»Trust no-one!« — Die Komplexität digitaler Systeme und das Problem ihrer Vertrauenswürdigkeit

Bernd Kulawik[a]

[a] Selbständiger Architektur- und Musikhistoriker sowie DH-Entwickler, Bern, Schweiz, be_kul@me.com

KURZDARSTELLUNG: Die Integrität und Authentizität von Forschungs-ergebnissen (i.d.R. noch Texten, aber auch Datenbanken) und der ihnen zugrunde liegenden Software sowie Daten wird mit fortschreitender IT-Nutzung in den Wissenschaften immer dringlicher. Dies betrifft sowohl die Integrität aktueller, bereits hochkomplexer Systeme als auch die Sicherung dieser Integrität für ihre zukünftigen Nachfolger. Natürlich betrifft diese Anforderung sowohl die Soft- als auch die Hardware. Hinzu kommt in den Wissenschaften aber noch die notwen-dige Forderung, dass digital gespeicherte Informationen — und das betrifft nicht nur die »Rohdaten« — nicht nur irgendwie sicher archiviert werden, sondern der zukünftigen Forschung zur *interaktiven* Nutzung jederzeit zur Verfügung stehen sollten, *ohne* dass die Ausgangsdaten verändert werden. Nimmt man an, dass dieser Punkt durch einfaches Kopieren unveränderlich gespeicherter Daten ver-nachlässigt werden kann, so bleibt immer noch das Problem, dass es keine wirk-lich sicheren IT-Systeme gibt und ihre Sicherheit mit zunehmender Komplexität und Zeitdauer sinkt, da Problemstellen erst nach und nach entdeckt werden.

1. EINLEITUNG

Das Problem der Vertrauenswürdigkeit, also Integrität und Authentizität digitaler Daten, ist nicht nur für Banken oder staatliche Einrich-tungen von existenzieller Bedeutung, sondern auch für die Wissenschaften — und darunter sogar besonders für die historischen und Geisteswissenschaften: Um die grundlegende Forderung an jegliche Wissenschaft zu erfül-len, dass ihre Ergebnisse jederzeit überprüfbar und ggf. durch Wiederholung der Forschung auf der Basis derselben Daten *nach*prüfbar sein *müssen*, kann im digitalen Zeitalter *nur* erfüllt werden, wenn diese Daten vor unbefugter oder auch nur unbeabsichtigter Veränderung ge-schützt werden, jedoch *trotzdem* jederzeit reproduziert und ggf. weiter- bzw. nachgenutzt werden können.

Dies ist bekanntlich nicht der Fall: Das lehren nicht nur immer wieder auftretende Einbrüche in Datenbanken jeglicher Art, sondern auch die täglichen Meldungen über Fehler in Soft- *und* Hardware, die solche Einbrüche und Daten-diebstähle erst ermöglichen. Zwar mag man einwenden, dass ein Daten*diebstahl* für die ohnehin hoffentlich bald offenen Datenbestän-de der historischen und Geisteswissenschaften nicht allzu gravierend wäre, solange die Origi-naldaten noch erhalten sind. Aber es sollte ein-leuchten bzw. bekannt sein, dass in nahezu jedem Fall, in dem Daten *gestohlen* werden können, diese auch *verändert* oder ersetzt werden können.

Dabei ist die hohe Komplexität heutiger Systeme und die gängige Methode, sie durch *noch höhere* Komplexität zusätzlicher Software sicherer machen zu wollen, — vorsichtig ausgedrückt — nicht sehr hilfreich.

2. DATENINTEGRITÄT – WOZU?

Zuerst einige Bemerkungen zur Notwendigkeit von Datenintegrität und -authentizität: Natür-lich könnte ein außenstehender (Nicht-)Wis-senschaftler fragen, ob wir uns und unsere Daten und Ergebnisse nicht vielleicht etwas zu wichtig nehmen angesichts der Probleme, vor denen die Menschheit steht oder die sich nur schon in »ernsthaften« Umgebungen, wie eben bspw. in der Bank- oder Gesundheitswesen oder in der staatlichen Verwaltung ergeben. Es lässt sich leicht zugeben, dass die Probleme der historischen und Geistes- oder Bildwissen-schaften dagegen vergleichsweise unbedeutend erscheinen mögen.

Aber dagegen ließe sich bereits einwenden, dass bspw. die Integrität der Datenbanken in Museen, Archiven, Bibliotheken oder Privat-sammlungen meist einmalige Kulturgüter von unschätzbarem Wert betrifft. Die Notwendig-keit der Sicherheit und Integrität der Daten folgt also direkt aus derjenigen der von ihnen beschriebenen Artefakte: Ohne diese scheint eine Menschheit, die sich ihrer Geschichte und also ihrer Identität(en) bewusst sein und jeder-zeit versichern können will, schlicht nicht vor-stellbar. Stellt man sich vor, alle oder auch nur die bedeutendsten Artefakte in solchen Samm-lungen wären jederzeit gegen Kopien aus-tauschbar, weil sich anhand der Daten über sie nicht mehr sagen lässt, ob es sich um Originale handelt oder nicht, und wenn jeder alles Be-liebige über diese Objekte, ihre Geschichte und Bedeutung verbreiten könnte, ohne

dass man diese »Informationen« prüfen oder sie als *fake news* widerlegen könnte, dann dürfte ein Fortbestehen menschlicher Kultur, wie wir sie uns vorstellen, langfristig kaum möglich sein.
Unsere Wissenschaften tragen zu dieser kultu-rellen Selbstvergewisserung bei; man könnte darin sogar ihren vorrangigen Sinn und Zweck sehen. Deshalb erscheint es ebenso unverzicht-bar, nicht nur die Integrität der Datenbanken in den Sammlungen zu sichern, sondern auch diejenige der über die Objekte erhobenen For-schungsdaten und der daraus abgeleiteten wis-senschaftlichen Erkenntnisse.
Der bekannte Fall einer von ihrem Verfasser offensichtlich auch als umwälzend angesehe-nen wissenschaftlichen Arbeit über ein altes Buch mit Zeichnungen, das sich im Nachhin-ein als Fälschung erwies, ist hierfür sicherlich markant und nicht nur aufgrund der Namhaf-tigkeit des Autors und des beforschten histori-schen Wissenschaftlers ein signifikantes Bei-spiel, das als Warnung dienen sollte.
Ein zweiter, damit bereits angeschnittener, wichtiger Aspekt, der kaum etwas mit den Artefakten und den Informationen über sie selbst zu tun hat, sondern »nur« mit ihrer wissenschaftlichen Bearbeitung, ist das Prob-lem eben dieser wissenschaftlichen Arbeit und des Wissenschaftsbetriebs selbst: Wenn es nicht mehr eindeutig bestimm- und nachweis-bar bleibt, wer was wann worüber wie erkannt und publiziert hat, ist die Integrität des Wis-senschaftsbetriebs selbst insgesamt nicht nur in Frage gestellt, sondern hinfällig.
Vor diesem Hintergrund muss es geradezu erstaunen, wie *wenig* Aufmerksamkeit die historischen und Geisteswissenschaften bisher der Datensicherheit, -integrität und -authentizität widmen. Vielleicht ist das bisher noch vorherrschende Verlassen auf die Publikation der wichtigsten Ergebnisse wissenschaftlicher Ar-beit auf langfristig stabilem Papier in Büchern mit einer gewissen Auflage hier die — viel-leicht trügerische — Basis des Vertrauens, dass schon niemand in der Lage sein werde, ein Buch in größerer Zahl in alle relevanten Bibliotheken zu schmuggeln, um bspw. das Primat für die Erkenntnis zu beanspruchen, die *Mona Lisa* sei tatsächlich ein Bild Picassos? (Immerhin ist durch einen der größeren Kunstfälscherskandale der letzten Jahre bekannt, dass noch eine *riesige* Zahl gefälschter Bilder in unseren Museen und von der Forschung als Originale angesehen wird: Die Voraussetzung dafür ist gerade die scheinbare, aber in solchen Fällen ebenfalls gefälschte historische Daten-spur durch Archivalien, die über Bestands-nachweise und Provenienz Auskunft geben. Einen historischen Kaufvertrag oder ein Testa-ment zu fälschen ist relativ kompliziert; in Bezug auf Bits und Bytes ist diese Hürde jedoch für jemanden, der in IT-Systeme einzu-dringen vermag, relativ gering.
Es sollte also kaum zweifelhaft sein, dass die Integrität von IT-Systemen und der mit ihnen erhobenen und gespeicherten Daten sowohl in der Dokumentation als auch in der Forschung von *grundlegendem* Interesse ist. Wir sollten uns also fragen, wem wir vertrauen (können)…

3. EINIGE SICHERHEITSPROBLEME

Im Folgenden möchte ich anhand der verschiedenen »Ebenen« eines Systems kurz erläutern, wo überall gravierende Sicherheits-probleme und also Bedrohungen der Daten-integrität und -authentizität auftreten können, um anschließend zu überlegen, ob und wie man diesen Problemen zukünftig besser begeg-nen kann, als dies bisher der Fall ist.

3.1 HARDWARE

Computer gehören bekanntlich zu den komplexesten Maschinen, die Menschen bisher gebaut haben. Wohl niemand von uns wäre heute noch in der Lage, mit ein paar Bastel- und Elektronikkenntnissen einen einfachen Computer für ein aktuelles Betriebssystem selbst zu bauen. Außer denjenigen, die direkt mit dem Aufbau eines bestimmten Fabrikats vertraut sind, wäre wohl erst recht niemand mehr in der Lage, angesichts des »Wirrwarrs« an Bauelementen sagen zu können, welche da-von notwendiger Bestandteil und welche bspw. zu Spionagezwecken eingebaut wurden.

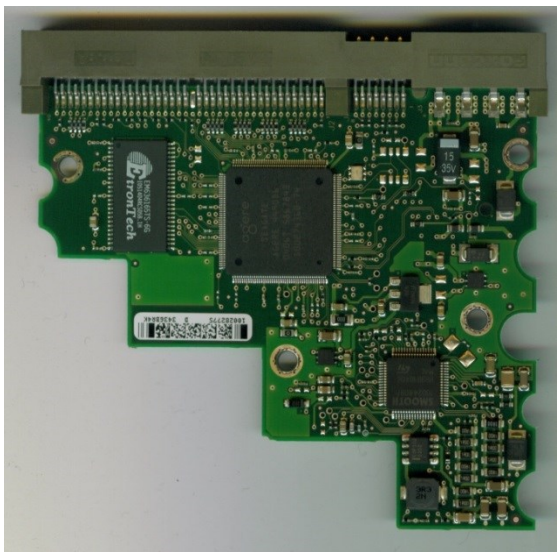

Abb. 1: Controller-Einheit für eine Festplatte (Dies ist also kein vollständiger Computer!)

Dass dies möglich ist, dürfte nicht erst seit den Verdächtigungen gegen den chinesischen Hersteller Huawei breiteren Kreisen bekannt sein. Wie ernst der Vorwurf ist, kann man daran erkennen, dass Huawei deshalb selbst angeboten hat, seine Hardware von unabhän-gigen Institutionen kontrollieren zu lassen, also seine

Geschäftsgeheimnisse zumindest teil-weise offen zu legen. Aber *dass* ein Verändern der Hardware zum Zweck des Ausspionierens der Nutzer auch heute schon tatsächlich geschieht, ist spätestens seit den Veröffent-lichungen Edward Snowdens bekannt.
Zwar führt der Trend zur Miniaturisierung und kompakten Herstellung zur Produktion soge-nannter *SoC*s = *System on a Chip*; d.h., alle notwendigen Teile des eigentlichen Rechners (ohne Bildschirm, Drucker und andere Peri-pherie) werden in einen einzelnen Chip integriert. Aber damit verschiebt sich das Problem nur: Es lässt sich also kaum aus-schließen, dass in einem heute oder in naher Zukunft verfügbaren System bereits auf der Hardware-Ebene Veränderungen vorgenom-men wurden oder werden können, die eine Kompromittierung des Systems erlauben.
Aber selbst so ein Eingriff ist nicht einmal nötig: Wie die vor wenigen Jahren bekannt gewordenen Probleme diverser Chiparchitek-turen zeigen (Stichworte *Spectre* und *Melt-down*), lassen sich scheinbar reguläre Funk-tionen in aktueller Hardware so ausnutzen, dass sie Angriffsmöglichkeiten eröffnen, mit denen entweder niemand zuvor gerechnet hat, rechnen konnte — oder rechnen wollte … oder die man einfach für vernachlässigbar hielt. Im Grunde ließen/lassen sie sich nur vermeiden, indem man auf wesentliche Funktionen ver-zichtet und damit massive Verluste an Rechen-geschwindigkeit hinnimmt. Grundsätzlich lässt sich daraus aber die Forderung ableiten, dass der *gesamte* Prozess vom Entwurf der Hard-ware bis zu ihrer Produktion *jederzeit* tran-sparent sein und von neutralen Spezialisten begutachtet werden müsste. Aber genau dies kann angesichts der Herstellerkonkurrenz und kriminellen oder Geheimdienstinteressen auf absehbare Zeit gar nicht der Fall sein.

3.2 SOFTWARE

Bei Software sieht es nicht nur ähnlich aus, sondern eher sogar noch viel schlechter: Denn während sich ein allzu massiver Eingriff in die Hardware vielleicht dadurch zeigt, dass diese nicht mehr (korrekt) funktioniert, gehört es bei Software-Angriffen auf IT-Systeme quasi »zum guten Ton«, dass diese *nicht* allzu leicht entdeckt werden können, die Angreifer also im Hintergrund nicht nur mitlesen oder Daten unbemerkt kopieren, sondern auch verändern können. Wie zahlreich die Möglichkeiten für solche Angriffe sind, lehren uns die (hoffent-lich) regelmäßigen Sicherheitsupdates der Software-Hersteller und die häufigen Mel-dungen über riesige katastrophale Datendieb-stähle aus IT-Systemen in Verwaltungen, Ban-ken, Sicherheitsorganen oder Krankenhäusern. Die inzwischen wohl häufigste Form einer solchen Kompromittierung von Daten dürfte aber die Verschlüsselung zu Erpressungszwek-ken sein (Stichwort *Ransomware*). Auch hier sind Angriffe natürlich wieder auf allen denkbaren Ebenen einer Software-Archi-tektur möglich: von der hardware-nahen Systemprogrammierung bspw. Betriebssystem-Kernen und Treibern, über *Middleware*, also sog. anwendungsneutrale Programme, bis hin zur einzelnen Anwendungssoftware (Daten-bank, Office-/Graphikprogramm, Eingabe- und Auswertungssoftware, Mailclients usw. usf.)
Wie hochkomplex und engstens miteinander verschränkt solche Software inzwischen ist, können die folgenden Schemata verdeutlichen:

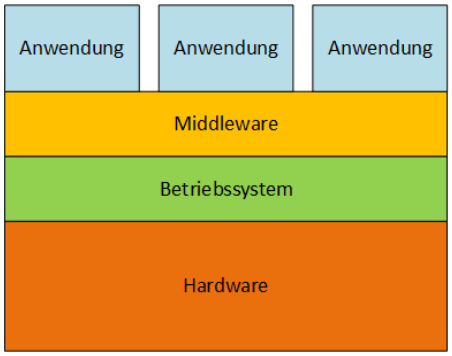

Abb. 2: Schema/Struktur eines IT-Systems

(Quelle: Wikipedia)

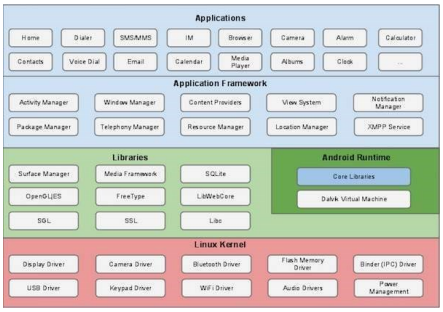

Abb. 3: Struktur des Android-Betriebssystems

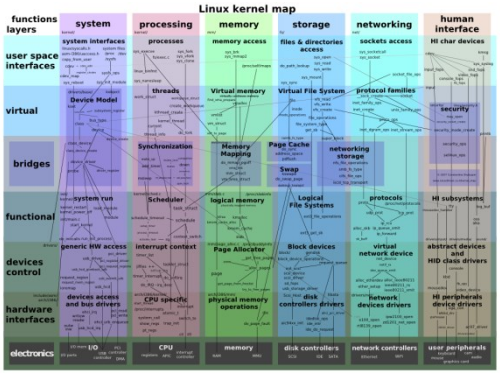

Abb. 4: Struktur des Linux-Kernels (= unter-ster Kasten in Abb. 3)

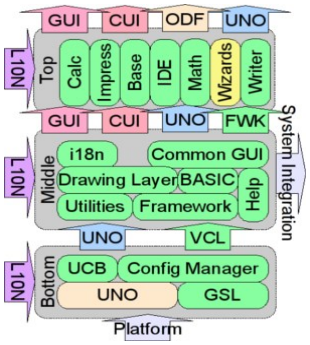

Abb. 5: Struktur einer Anwendung (hier: Open Office)

Ohne diese Abbildungen überhaupt genau lesen (können) zu müssen, dürfte deutlich wer-den, *wie komplex* heutige IT-Systeme bzgl. ihrer Software sind. Und da jeder Punkt, jede Verbindung in so einem Schema einen poten-tiellen Angriffspunkt darstellt: *wie* angreifbar diese Systeme sind, zumal sie vielfältigste Komponenten wie Teilprogramme und Pro-grammbibliotheken, Skripte und Routinen be-inhalten, die kaum alle überprüfbar sind.

Es braucht hier wohl nicht ausgeführt zu werden, dass vor diesem Hintergrund sog. *closed source* Software, also Binärcode, dessen menschenlesbarer Quellcode vom Benutzer nicht eingesehen oder gar kontrolliert werden kann, ein riesiges *schwarzes Loch* darstellt, in dem sich alles und jedes verbergen kann. Dies wird nicht nur deutlich durch die vielen Sicherheitsupdates, die Firmen wie Microsoft oder — besonders auffällig — Adobe bspw. regelmäßig verteilen *müssen* — und zwar nur für diejenigen Lücken, die zumeist durch andere trotzdem und i.d.R. unerwartet entdeckt worden sind. D.h., man muss davon ausgehen, dass sich in solcher Software noch *viel mehr* Fehler und damit Angriffsmöglichkeiten ver-bergen, als selbst den Herstellern bekannt ist. Und hinzu kommen vermutlich auch noch weitere Fehler, die bspw. bewusst eingebaut wurden, wie die *Hintertüren* in bestimmte viel-genutzte Software, die bspw. wieder kürzlich von US-Geheimdiensten gefordert wurden — natürlich *nur* für den »guten« Zweck und zu unser aller Sicherheit…

3.3 »WETWARE«

Unter Informatikern gibt es in verschiedensten Varianten den etwas sarkastischen Spruch, dass das Hauptproblem der gesamten IT zwischen Tastatur und Bildschirm liege, womit die Anwender gemeint sind. Denn selbst die sicherste Hard- und Software kann natürlich durch irrtümliche, falsche Bedienung beein-trächtigt werden. Kein Mensch — im Jargon gelegentlich abschätzig als »Wetware« bezeichnet — ist frei von Fehlern, und also auch nichts, was Menschen erdacht und hergestellt haben wie z.B. Hard- und Software.

Die Kompromittierung von IT-Systemen durch irrtümliche oder absichtliche Fehlbedienung seitens der menschlichen Benutzer zu ver-hindern, erscheint noch weniger möglich als die Vermeidung oder das vollständige Auf-spüren von Lücken und Angriffsmöglichkeiten in Hard- und Software. Fortwährende Schu-lungen könnten theoretisch zwar Abhilfe schaffen, sind aber mindestens genauso unbeliebt wie erfahrungsgemäß wirkungslos. Und Systeme, in denen Fehlbedienungen oder gar Angriffe schon auf der Ebene der *Konzep-tion* durch zusätzliche Sicherheitsmaßnahmen verhindert werden sollen, werden i.d.R. als Be-vormundung empfunden — und das nicht erst, wenn der Nutzer Administratorenrechte benö-tigt, um eine vergleichsweise kleine Verän-derung vorzunehmen…

4. ZEIT UND VERSCHLÜSSELUNG

Nach all dem scheint es nahezu aussichtslos, das Problem der Datenintegrität und -authen-tizität *überhaupt* dauerhaft lösen zu können. Und diese Unmöglichkeit potenziert sich noch, wenn man berücksichtigt, dass selbst ein heute absolut sicheres System aus Hard- und Soft-ware, bedient von vollkommenen Nutzern, die sich aller möglichen Irrtümer und Sicherheits-vorkehrungen bewusst sind, dass selbst ein solches System nicht »ewig« existieren *kann*. Im Gegenteil: Kaum eine Klasse von Systemen ist so schnellen und immer schneller auch: grundlegenden Veränderungen unterwor-fen, wie IT-Systeme: Da kaum ein System länger als 20 Jahre unverändert existiert und vor allem: *sicher* lauffähig ist und nur einzel-ne, sehr simple Datenformate wie TXT und PDF/A *vielleicht* bis zu 50 Jahren »haltbar« sind, kann man die *Halbwertzeit* solcher Systeme ruhig mit nur 10 Jahren angeben. D.h., von bspw. x über 3–10 Jahre geförderten Projekten, die vermutlich ihre eigene, an die spezifischen Bedürfnisse angepasste Software entwickelt und verwendet und mit dieser fleißig Daten sammelt haben, dürften 10 Jahre nach Abschluss des Projekts nur die Hälfte noch problemlos benutzbar sein… Die etab-lierte Unsitte, drittmittelfinanzierte Mitarbeiter nach Projektende zu entlassen, so dass auch das Wissen zum Weiterbetrieb und bzgl. mög-licher oder nötiger Anpassungen i.d.R. ver-loren geht bzw. eigentlich: »weggeworfen« wird, leistet ein Übriges dazu, solche Projekte nach Ende der Förderung in »lebende (Daten-) Leichen« zu verwandeln, für deren »Wieder-belebung« weder Geld oder neuere Hard- und Software noch eingearbeitete Mitarbeiter zur Verfügung stehen. Im Wortsinne gilt also, wovor Vinton Cerf bereits (bzw. eigentlich *erst*) 2015 warnte:

»We are nonchalantly throwing all of our data into what could become an information black hole without realising it. We digitise things because we

think we will preserve them, but what we don't understand is that unless we take other steps, those digital versions may not be any better, and may even be worse, than the artefacts we digitised.« [1]
Und selbst die geschätzte »Halbwertzeit« von 10 Jahren dürfte in vieler Hinsicht und bzgl. der hier im Fokus stehenden Datensicherheit und -integrität noch als viel zu optimistisch anzusehen sein: Denn es dürfte *kein* System geben, dass *ohne* Sicherheitsupdates auch nur über diese 10 Jahre als sicher oder auch nur lauffähig angesehen werden könnte. D.h. — und das ist IT-Anwendern gerade in Samm-lungen und Archiven sehr bewusst —, die *gesamten* Daten, i.d.R. inklusive der zu ihrer Nutzung notwendigen Software, *müssen regel-mäßig* auf neue Versionen derselben Software, irgendwann auf möglichst ähnliche Nachfolgesysteme und — *last but not least* — auch auf neue *Hardware* übertragen werden. (Dabei sei einmal außer Acht gelassen, dass kaum ein Projekt in dem Sinne als »abgeschlossen« angesehen werden kann, dass darin keine neuen Daten mehr aufgenommen werden müssten, die erst *nach* Projektende bekannt werden, dass also diese Projekte und ihre Systeme eigentlich nur dann sinnvoll sind, wenn auch solche Daten aufnehmen und sicher konservieren können. Insofern sind samm-lungszentrierte Datenbanken ohne Möglichkeit der Vernetzung digitaler Objekte über das Internet eigentlich obsolet…)
Nicht nur stellt sich das Problem der Sicherheit bzw. Kompromittierbarkeit damit für jedes System wieder neu, sondern der Übertragungs-schritt auf neue, zukünftige und also heute *per definitionem* noch unbekannte Hard- und Soft-ware selbst, der *nicht* nur in einem einfachen Kopiervorgang bestehen *kann* bzw. wird, potentiert die Angriffsmöglichkeiten.
Vertraut man erfahrenen IT-Spezialisten wie Vinton Cerf ([1] und [2]) oder Alan Kay [3], die seit Jahrzehnten wesentlich zur Konzeption und Gestaltung heutiger IT-Systeme beige-tragen *und* deren Entwicklung seither aufmerk-samst verfolgt haben, so wird deutlich, dass es zwar über kurz oder lang Emulationen ganzer Systemen inklusive Hard- *und* Software geben *muss*. Aber es dürfte auch klar sein, dass hier-bei nicht nur diese Systeme selbst, sondern auch die *Host*systeme, in denen diese Emula-tionen laufen sollen, immer wieder vor den beschriebenen Problemen stehen werden.

Und ähnliches gilt für die heute i.d.R. als Allheilmittel gegen Angriffe angesehene Ver-schlüsselung im weitesten Sinne: Denn diese fügt der bisherigen Komplexität weitere Ebe-nen und Verästelungen hinzu. Beispielsweise
wenn Hardware-Hersteller versuchen, durch *Trusted Computing*-Hardware unauthorisierte Zugriffe (oder auch nur die Installation alternativer, z.B. freier Betriebssysteme) zu verhindern;
wenn das Einloggen in Systeme nur mit sog. *Tokens* möglich ist, also z.B. USB-Sticks mit einem darauf instal-lierten System, die quasi als Schlüssel fungieren — auch noch in 100 Jahren? — ; oder
durch Verschlüsselung einzelner Be-reiche oder ganzer Festplatten mittels zusätzlicher Software, die dann nur mit Passwort/Token zugänglich sind.
In allen diesen und ähnlichen Fällen *erhöht* sich die Komplexität der zu kontrollierenden und vor Angriffen zu schützenden Hard- *und* Software um mindestens eine Größenordung bzw. sie potenziert sich sogar. Und natürlich müssten auch für diese Zusatz-Systeme wieder dauerhafte Lösungen entwickelt werden, damit die Benutzer der Zukunft noch Zugriff auf die Daten erhalten könn(t)en…

5. EIN LÖSUNGSVORSCHLAG

An dieser Stelle hatte ich in den letzten Jahren bereits — vor allem mit Bezug auf die lang-fristige Datenverfüg- und -nachnutzbarkeit — kurz skizziert, wie dieses m.E. bis heute unge-löste Problem angegangen werden könnte ([4], [5], [6]): Dabei bin ich zu der — bisher eigent-lich nicht erschütterten — Überzeugung ge-langt, dass der »Wildwuchs« der Hard- wie Software-Entwicklung der letzten Jahrzehnte eigentlich viel zu weit fortgeschritten ist, als dass man dort hinein noch eine halbwegs zu-kunftsfähige oder sogar zukunfts*sichere* »Schneise schlagen« oder — auf der Basis des Existierenden — irgendeine sichere Struktur schaffen könnte, um das Problem wirklich zu lösen. Hier seien im Folgenden einige Punkte genannt, die mir aufgrund meines zweifellos beschränkten Horizonts als notwendige Be-standteile eines Lösungsansatzes erscheinen.

5.1 VOLLSTÄNDIGER »REBOOT« (?)

Da wir zwar wissen, dass und manchmal sogar wie und wo die aktuell existierenden Hard- und Softwaresysteme gravierende Mängel haben, um die herum bisher oft eher *work-arounds* geschaffen wurden, weil ihre *grund-sätzliche* Behebung unabsehbare Konsequen-zen für das jeweilige Gesamtsystem hätte (Stichwort: *A20-gate*), scheint es mir unver-meidlich, unter Berücksichtigung dieses Wis-sens *und* des heutigen Wissensstandes über die bereits absehbaren Entwicklungen in der Zukunft einen vollständigen »Neustart« zu wagen: Dabei ist »Neustart« oder »Reboot« eigentlich eine irreführende Bezeichnung, denn sie suggeriert, dass man vielleicht noch auf der Basis *existierender* Systeme nach einigen »Tuning-Maßnahmen« und eben dem Neustart desselben Systems die

skizzierten Probleme beheben könnte. Ich bin überzeugt, dass dies jedoch *nicht* der Fall ist und sein kann: Nicht nur weben »Betriebsgeheimnissen« bzgl. Hardware oder Software-Quellcode auf allen vorstellbaren Ebenen »schleppen« wir — z.T. seit Jahrzehnten — viel zu viele Fehlermög-lichkeiten mit. Auch grundlegende Konzepte, auf denen nahezu alle heute existierenden Systeme beruhen, dürften einer kritischen Überprüfung aus *heutiger* Sicht kaum noch standhalten, da sie aus einer Zeit stammen, in der bspw. Systeme noch so teuer waren, dass man allein schon aus *Kosten*gründen *work-arounds* hinnahm oder einbaute, die sich aus *Sicherheits*gründen eigentlich verbieten…
D.h., eigentlich müsste man vermutlich sowohl in der Hard- als auch in der Software-Entwick-lung noch einmal *from scratch* beginnen…

5.1.1 OFFENE HARD- UND SOFTWARE

Die grundsätzliche Forderung an den gesamten Prozess muss seine absolute Offenheit, Trans-parenz und Freiheit sein: Damit ist gemeint, dass schon die Diskussion über das Design aller Systemkomponenten offen, in der Öffent-lichkeit stattfinden und frei von »Hinterzim-mer«-Verhandlungen sein *muss*. Jeder Schritt, der dabei nicht von potentiell allen Beteiligten und Betroffenen öffentlich und nach wissen-schaftlichen Maßstäben nachvollziehbar wäre, würde das *gesamte* zu schaffende neue System bereits von vornherein kompromittieren.
Natürlich würden sich kommerzielle Herstel-ler, Patentbefürworter oder auch »interessierte Kreise« wie bspw. Überwachungsbefürworter *vehement* dagegen wehren und den Untergang der freien Wirtschaft, des Wohlstands und überhaupt der Welt heraufbeschwören…, aber ich denke, man sollte ihnen keinen Glauben schenken: Die wichtigsten, folgenreichsten und schnellsten Verbesserungen (nicht nur) in der Technik und Kultur entstanden immer dort, wo grundlegende Verfahren und Kenntnisse frei zugänglich waren und nicht Einzelne im »stil-len Kämmerlein« an geheimnisumwitterten Lösungen arbeiteten, die sie dann ggf. paten-tieren konnten, um andere von parallelen Ent-wicklungen abzuhalten. Sondern sie entstanden dort, wo neue Erkenntnisse offen geteilt und bspw. ohne kostspielige Lizenzen oder dro-hende Einflussnahmen der »Erfinder« ent-stehen und sich verbreiten konnten: Hätten die »Väter des Internet«, Vint Cerf und Rob Kahn oder Alan Kay und seine Mitarbeiter anders gehandelt, wäre das Internet heute vermutlich nicht existent… und XEROX, in dessen *Palo Alto Research Center* (PARC) die graphischen Benutzeroberflächen entwickelt wurde, wäre heute die größte IT-Firma der Welt, nicht Microsoft oder Apple, die sich umstandslos der Konzepte aus dem XEROX PARC bedienten.
Entsprechend und wie zur Bestätigung dieser These bzw. Forderung stammt das mit großem Abstand am häufigsten heute auf IT-Systemen eingesetzte Betriebssystem *eben nicht* aus Red-mond oder Cupertino, sondern heißt Linux und liegt in einer Vielzahl spezifisch angepasster Varianten vor, deren bekannteste sicherlich Googles Handy-Betriebssystem *Android* ist. Linux stammt aber eben von einer Vielzahl freier und zunehmend auch angestellter Ent-wickler, die ihre Arbeitsergebnisse gemäß der GNU GPL allen anderen Nutzern wiederum zur Verfügung stellen (müssen). NUR dieses Merkmal hat dazu geführt, dass sich Linux weiter verbreiten und schneller entwickeln konnte, als einzelne Firmen es je hätte leisten können. Nicht nur Handheld-Computer laufen überwiegend damit, sondern auch alle Super-computer in den obersten Rängen der TOP500-Liste. *Kein* anderes System verfügt über diese Bandbreite an Einsatzmöglichkeiten!
Die Offenheit und Transparenz sowie Freiheit bzgl. der Wiederverwendung und Weiterent-wicklung ist aber unter dem hier interessieren-den Blickwinkel weniger wegen der Entwick-lungsgeschwindigkeit und -freiheit wichtig, sondern vor allem wegen der Vertrauenswür-digkeit der Ergebnisse: *Closed Source* Soft-ware ebenso wie Hardware *kann* eben *per se nicht vertrauenswürdig sein*, denn das Vertrau-en muss bzw. müsste *nur* auf der Versicherung der Hersteller beruhen, alles schon irgendwie *richtig* gemacht, keine Fehler (un)absichtlich eingebaut zu haben und die Macht über sein Produkt niemals missbrauchen zu wollen.

5.1.2 HARDWARE

Beim Design der Hardware sollte nicht nur von vornherein darauf geachtet werden, bekannte konzeptionelle Fehler zu vermeiden, sondern bspw. auch darauf, größtmögliche Energie-effizienz zu erreichen: Der rasante Fortschritt der Digitalisierung auf allen Gebieten sorgt bereits heute dafür, dass sie einen Großteil der bereitstellbaren Energie verbraucht. Und mit dem Aufschließen der sog. *Dritten Welt* wird sich dieser Energieverbrauch zweifellos ver-vielfachen. Bzgl. eines sparsamen Umgangs mit Ressourcen (Stichwort »seltene Erden«) müsste außerdem eine weitgehende Modulari-tät verlangt werden, die es ermöglicht, einzelne Komponenten auszutauschen *ohne* gesamte Systeme zu Elektroschrott zu machen, dessen Recycling — so es denn überhaupt stattfindet — auf die Müllberge von Nigeria oder Indonesien »outgesourcet« wird, wo er die Gesundheit der Menschen massivst gefährdet.

5.1.3 SOFTWARE

Auch die Software sollte nicht nur modular, sondern auch energiesparend konzipiert werden: Seit mindestens zwei Jahrzehnten wird bspw. die *Wintel*-Allianz beklagt, also die unheilige Allianz von *Microsoft Windows* und *Intel*, die regelmäßig dazu führt, dass alle Hardware-Fortschritte, die sich in einer Ver-vielfachung der Rechengeschwindigkeiten und einem niedrigeren Stromverbrauch nieder-schlagen sollten, von den Anforderungen der nächsten Software-Generation wieder »aufge-fressen« werden: Während sich die Rechen-geschwindigkeiten und Speichergrößen ver-tausendfacht haben, ist zwar auch die Zahl der (meist kaum benötigten) Optionen der Soft-ware gewachsen, aber *nicht* bzw. kaum deren Arbeitsgeschwindigkeit.
Auch vermag heute wohl kaum noch jemand — mit Ausnahme einiger forensischer Spezia-listen — *überhaupt* zu sagen, welche Daten sein Betriebssystem, sein Browser, sein Office-Paket oder die Mail- und Kalendersoftware an den Hersteller und seine »Industriepartner« (oder Geheimdienste: Stichwort *NSA-key*) wei-terleitet… Datensicherheit und -integrität wären das Gegenteil dessen…

5.2 VERTRAUEN

Dass »Vertrauen die Grundlage von allem« sei, behauptet nicht nur die Werbung eines selbst nicht sehr vertrauenswürdigen Bankhauses, sondern folgt in diesem Zusammenhang *zwin-gend* aus den Anforderungen zur Dateninte-grität, -authentizität und -sicherheit. Vertrauen kann aber gerade *nicht* auf Zusicherungen oder »Ehrenworten« beruhen, sondern *nur* durch einen jederzeit und durch jeden kontrollier-baren Prozess hergestellt und erhalten werden.
Gegenwärtig funktioniert nicht nur die Produktion von IT-Systemen jedoch *nicht* nach diesem Grundsatz, sondern auch die wissen-schaftlichen Wissens. Auch dies wäre im Inter-esse einer zukünftig als sicher(er) anzusehen-den wissenschaftlichen IT-Nutzung zu ändern.

5.2.2 NEUE FORM DES *PEER REVIEW*

Die historisch gewachsene Form des (im Idealfall: *double-blind*) *Peer Review* geht bekanntlich darauf zurück, dass Einreichungen bei wissenschaftlichen Zeitschriften möglichst von einigen Spezialisten derselben Fachrich-tung und ohne Voreingenommenheit beurteilt werden sollen. Nun ist dieses Verfahren nicht erst seit dem Bekanntwerden von *Zitierseil-schaften* zumindest fragwürdig. Allein die rasant fortschreitende Spezialisierung führt — selbst in den historischen und Geisteswissen-schaften dazu —, dass die vorausgesetzte Anonymität von Reviewer-Seite leicht zu durchbrechen ist, da *man sich kennt* und des-halb relativ gut abschätzen kann, wer ein ein-gereichtes Paper verfasst haben dürfte.
Hinzu kommt noch ein anderer Schwachpunkt: Die bewusst herbeigeführte Knappheit an For-schungsgeldern führt zu einem Kampf um Drittmittel, in dem es für die meisten Beteilig-ten buchstäblich um die Existenz geht. Und selbst die wenigen »Auserwählten«, die sich im Prinzip auf unbefristeten Stellen einer gewissen Absicherung erfreuen dürfen, stehen aufgrund der Durchökonomisierung und der damit einhergehenden Forderung nach *Mess-und Bewertbarkeit* wissenschaftlicher For-schung vor dem Problem, im Kampf um Dritt-mittel zum Erfolg verdammt zu sein. Dass daraus kein gesundes, dem *gemeinsamen* Wis-sensfortschritt förderliches Klima ensteht, ist nicht erst seit dem »Auffliegen« diverser massiver Betrugsfälle bekannt und wird auch nicht erst seitdem beklagt.
Bzgl. der Datenintegrität sind diese Entwick-lungen als ebenso katastrophal wie diejenigen der IT einzuschätzen, weshalb ein Umdenken angebracht erscheint. Glücklicherweise spricht heute *prinzipiell nichts* — außer dem Macht-verlust interessierter Kreise — mehr dagegen, die Begutachtungsprozesse vollständig offen zu gestalten: Wenn jeder mit seinem Klarna-men für die Bewertung der Arbeit eines an-deren einstehen muss, ist das Ende der Seil-schaften erreicht. Und für die »Bewertung« eines Wissenschaftlers wären auch nicht mehr nur seine Texte der einzige Maßstab, sondern ebenso seine (Fehl-) Urteile über andere…

5.2.3 NEUER WISSENSCHAFTSBETRIEB

Letzlich könnte dies zu einem vollkommen neuen Wissenschaftsbetrieb oder -modell führen, in dem sich *wirklich* Qualität durchsetzen könnte. In dem aber vor allem dank *Open Access* und *Open Data* Ergebnisse jederzeit nachprüfbar wären. Unterstützt würde dies durch die freie Verfügbarkeit offener IT-Systeme, deren Benutzung bspw. von Förder-institutionen verpflichtend eingefordert werden könnte. Momentan verlangen diese stattdessen vom einzelnen Antragsteller, sich mit den un-gelösten und für Nicht-IT-Spezialisten unlös-baren Problemen des Forschungsdatenmanage-ments nicht nur zu befassen, sondern auch Lösungen selbst vorzuschlagen bzw. sogar zu entwickeln, welche die Verfügbarkeit, *Sicher-heit und Integrität* der erhobenen Daten und erarbeiteten Ergebnisse für die Zukunft sicher stellen sollen. Selbst für einen Zeitraum von 10–15 Jahren ist dies (wie oben erläutert) eigentlich gar nicht realisierbar, die Forderung also unrealistisch und unfair! Deshalb wären diese Institutionen m.E. in der Pflicht, die

Mit-tel für die Erfüllung ihrer Forderungen *selbst* bereit zu stellen.
Ich bin überzeugt, dass ein *offener* Umgang mit Forschungsdaten und -ergebnissen lang-fristig dazu führen würde, dass das aus dem allgegenwärtigen Kampf aller gegen aller um Drittmittel und »Meriten« ein *Miteinander* werden könnte, das erst als solches ernstzu-nehmender Wissenschaft würdig wäre.

6. UMSETZUNG

Das alles mag dem einen oder der anderen als viel zu utopische »Zukunftsmusik« erscheinen, deren Realisierung nicht zuletzt durch unser Wirtschafts- und das davon leider viel zu sehr abhängige politische System verhindert werde. Da sich aber langsam nicht nur bzgl. ökolo-gischer Fragen die Einsicht verbreitet, dass es »*so* nicht mehr weiter gehen kann«, ist es vielleicht auch denkbar, eine IT-Umgebung zu entwickeln, die dem oben Beschriebenem nahe kommen und so — neben anderen Problemen — eben auch das der Vertrauenswürdigkeit der Forschungsdaten und -ergebnisse gewährlei-sten könnte. M.E. wäre es dazu notwendig, eine dauerhafte, *internationale* Institution zu schaffen, die den Konzeptions- und Entwick-lungsprozess der Hard- und Software koordi-niert und kontinuierlich lenkt. Wenn dann ab-sehbar wäre, dass ein solches System zukünftig nicht nur in allen mit Steuermitteln geförderten Forschungseinrichtungen Standard wird, dürfte sich auch die (ja ebenfalls aus Steuermitteln finanzierte) universitäre Ausbildung vielleicht dahingehend orientieren, forschend und ent-wickelnd an diesem Prozess teilzunehmen.
Diese Institution sollte dann die Herstellung, Betreuung und Weiterentwicklung insbeson-dere der Software langfristig absichern (kön-nen), während die Herstellung der Hardware bspw. an lizensierte Auftragnehmer delegiert werden könnte, was den sicherlich nicht un-willkommenden Nebeneffekt haben dürfte, dass die Konkurrenz tatsächlich das Geschäft beleben und monopolistische Fantasiepreise verhindern würde.
Diese Institution würde das Software-System dann frei für jeden zur Verfügung stellen, der es nutzen möchte; für aus Steuermitteln geför-derte Forschung wäre seine Verwendung sogar verpflichtend. Damit könnte zugleich gesichert werden, dass die so erhobenen Daten und die Forschungsergebnisse nach Abschluss eines Projekts an diese Institution zur weiteren Auf-bewahrung zurück übergeben werden und sie so dort allen zukünftigen Nutzern zur Ver-fügung stehen könnten.
Diese Institution wäre natürlich verpflichtet, darauf zu achten, dass das von ihr betreute System stabil und *sehr vorsichtig* so weiter-entwickelt wird, dass einmal erhobene Daten soweit irgendwie absehbar verfüg- und nutzbar bleiben. D.h., Anpassungen der Software an bestimmte spezifische Anforderungen eines einzelnen Projekts dürften nur nach Rück-sprache mit dieser Institution und mit deren Einwilligung vorgenommen werden.
Das Ganze mag viel zu »zentralistisch« klin-gen, als den meisten lieb ist, vielleicht wird darin sogar eine Bedrohung der Freiheit der Wissenschaft gesehen — eine Behauptung, die sich bereits im Kampf um die Durchsetzung des *Open Access* als ideologische Propaganda erwiesen hat… Aber angesichts der Alter-native, unsere Forschungsdaten und -ergebnis-se in absehbarer Zeit *voll-ständig* zu verlieren und eben ihre Vertrauenswürdigkeit und damit *in the long run* ihre Wissenschaftlichkeit selbst nicht mehr garantieren zu können, erscheint mir ein — das sei noch einmal betont — auf allen Ebenen und in jeder Phase *offener, transparenter und freier* Prozess nicht nur als das »geringere Übel«, sondern sogar als die einzig denkbare und vernünftige Lösung.
Und die Kosten? Sicherlich dürften diese zu Beginn im höheren dreistelligen Millionen-bereich liegen, insbesondere, wenn man die besten IT-Spezialisten einbinden und angemes-sen bezahlen wollte. Angesichts dessen aber, was auf dem Spiel steht und was andererseits an Mitteln für Kriege, fragwürdige Infrastruk-turprojekte oder gar die Rettung von Banken (bzw. deren Aktionären) vor der Pleite auf-gewandt wurde und aufgewendet wird, wären nahezu alle denkbaren Beträge jedoch die sprichwörtlichen *Peanuts*. Und möchte wirk-lich jemand in Frage stellen, dass die Bewah-rung und zukünftige Sicherung wissenschaft-licher Forschung uns *mehr* wert sein sollte als ein Kampfflugzeug, das (zum Glück) nicht oder nur bei Schönwetter fliegt, ein Flughafen oder Bahnhof, der minimalen Sicherheitsvor-gaben nicht entspricht und deshalb vermut-lich/hoffentlich nie in Betrieb gehen darf, oder das Wohlbefinden von Aktionären, die sich bei ihren Wetten auf Kurse und »Wertpapiere« wissentlich verspekuliert haben?
Die Mittel sind also da, das Know-how ist da, der Bedarf ist da und das Interesse, ihn zu erfüllen, ebenso: Worauf also warten wir noch? Wem wollen wir sonst weiter vertrauen? Uns selbst oder dem »freien Markt«?

7. ABBILDUNGSNACHWEISE

Abb. 1: ComputerService Wöhler c-s-woeh-ler.de/produkt/elektronik-platine-festplatte-st340014a-seagate/?v=3a52f3c22ed6
Abb. 2: commons.wikimedia.org/w/index.-php?curid=36493238
Abb. 3: www.tutorialspoint.com/android/an-droid_architecture.htm

Abb. 4: commons.wikimedia.org/wiki/File:Linux_kernel_map.svg
Abb. 5: wiki.openoffice.org/wiki/Architecture

8. LITERATURHINWEISE

https://www.theguardian.com/technology/2015/feb/13/google-boss-warns-forgotten-century-email-photos-vint-cerf

Vint Cerf auf der 25. Jahrestagung des W3C über das *digital vellum*: https://vimeo.com/110794988 und ausführlicher: www.youtube.com/watch?v=STeLOogWqWk

Nguyen, Long Tien; Kay, Alan: The Cuneiform Tablets of 2015. Viewpoints Research Institute, VPRI Technical Report TR-2015-004, Los Angeles: 2015, online www.vpri.org/pdf/tr2015004_cuneiform.pdf]

Kulawik, Bernd: Digitales Kuratieren – und dann? – In: Konferenzband EVA-Berlin 2016, S. 75–82.

Kulawik, Bernd: Wie man das Verschwinden unserer Daten im »digitalen Schwarzen Loch« und ein »Dunkles Informationszeitalter« verhindern könnte. – In. Konferenzband EVA-Berlin 2017, S. 220–227.

Kulawik, Bernd: Digitale Zwillinge sollten sich nicht zu sehr ähneln und «getrennt wohnen». – In: Konferenzband EVA-Berlin 2018, S. 101–105.