 - Aufbewahrung nicht unbedingt professionell
 - In Schubladen auf DVDs und Festplatten
 - Auffindbarkeit eher selten gesichert
 - Aus Sicht des Experimentators ist das aber völlig in Ordnung
- Frage nach der wissenschaftlichen Integrität
 - DFG: 10 Jahre nach Projektende (=22 bei SFB)

Archiv-Standard-Modell OAIS



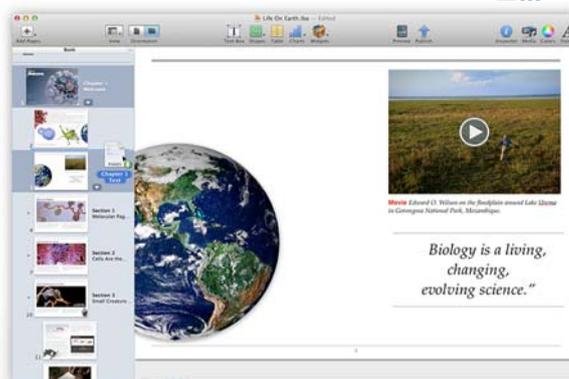
- Klar und einfach!?
–Bitstream ja, technischer Rest nicht unbedingt

Apple

- *With galleries, video, interactive diagrams, 3D objects, and more, these books bring content to life in ways the printed page never could.*

Ignorieren??

- Möglicherweise ist dies die Zukunft!
- Entweder man liest 10 Seiten „Bleiwüste“, um die Reproduktion der Zellen zu verstehen - oder man schaut sich ein kurzes (und gutes) Filmchen an.

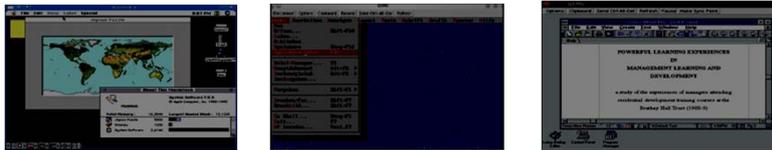


Wie archiviert man das?

Emulation von Originalumgebungen



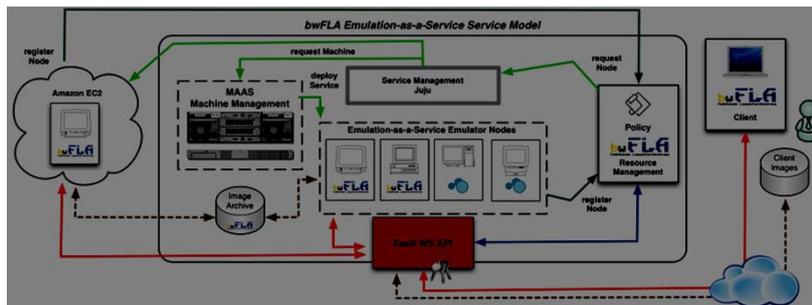
- Alternative zur Migration des Objekts – Nutzung seiner Originalumgebung
 - Annahme: Formate funktionieren am besten in ihrer Originalapplikation
 - Software funktioniert in Umgebung am besten, für die sie geschrieben wurde
- Deshalb: Wiederherstellung, Bewahrung von Originalumgebungen



EaaS - bwFLA Betriebsmodell



- Technische Komplexität von Emulation-as-a-Service
 - Zentral gemanagte Emulatoren und Workflows
 - Deployment der Emulatoren in der Cloud (Bspw. Amazon EC2)
 - Deployment in lokalen Clusters/GRIDS
 - Vorhalten von Ressourcen auf EaaS Seite notwendig



Virtuelle Forschungsumgebung II



- Projekt **bwLehrpool**:
- Ursprünglicher Grundgedanke: Wie können PC-Pools betrieben werden bei
 - rasch wechselnden Anforderungen
 - Arabisch, SPSS, Chinesisch, e-Learning
 - Sichere Umgebungen
 - E-Klausuren
 - Wenig Personal
 - D.h. Dozentensonderwünsche können nicht bedacht werden
 - Trennung von Betrieb und Nutzerwünschen
 - Verteilung der Verantwortlichkeiten auf mehrere RZs (inkl. FH)

Virtuelle Forschungsumgebung II

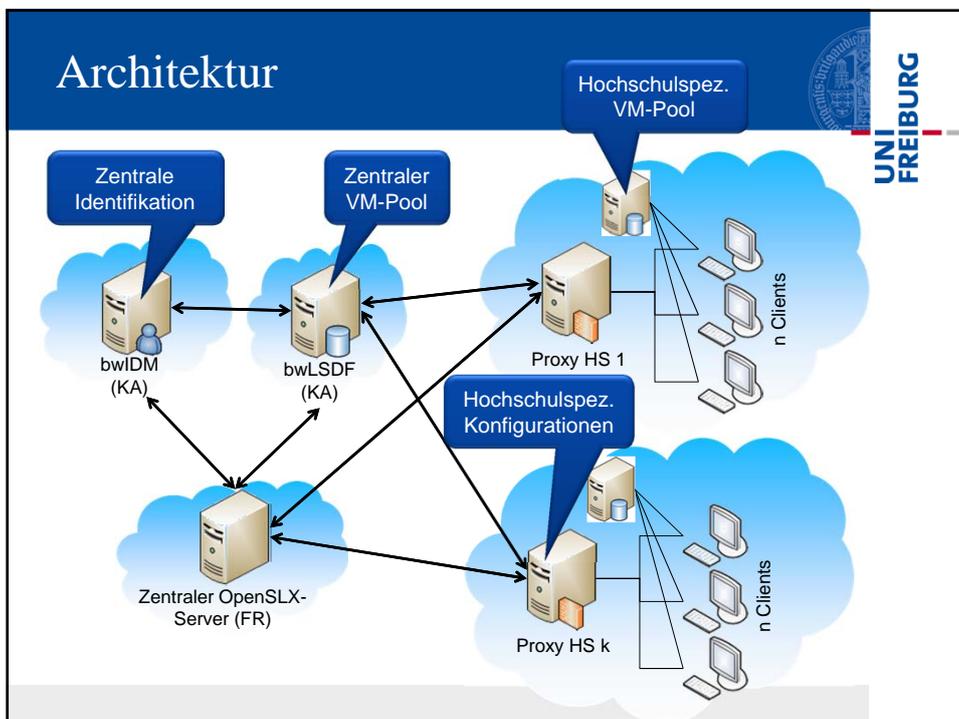


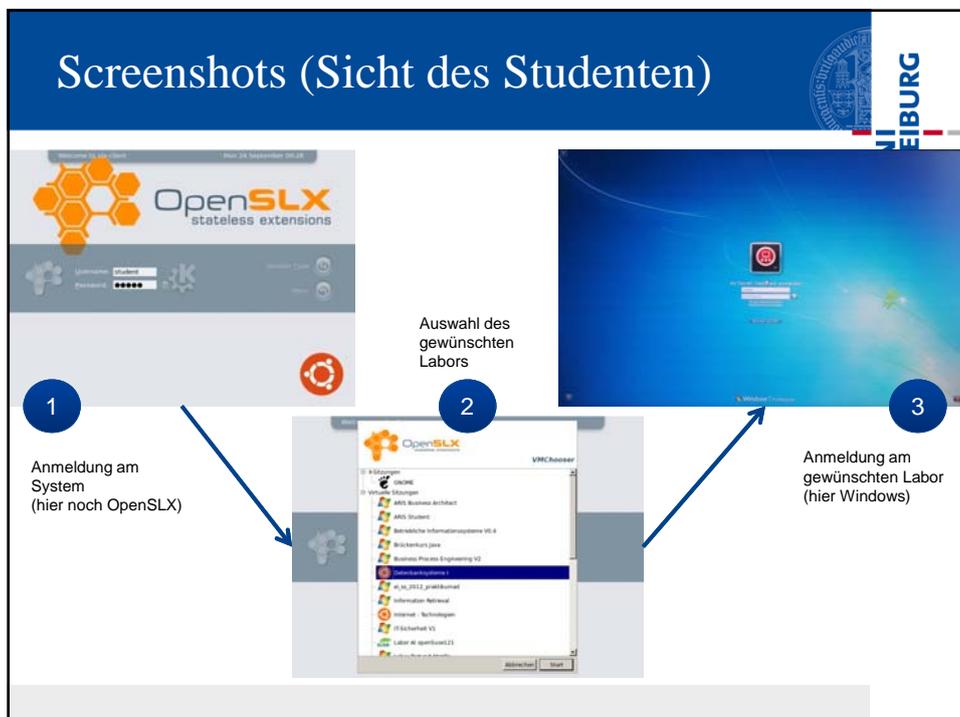
- Lösung:
 - Remote boot einer kleinen (Linux-)Umgebung
 - Nutzerumgebungen laufen in einer VM
 - Und werden vom Nutzer/Dozenten selbst gewartet
 - RZs liefern Standardimages
 - Start der Nutzerumgebungen zum Bedarfszeitpunkt
 - Heutige Hardware / Netze schnell genug
 - Hilfestellung:
 - Standardisierte Umgebungen durch Ausschreibungen: bw-PC, bw-Laptop
- Erkenntnis:
 - damit kann man auch wissenschaftliche Arbeitsplätze betreiben
 - Ohne Eingriff in die angeblich notwendige Freiheit (Parallelexistenz)
- Liefert archivierbare Virtuelle Maschinen
 - Was in bwFLA weiter behandelt werden kann

Vorteile



zentrale Infrastruktur	}	<ul style="list-style-type: none"> • Kostenreduzierung • Verringerung Administrationsaufwand • Konsolidierung
Raum- und Hardware-unabhängigkeit	}	<ul style="list-style-type: none"> • Flexibilität • Räumliche und zeitliche Unabhängigkeit
Abgesicherte Umgebungen	}	<ul style="list-style-type: none"> • e-Prüfungen • Komplexe Systemumgebungen • Eigenes Netzwerk
Kooperative Entwicklung von Lehrumgebungen	}	<ul style="list-style-type: none"> • Einbindung und Anpassung in Echtzeit; • Administrative Unabhängigkeit der Professoren und Lehrbeauftragten





Fazit

- Der Kreis schließt sich langsam
- Mit Serviceangeboten und Aufklärung kann man *teilweise* verhindern, dass unreflektiert den „Rattenfängern“ gefolgt wird
 - bwFDM, bwSync&Share, bwIDM
- Mit neuen Konzepten kann man *teilweise* verhindern, dass die Hoheit über die Daten verloren geht
 - bwFLA, bwLehrpool
- Mehr als der klassische Datenschutz:
 - Funktionalität, ohne dass die Daten nach außen sickern
- Rom ist auch nicht an einem Tag gebaut worden
 - Wie gesagt: wir versuchen es wenigstens

22.10.2014

Datenschutz und Datenhoheit

37

Fazit

- So hatten wir es uns damals nicht vorgestellt, als „wir“ das Internet erfunden haben
- Zum Glück hatte niemand daran gedacht, sonst wäre die **Weiterentwicklung / Evolution / Revolution** die das Internet ausgelöst hat, niemals geschehen
- *Kolumbus wusste auch nicht, was er anrichten würde, als er nach Indien segeln wollte.*

22.10.2014

Datenschutz und Datenhoheit

38