KRITERIEN ZUR BEURTEILUNG DER BEWEISKRAFT VON DATEN

Michael Sonntag

Abstract: In Gerichts- und Verwaltungsverfahren nehmen Daten als Beweismittel einen immer größeren Raum ein, da Private wie auch Firmen ohne IT-Ausrüstung und elektronische Kommunikation kaum noch auskommen bzw erfolgreich sein können. An die Stelle handschriftlicher Aufzeichnugen oder mit Schreibmaschine getippter Briefe treten el. Dokumente, E-Mails und Chat-Nachrichten. Da diese alle leicht änderbar sind, stellt sich bei Streitigkeiten die Frage: Sind sie echt oder manipuliert? Stammen sie vom angegebenen Absender/Autor? Stimmen sie mit der Realität überein? Dieser Beitrag arbeitet Kriterien heraus, nach denen der Wert solcher Beweise beurteilt bzw bewertet werden kann.

INHALTSVERZEICHNIS

A. EINLEITUNG

Elektronische Daten als Beweise werden immer wichtiger, da sie auch im normalen Geschäftsleben und bei Privatpersonen einen immer größeren Raum einnehmen. Darüber hinaus sind viele moderne Geräte weniger Waschmaschinen oder Fernseher, als vielmehr «Computer mit für diese ungewöhnlichen Zusatzfähigkeiten». Bei vielen Gerichts- oder Verwaltungsverfahren ist es daher notwendig, deren «Inhalte», dh Daten als Beweise einzubrigen. Da Daten selbst aber praktisch nie sichtbar (oder sonst mit menschlichen Sinnen direkt wahrnehmbar) sind, muss auf deren Ausdrucke, eine auf einem Computer einzusehende Kopie, Auswertungen von Experten etc zurückgegriffen werden. Es ist daher erforderlich festzustellen, wie der Beweiswert von Daten einzuschätzen ist: Sind sie wenig wert, da trivial zu fälschen, oder sehr hochwertig, da dies

fast unmöglich ist bzw mit sehr hoher Wahrscheinlichkeit nicht erfolgte? Leider wird dies in jüngster Zeit oft ausschließlich unter dem Gesichtspunkt elektronischer Signaturen (→ eIDAS-VO[1]) diskutiert[2]. Diese sind zwar sehr hilfreich – aber nur sofern vorhanden. In den meisten vor Gericht landenden Fällen fehlen sie jedoch, denn wie viele (insb kleine und mittlere) Unternehmen sichern alle irgendwo anfallenden Daten (Logdateien der Firewall, Office-Dateien der Korrespondenz, Buchhaltungs-Datenbank...) schon mit qualifizierten Signaturen?

Als weiteres Beispiel der Problematik des Beweiswertes el. Daten können die so genannten «Deep Fakes» dienen, Fotos oder Videos, in welchen die Gesichter durch die anderer Personen ersetzt wurden. Durch Menschen ist eine optische Erkennung der Manipulation nicht möglich, und ein Bild/Video kann problemlos unter einem passenden Dateinamen auf einem fabriksneuen USB-Stick gespeichert werden. Damit liegt ein unverändertes Original vor, dem natürlich vollauf zu vertrauen ist – oder doch nicht? So könnte man zB die Frage stellen, mit welcher Kamera die Aufnahme erfolgte (→ Dead Pixel), wie die Metadaten aussehen, ob ein fast identisches Bild mit anderem Gesicht bekannt oder aufzufinden ist, von wem der Memory-Stick stammt etc. Es ist daher wichtig zu untersuchen, welche Kriterien geeignet sein können, den Beweiswert von Daten zu erhöhen oder zu reduzieren, sodass man besser beurteilen kann, ob vorliegende Daten echt oder manipuliert sind.

B. PROBLEME VON DATEN ALS BEWEIS

Bei Daten stellen sich eine Reihe von Problemen bei ihrer Würdigung, die sich in anderen Bereichen nicht oder nur in geringerem Ausmaß stellen.

---

1 Verordnung (EU) Nr. 910/2014 des Europäischen Parlaments und des Rates vom 23. Juli 2014 über elektronische Identifizierung und Vertrauensdienste für elektronische Transaktionen im Binnenmarkt und zur Aufhebung der Richtlinie 1999/93/EG, ABl L 257 vom 28.8.2014, 73.

2 Als Beispiel siehe etwa JANDT/MICHAELK/DIETRICH, Wie hoch ist der (Beweis-)Wert digitaler Dokumente? DuD 10/2015, 687.

I. Wie genau sind die Daten?

Computer arbeiten im Binärcode, kennen daher nur 0 oder 1. Die Genauigkeit sollte daher unübertroffen sein – doch dies ist eine Täuschung. Wird beispielsweise eine Temperatur gemessen und gespeichert, so kann dies sehr genau erfolgen – oder auch nicht. Unter der Annahme, dass der Sensor auf 0,2 °C genau misst, wird zB der Wert 21,534088 °C gespeichert, da sich dies aus der Korrekturformel des Sensors ergibt (Roh-Messwert angepasst an dessen Nichtlinearität. Dies ist für einen Computer trivial und kein Problem. Es können daher problemlos Auswertungen auf Millionstel Grad durchgeführt werden, welche der Computer bis auf die letzte Stelle genau ausrechnet. Selbst wenn alle Formeln absolut korrekt sind, wird hiermit jedoch Genauigkeit nur vorgetäuscht: Was Anfangs nicht vorhanden war (Sensor: 0,2 °C !), kann nachher nicht mehr besser werden. Es ist daher wichtig, die Ursprungs-Genauigkeit (im weitesten Sinne, zB auch, welche Daten überhaupt erfasst werden) zu berücksichtigen.

Ähnlich dazu ergeben sich durch die Speicherung im Binärcode weitere Probleme. Was ist genauer: Der Text «Ein Drittel», der Text «1/3», die Zahl «0.3333333333333333», ein Objekt das als Zähler 1 und als Nenner 3 speichert? Theoretisch sind diese Daten identisch, doch bei der Weiterverarbeitung können unterschiedliche Ergebnisse entstehen, obwohl hier überall die (fast) identische Genauigkeit als Eingabe verwendet wurde. Es kommt daher zusätzlich auf die Verarbeitung der Daten (und etwaige «gekürzte» Zwischenergebnisse) an, bis diese ausgegeben werden.

II. Was ist das Original?

Bei einem Papierdokument ist diese Frage, was das Original und was eine Kopie ist, im Allgemeinen leicht zu beantworten. Doch bei Daten ist dieses Konzept problematisch, da (korrekte) Kopien absolut identisch mit dem Original sind. Ist daher der Datenträger (der von den darauf enthaltenen Daten strikt zu unterscheiden ist) maßgeblich? In anderen Ländern ist bzw war dies bis vor eingen Jahren ein relevantes Problem, da nur «Originale» vor Gericht vorgelegt werden durften (sofern sie noch existieren), und Ausdrucke von Daten nicht als Original galten[3].

[3] Dies ist daher nur eingeschränkt eine Frage des Beweiswertes und mehr eine der Zulässigkeit von Beweismitteln: DEUTSCH, Die Beweiskraft elektronischer Dokumente, JurPC Web-Dok.

Abgesehen von der rechtlichen Beurteilung ist dies ein praktisches Problem: Je mehr Kopien das Vorgelegte vom «Original» entfernt ist, desto eher können Übertragungsfehler und Veränderungen vorkommen, da länger und für mehr Personen Möglichkeiten zur Manipulation existierten.

III. Woher stammen die Daten?

Ein Mobiltelefon Modell X des Herstellers Y sieht exakt so aus, wie alle anderen Exemplare dieses Typs; gleiches gilt für Datenträger. Dennoch müssen sie einer Person zugeordnet werden. Dies ist kein nur in der IT auftretendes Problem, aber aufgrund der exakten Gleichheit (meist keine erkennbaren Fertigungstoleranzen) der Exemplare besonders problematisch. Es ist daher notwendig, die Geräte oder Medien zu individualisieren, zB über unveränderbare Seriennummern. Diese können bereits vorhanden sein (zB die IMEI bei Mobiltelefonen), oder auch nicht (USB-Sticks sollten weltweit eindeutige Seriennummern besitzen, doch insb bei Billigprodukten fehlt diese oder ist für alle Exemplare einer Baureihe identisch). Denn erst wenn exakt bekannt ist, um **welches** Gerät es sich handelt, besteht die Möglichkeit festzustellen, **wessen** Gerät es ist bzw war.

IV. Sind die vorgelegten Daten vollständig?

Bei Cloud- oder anderen Online-Diensten ist es fast nie möglich, Datenträger zu erhalten, sondern es ist nur ein Online-Zugriff über das Internet möglich. Damit stellt sich die Frage, ob das, was abgefragt wurde, auch tatsächlich alles ist, was vorhanden ist oder war. Ansonsten könnte durch die Auswahl sichergestellter Daten ein falscher Eindruck erzeugt werden. Weiters können etwa bereits gelöschte Daten nicht gesichert werden, was hingegen bei einer vollständigen Bit-Kopie einer Festplatte der Fall ist. Hinzu kommt, dass der Online-Zugriff auf die Daten zur Sicherstellung nicht unbedingt

188/2000, https://www.jurpc.de/jurpc/show?id=20000188 (16.9.2019). Siehe zB den Civil Evidence Act 1995, insb § 4, http://www.legislation.gov.uk/ukpga/1995/38/contents (2.10.2019), durch den erst die allgemeine Zulässigkeit von elektronischen Dokumenten eingeführt wurde (und damit von «Zulässigkeit» zu «Beweiswert» gewechselt wurde).

der erste sein muss. Der Vorwurf, sich zuerst anzumelden, «Problematisches» zu löschen und sich abzumelden, und erst anschließend offiziell die Beweise sichern, liegt auf der Hand und ist zB durch eine geeignete Dokumentation des Vorgangs zu entkräften. Derartige Manipulationen sind jedoch von jedem beliebigen System mit Internetzugriff aus möglich, sofern die erforderlichen Passworte etc bekannt sind; dies gilt daher exakt gleich für Verdächtige wie Dritte. Möglichkeiten für rasche Sperrung, umgehende Beweiserhebung etc sind daher besonders wichtig, ebenso wie Dokumentation und Nachvollziehbarkeit (aber eben nicht notwendigerweise Wiederholbarkeit) der Datensammlung.

V. Wurden die Daten verändert?

Wird eine 1 in eine 0 verändert, so ist dies anschließend nicht mehr nachvollziehbar: Bei praktisch keinem modernen Speichermedium ist es möglich, den vorherigen Dateninhalt festzustellen[4]. Es ist daher besonders wichtig sicherzustellen, dass keine undokumentierten Änderungen erfolgen.

Bei der Untersuchung «lebender» Systeme, wie zB eines kompromittierten Webservers durch die Anzeige von Daten darin oder das Ausführen zusätzlicher Programme darauf ist dieses Problem besonders ausgeprägt: Jede Untersuchung muss das System zwangsweise verändern, und während der Untersuchung ändert es sich selbst sowie potentiell zusätzlich durch Aktionen Dritter (hier etwa Web-Zugriffe von außen). Dies kann zwar reduziert, aber nicht verhindert werden. In derartigen Fällen ist daher weder eine Wiederholbarkeit der Untersuchung noch ein konsistentes Bild zu einem einzelnen Zeitpunkt (während der erste Teil kopiert/dargestellt wird, ändert sich der Rest und

[4] WRIGTH/KLEIMANN/SUNDHAR, Overwriting Hard Drive Data: The Great Wiping Controversy, In Sekar/Pujari (Hrsg.), ICISS 2008, LNCS 5352 (2008), 243. Die Wahrscheinlichkeit für die Rekonstruktion eines einzelnen Bytes auf einer zuvor nie benutzten Festplatte nach einmaligem Überschreiben beträgt ca. 32,8%, bzw für 4 Bytes (zB eine IP-Adresse) 1,16%. Durch die seither noch stark verdichtete Speicherung der Einzelbits ist eine Rekonstruktion heute vollständig auszuschließend. Ebenso ist nach Kissel/Regenscheid/Scholl/Stine, NIST SP800-88 Rev. 1, Guidelines for Media Sanitization (2014) https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-88r1.pdf (2.10.2019) einmaliges Überschreiben ausreichend. Für SSD gilt dies gleichermaßen, doch ist dort ein vollständiges Überschreiben jedes einzelnen Bits sehr viel schwieriger, bzw praktisch nur über das «Secure Erase» Kommando möglich (sodass zB auch kryptographische Methoden oder physische Zerstörung eingesetzt werden).

umgekehrt) gegeben. Damit ist keine Möglichkeit vorhanden, das System identisch und funktionsfähig zu reproduzieren, um weitere Analysen durchzuführen.

VI. Können Probleme lokalisiert werden?

Wird ein Papierdokument nachträglich verändert oder fehlt eine Seite, so kann das «Problem» auf bestimmte Bereiche eingeschränkt werden: Der Rest des Dokuments ist weiterhin verlässlich. Bei elektronischen Daten ist dies oft nicht möglich. Es wurde zB nur ein Teil eines Dokumentes beschädigt, oder es konnten nur Fragmente rekonstruiert werden, doch bei der üblichen Sicherung mittels Hashwerten (siehe unten) wird durch die kleinste Veränderung (ein einziges Bit), der Beweiswert des gesamten Dokumentes vernichtet – es ist unmöglich festzustellen, wo eine Änderung erfolgte, und ob diese Änderung (oder in gewissen Grenzen die fehlenden Teile) aus einem Bit oder vielen Megabyte bestand. Gleiches gilt für elektronische Signaturen: Kleine/Große Probleme führen nicht zu einer leichten/starken Reduktion des Beweiswertes, sondern schon der kleinste Fehler vernichtet diesen fast vollständig.

VII. Wann wurden die Daten erzeugt?

Bei vielen physischen Beweisen kann ihr Alter zumindest grob abgeschätzt werden. Bei Daten ist dies unmöglich. Eine gewisse Approximation ist durch das Trägermedium möglich, doch in vielen Fällen ist dieses irrelevant (Online-Daten) oder wenig hilfreich (etwa bei Backups durch deren regelmäßige und wiederholte Nutzung). Es ist daher nur durch zusätzliche Maßnahmen möglich, zumindest eine Eingrenzung des Zeitbereichs vorzunehmen. Hier ist ein wichtiger Unterschied, dass diese Maßnahmen meistens nur den Zeitpunkt der Beweiserhebung belegen, aber nicht den der eigentlichen Datenerzeugung.

Ein weiterer Aspekt ist, dass sich Daten in Online-Diensten rasch und häufig ändern. So ist die Beantwortung der Frage «Was ist heute auf der Startseite von orf.at?» leicht, während die Beantwortung der Frage nach dem Inhalt vor einem Jahr sehr schwer ist, und vielfach unmöglich sein wird (uU sogar für den Betreiber). Damit bestimmt der Zeitpunkt der Beweiserhebung sehr stark den Inhalt des Beweises – dies sollte zumindest klargestellt und dokumentiert sein, um Manipulationen durch die Auswahl des Zeitpunktes zu erschweren.

VIII. Durch wen bzw wo wurden die Daten erhoben?

Ein Fingerabdruck verändert sich nicht, egal wer ihn findet und sicherstellt (abgesehen von absichtlicher Manipulation und Fälschung). Doch bei Online-Diensten ist dies nicht gegeben und Unterschiede können durch sehr vielfältige und schwer nachvollziehbare Gründe entstehen. Ein Beispiel hierfür ist das Google-Suchergebnis nach einem bestimmten Suchbegriff. Je nach vorherigen Suchen, Anmeldung bei Google (weitere Daten über die Person von früher bekannt), früheren Anfragen von dieser IP-Adresse (evtl dynamisch vergeben und daher wenige Minuten zuvor von jemand Anderem genutzt), vorher besuchten Webseiten etc wird eine andere Ergebnisliste erstellt. Selbst wenn diese Liste von einer Person auf einem Rechner zum selben Zeitpunkt lediglich mit zwei unterschiedlichen Browsern erstellt wird, ist sie nicht immer identisch. Da nicht exakt bekannt ist, wie stark einzelne Elemente die Ergebnisse beeinflussen (und Details streng geheim gehalten werden), ist weder eine Wiederholung noch ein Vergleich mit anderen Personen möglich. Es kann nicht einmal gesagt werden, welcher «Anteil» an der Ergebnisliste «objektiv» (=unabhängig von Person/Rechner/Browser/...) ist, und welcher davon auf individuelle (und schon gar nicht welche spezifischen) Elemente zurückzuführen ist. Damit verändern jedoch die Person bzw die Umstände der Beweiserhebung den sichergestellten Inhalt.

IX. Wie wurden die Daten erzeugt?

Die am besten gesicherten Daten sind nutzlos, wenn sie nicht für den Beweis relevante Elemente enthalten. Daraus folgt gleichermaßen, dass es wichtig ist, wie die Daten erzeugt wurden. Werden beispielsweise nur «Probleme» aufgezeichnet, so lässt sich aus der Abwesenheit erfolgreicher Transaktionen/Aufgaben/etc offensichtlich nichts schließen. Doch in viele Fällen ist dies deutlich unklarer und schwerer zu erkennen, zB wenn keine offizielle Dokumentation verfügbar ist, was exakt gespeichert wird. Auch Programmfehler (stürzt das Programm ab, kann es dieses Faktum bzw den Grund hierfür uU nicht mehr speichern) sowie Ausnahmesituationen (eine der ersten Aktionen von Hackern ist es, Spuren des Eindringens bzw der hierfür verwendeten Methode zu beseitigen) sind zu berücksichtigen.

X. Was sind die technischen Möglichkeiten?

Elektronische Signaturen genießen echte Beweiserleichterungen (siehe Art 25 Abs 2 eIDAS-VO iVm § 294 ZPO), da auf die hierfür eingesetzte Technik stark vertraut wird. Dies ist jedoch potentiell problematisch: Technik, die heute sicher ist, kann es morgen nicht mehr sein. Noch schlimmer ist es, wenn sich heute herausstellt, dass sie es schon seit einem Jahr nicht mehr war. Viele Technologien sind daher mit einem Ablaufdatum zu versehen bzw ist bei ihrer Bewertung zusätzlich ein Zeitpunkt zu berücksichtigen: Was waren zu dem Zeitpunkt, zu dem die Daten gesichert wurden (und die zwischenzeitlich nicht verändert wurden!), die technologischen Möglichkeiten? Waren sie damals als sicher anzusehen bzw sind sie es noch heute (dh keine Möglichkeit nachträglicher Erzeugung oder unerkannter Veränderungen)?

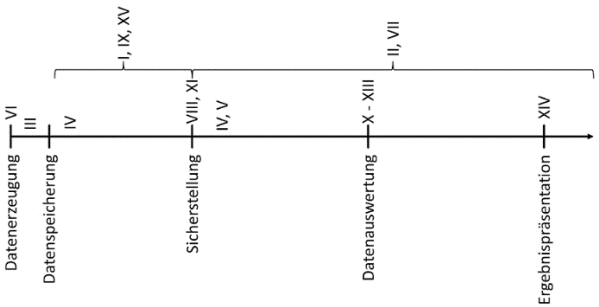
XI. Wie wurden die Daten ausgewählt und aufbereitet?

Gewöhnliche PCs besitzten aktuell eine Speicherkapazität von mehreren Terabyte – ohne die Berücksichtigung von USB-Sticks, externen Festplatten und Cloud-Speicher. Es ist daher zeitlich bzw finanziell unmöglich, in einem Verfahren **alle** Daten auszuwerten oder darzustellen. Stattdessen muss eine Auswahl erfolgen: Wie passierte etwas? Wer führte eine Aktion durch? Wer erstellte die Regeln/Prinzipien für eine Verarbeitung? Dies beeinflusst sowohl die Sicherstellung der Daten wie auch deren Auswertung. So ist etwa die Auswahl der zu beantwortenden Fragen ein bekanntes Problem bei Sachverständigen-Gutachten, bzw eine Chance zur «Manipulation» des Ergebnisses. Wird sehr speziell nach Details gefragt, werden andere, uU sehr wohl relevante, Daten nicht gesichert und untersucht, und damit effektiv ignoriert.

Darüber hinaus ist die Aufbereitung der Daten indirekt relevant: Sie verändert nicht unbedingt den Beweiswert, aber dessen Wahrnehmung. Wird etwa eine ausgewählte E-Mail mit problematischem Inhalt und dem Text «17/100» präsentiert, so ist dies nur «ein» Problem. Eine grafische Darstellung von 100 E-Mails (so klein, dass unleserlich), mit Zoom auf den Text einer davon lässt ihren Wert als nur eine von sehr vielen deutlich höher erscheinen. Dies ist wichtig zu berücksichtigen, wird aber im Folgenden nicht mehr behandelt, da es kein Problem des Beweises an sich ist, sondern nur seiner Präsentation.

C. BEURTEILUNGSKRITERIEN

In diesem Abschnitt werden diverse Kriterien dargestellt, anhand derer die Beweiskraft elektronischer Daten beurteilt werden kann. Die große Anzahl darf nicht dazu verleiten zu glauben, in jedem Fall wären alle vollständig anzuwenden/auszuwerten, oder die Daten wären als Beweis wertlos. Eine vollständige Umsetzung ist zwar wünschenswert und würde hervorragende Beweise liefern, aber nicht alle Kriterien sind überall anwendbar und dies ist weder zu erwarten noch kostengünstig. Für einen besseren Überblick wird in der Abbildung dargestellt, zu welchem Zeitpunkt des Lebenszyklus von Daten die jeweiligen Kriterien die wichtigste Rolle spielen bzw anzuwenden sind. Offensichtlich fehlt hier die endgültige Löschung der Daten – doch diese ist für Beweise nur negativ relevant, im Sinne ihrer Verhinderung um sie mit ihrem Beweiswert zu erhalten.



Die Kriterien im Lebenszyklus von Daten

I. Sichere Aufbewahrung bis zur Sicherstellung

Zwischen der Erzeugung der Daten und deren Sicherstellung liegt uU eine lange Zeitspanne, in welcher diverse Manipulationen möglich sind[5]. Für den Beweiswert ist es

[5] Siehe als Beispiel SMS: Durch ein Mobiltelefon (normale Benutzerschnittstelle) können diese gar nicht verändert werden; bei Zugriff auf die SIM-Karte ist dies jedoch mit trivialen Mitteln möglich und kann später nicht erkannt werden. KNASMÜLLER, Zur Beweiskraft von elektronischen Nachrichten (SMS, E-Mail, Facebook und WhatsApp), Sachverständige 4/2015, 204

daher wichitg festzustellen, wie bzw wo die Daten in diesem Zeitraum aufbewahrt wurden. Kriterien hierfür sind uA (neben explizit diskutierten wie zB Signaturen):

- Sind die Daten dauernd in Verwendung, so werden sie durch Programmfehler leichter verändert, als wenn sie nur einmal geschrieben und dann nicht mehr modifiziert werden. Analog dazu sind Backups auf Offline-Medien vor Manipulationen deutlich sicherer als im aktiven System gespeichert Daten.
- Werden die Daten auf einem anderen System gespeichert, so sind Manipulationen schwieriger. Aus diesem Grund erfolgt zB zentralisiertes Logging, da unmittelbar nach der Datenerzeugung diese an ein separates (und sonst nicht zugängliches) System gesendet und dort gespeichert werden.
- Um so weniger Personen Zugriff auf diese Daten besitzen und je unabhängiger diese von den Verfahrensparteien sind, desto höher ist der Beweiswert.

II. Sichere Aufbewahrung nach der Sicherstellung

Eine beliebte Rechtfertigung ist, dass problematische Daten erst nachträglich auf die eigenen Datenträger gespeichert wurden oder durch die Polizei, den Sachverständigen, unbekannte Dritte etc verändert wurden. Technisch ist dies fast nie auszuschließen, denn diese Möglichkeit besteht im Grundsatz immer. Für den Beweiswert ist es daher wichtig, Maßnahmen zu ergreifen, die derartige Argumente entkräften oder sehr viel unwahrscheinlicher machen. Hierzu zählen uA:

- Hashwerte und Signaturen: Siehe unten. Hier relevant ist, dass diese Maßnahmen möglichst bald durchgeführt werden sollten. Idealerweise erfolgen sie sofort vor Ort bei der Untersuchung/Sicherstellung/Übergabe etc. Aufgrund der umfangreichen Datenmenge moderner Datenträger ist dies jedoch vor Ort oft unpraktisch, da eine Duplikation und Hashberechnung viele Stunden dauern kann.
- Vornahme der Sicherstellung durch vertrauenswürdige Personen ohne Interesse am Ausgang des Verfahrens. Dies ist jedoch kein technisches Mittel, sondern bezieht sich auf den Vorgang.
- Speicherung auf WORM-Systemen: Medien, welche nur ein einziges Mal beschrieben werden können («Write Once Read Many»), oder Systeme, welche ein Überschreiben technisch unterbinden, verhindern zuverlässig spätere Änderungen.

Hierbei ist es wichtig, die Identität der Medien sicherzustellen, da diese ansonsten trivial durch Neue mit modifizierten Inhalten ersetzt werden könnten.

– Absicherung des Zugangs: Es muss durch technische und/oder organisatorische Maßnahmen sichergestellt werden, dass nur berechtigte Personen Zugang zu den Daten haben, sowohl im Hinblick auf Geheimhaltung wie auch zur Verhinderung von Modifikationen.

Hier ist weiters sicherzustellen, dass eine autorisierte Weitergabe, etwa zur weiteren Auswertung, auch tatsächlich erfolgen kann, zB durch den Export der Daten in einem standardisierten Format. Denn bereits die Möglichkeit einer Zweitüberprüfung (eine tatsächlich Erfolgreiche natürlich umso mehr), erhöht den Beweiswert von Daten.

III. Integritätsicherung und Personenzuordnung: Elektronische Signaturen

Wichtig ist bei elektronischen Signaturen insb ihr technischer Aspekt, dass die Daten ohne Invalidierung der Signatur nachträglich durch andere Personen nicht mehr geändert werden konnten. Für den Beweiswert ist zusätzlich in bestimmten Fällen wichtig, wer die Signatur erstellt hat, da diese Person, im Gegensatz zu allen anderen, auch nachträglich noch Modifikationen (und eine erneute Signatur) hätte vornehmen können.

Rechtlich ist zu beachten, dass eine Signatur die Vermutung begründet, die damit bestätigte Aussage stamme von der signierenden Person. Dies passt gut auf Verträge, doch bei bloßen Daten, zB einer Folge an Messwerten, ist diese Vermutung nutzlos, da es sich nicht um eine Willens- oder Wissenserklärung handelt[6]. In diesem Fall ist daran zu denken, dass die «Erklärung» einer Signatur sein könnte, dass diese Daten zum Signierungszeitpunkt vorlagen – dies jedoch keine Aussage über deren Quelle, Korrektheit, Akzeptanz etc beinhaltet. Effektiv bleibt daher lediglich der Integritätsaspekt übrig, welcher technisch sehr gut gesichert ist.

---

6 Dazu näher KNOPP, Rechtliche Perspektiven zur digitalen Beweisführung, In: FISCHER/MAEHLE/REISCHUK, Informatik 2009 – Im Focus das Leben, Bonn 2009, 169/1552.

Ein weiteres Problem von Signaturen ist ihr Erstellungszeitpunkt. Dieser ist bei einer Signatur nicht aus ihr selbst heraus festzustellen – hierfür wird ein Zeitstempel (siehe unten) benötigt. Eine Signatur beweist daher eine Unveränderheit des Inhalts, aber nicht den Zeitpunkt, ab dem dies gilt, oder ab wann die Daten existierten. Der Zeitfaktor ist bei Signaturen im Hinblick auf den Beweiswert zusätzlich in einem anderen Aspekt wichtig, sofern es sich um große Zeiträume handelt: Sowohl Algorithmen als auch Schlüssellängen von Signaturen müssen im Laufe der Zeit geändert werden. Was früher als sicher angesehen werden konnte, ist es heute nicht mehr. Daher ist uU eine Nachsignierung nötig, oder der Beweiswert der Daten ist zu reduzieren, sofern deren Existenz zu einem früheren Zeitpunkt nicht nachgewiesen werden kann[7]. Dies bedeutet daher, Beweise zu benötigen, damit andere Beweise vertrauenswürdig sind. Damit entsteht eine Hierarchie an Beweisen, welche explizit darzustellen ist um sicherzugehen, dass bei der Bewertung auch wirklich an der Spitze begonnen wird, da sich alle Unsicherheiten nach unten hin fortplanzen.

IV. Zeitliche Einordnung

Zeitstempel dienen dazu, Daten eine sichere zeitliche Zuordnung zu geben. Allerdings funktionieren sie nur in eine Richtung: Die Daten müssen vor diesem Zeitpunkt existiert haben. Wie lange vorher (Sekunden oder Jahre?) wird dadurch nicht ausgesagt. Auch ist ein Beweis der (noch) Nicht-Existenz bestimmter Daten zu einem Zeitpunkt derzeit iA unmöglich. Zeitstempel erfordern die dauerhafte Existenz der Daten selbst: Die Bestätigung «Der Kaufvertrag über das Auto zwischen X und Y wurde am Z.Z.Z geschlossen» sagt für sich alleine bereits etwas über die bestätigten Daten aus. Demgegenüber besteht ein Zeitstempel nur aus einem Hashwert, aus welchem gar nichts über die Daten gewonnen werden kann. Andere zeitliche Bestätigungen sind entweder keine Daten (sondern zB Zeugenaussagen) oder sind als Redudanz oder anderweitig bestätigende Daten (siehe unten) einzuordnen.

[7] Siehe BSI technische Richtlinie 03125. Beweiswerterhaltung kryptographisch signierter Dokumente. Anlage TR-ESOR-B: Profilierung für Bundesbehörden.

Für den Beweiswert von Zeitstempeln sind folgende Punkte relevant:

- Von wem wurde der Zeitstempel erstellt? Nur wenn dies ein unabhängiger Dritter ist, erbringen sie eine vertrauenswürdige Bestätigung, da durch eine Manipulation der verwendeten Uhr eine Rückdatierung trivial ist.
- In den meisten Fällen werden (zB aus Geheimhaltungsgründen) nicht die vollständigen Daten bestätigt, sondern nur deren Hashwert. Es sind daher zusätzlich die Elemente des nächsten Abschnitts zu berücksichtigen.
- Da es sich bei Zeitstempeln technisch gesehen um elektronische Signaturen handelt, sind auch die Betrachtungen des vorigen Abschnitts relevant.
- Potentiell relevant (in der Praxis wohl nur sehr selten) ist die Genauigkeit der Zeitangabe und ihrer Quelle. Normalerweise handelt es sich um Atomuhren und Zeitangaben mit Sekundengenauigkeit. Da Zeitstempel über das Internet erfolgen (mit Verschlüsselung zur Sicherung, dh mehrfache Verzögerung der Kommunikationslatenz), sind genauere Zeitangaben meist sinnlos.

V. Unverändertheit der Daten: Hashwerte

Um Sicherzustellen, dass Daten unverändert sind, werden hauptsächlich Hashwerte eingesetzt. Hierbei ist zu berücksichtigen, dass Hashwerte mathematisch notwendig bestimmten Einschränkungen unterliegen: Es werden zB unendlich viele Dokumente auf den identischen Hashwert abgebildet. Die Qualität eines Hash-Algorithmus besteht vereinfacht dargestellt darin, dass es sich hierbei um «uninteressante» Alternativen (komplett anders, also keine bloß leicht veränderte Variante) handelt, und es möglichst schwer sein soll, solche äquivalente Daten zu finden. Dem Algorithmus ist daher besondere Beachtung zu schenken: Handelt es sich um einen, der heute noch als sicher gilt? Liegt die Erstellung des Hashwertes lange Zeit zurück (und ist der Hash-Algorithmus inzwischen zB gebrochen worden), so ist wie bei Signaturen ein Beweis der Existenz zu einem früheren Zeitpunkt – als der Algorithmus noch sicher war – nötig. Dies erfolgt typischerweise durch Zeitstempel – welche jedoch selbst auf Hashwerten beruhen. Daher sind für diese unterschiedliche Hash-Algorithmen oder andere Sicherheitsmaßnahmen erforderlich.

Der Vorteil von Hashwerten, dass schon kleinste Änderungen das Ergebnis ändern, ist gleichzeitig ein Nachteil: Wird aus irgendeinem Grund auch nur ein Bit verändert und ist dessen Position nicht mehr bekannt/feststellbar, so sind – wegen des dadurch veränderten Hashwertes – **alle** Daten nutzlos. Es sollte daher nicht nur auf den Hashwert alleine geachtet werden, und selbst bei Fehlern muss der Beweiswert nicht immer auf Null gesetzt werden, sofern andere Maßnahmen, wie etwa besonderer Zugriffs-/Änderungsschutz, gesetzt wurden. Als Gegenmaßnahme ist es zB möglich, nicht nur einen Hashwert für die Daten als Ganzes zu erzeugen, sondern ebenso für Teile (zB Dateien, Blöcke, zusammengehörige Daten etc), sodass nur der Beweiswert eines (hoffentlich weniger relevanten) Teils reduziert wird, ohne andere Teile zu invalidieren.

Die Sensitivität eines Hashwertes gegenüber Änderungen ist in einem weiteren Anwendungsbereich wichtig: Dem Vergleich von Bildern (typischerweise Kinderpornographie oÄ illegale Abbildungen – diese dürfen nicht aufbewahrt/weitergegeben werden, deren Hashwerte jedoch problemlos). Um ähnliche/umcodierte/... Bilder dennoch identifzieren zu können, werden «unscharfe Hashwerte» (zB perceptual hashing) eingesetzt. Diese erzeugen bei «optisch ähnlichen» Bildern den identischen Hashwert, auch wenn die Binärdaten abweichen. Dies bedeutet jedoch, dass ein zusätzlicher (typ menschlicher) Vergleich bzw Überprüfung notwendig ist, da ein identischer Hashwert auch durch zufällige oder absichtliche Fehler erzeugt werden kann. Für den Beweiswert ist daher bei diesen Arten wichtig festzustellen, um welche Art von Hash es sich handelt, welcher Algorithmus verwendet wurde, und wie leicht dieser zu beeinflussen ist bzw ob eine Nachkontrolle erfolgte.

Ähnlich zu WORM-Medien ist ein Hashwert nur so viel wert, wie seine Dokumentation bzw Aufbewahrung die Unverändertheit garantiert. Denn werden die Daten geändert, kann trivial der neue Hashwert berechnet und der alte ersetzt werden. Da es sich um sehr lange Zahlen handelt, ist praktisch ausgeschlossen, dass Menschen sich diesen merken. Der sicheren Dokumentation, sodass eine nachträgliche Änderung effektiv unmöglich ist, ist daher große Aufmerksamkeit für den Beweiswert zu schenken. Eine gewisse Absicherung können Zeitstempel bieten – diese beweisen ein Mindestalter der Daten (=des Hashwertes) und könnten uU – in Zusammenhang mit einer Speicherung beim Ersteller des Zeitstempels – auch nachweisen, dass keine sonstigen (und zB nicht

durch andere Dokumente erklärbaren) Werte (=frühere Zeitstempel vor einer Modifikation) des selben Kunden existieren[8].

VI. Darlegung der Datenerzeugung

Die Art der Datenerzeugung lässt einen Rückschluss auf ihren Beweiswert zu, bzw wofür sie sich überhaupt als Beweis eignen. So sind zB manuell eingegebene Daten nur so verlässlich, wie sie automatisiert überprüft und zeitnahe eingegeben werden, bzw wie vertrauenswürdig, sorgfältig etc die sie eingebende Person ist. Hierher gehört weiters, ob die Daten zwischen ihrer tatsächlichen Erzeugung bis hin zur Aufzeichnung über Kommunikationsnetze übertragen werden. Dies wäre eine typische Stelle für Fehler bzw Manipulationen. Stammen die Daten zB aus anderen Geräten, ist zusätzlich zu berücksichtigen, ob auch diese korrekt funktionierten. Weiters lässt sich daraus ersehen, welche Inhalte überhaupt gespeichert werden: Alles Vorliegende bzw wie die Auswahl erfolgt. Damit wird vermieden, Vorhandenes irrtümlich als vollständig anzusehen, und so einen falschen Gegenschluss, «Nicht aufgezeichnet, daher nicht passiert» zu ziehen.

Ein weiterer relevanter Aspekt ist der Grund der Datenerzeugung: Werden die Daten lediglich zur internen Dokumentation erzeugt, ist von geringer Genauigkeit und schwacher Absicherung auszugehen und der Beweiswert daher nicht besonders hoch. Handelt es sich jedoch um Daten, die unmittelbar für weitere betriebliche Funktionen verwendet werden, kann auf sie sehr stark vertraut werden, da ansonsten später Fehler auftreten würden. Sollen sie schließlich von vornherein geplant später als Beweis dienen (zB gegenüber dem Finanzamt), ist von besseren Schutzmaßnahmen und höherer Genauigkeit und Verlässlichkeit auszugehen – bzw umgekehrt von gezielter Manipulation von Anfang an. Diese Punkte können als Anhaltspunkt dienen, aber eine genauere Überprüfung (sofern zeitlich/finanziell/… möglich) bietet verlässlichere Ergebnisse über den Beweiswert.

---

[8] Dies würde jedoch erforderm, dass für alle anderen von dieser Person/Institution erzeugten Zeitstempel die zugehörigen Daten vorgelegt werden (denn aus dem Hashwert ist kein Rückschluss auf die Daten möglich). Aus Geheimhaltungsgründen (oder zB bei Verlust auch nur eines einzigen Dokuments) kann dies sehr problematisch sein. In der Praxis dürfte dies daher nicht vorkommen bzw zielführend sein.

Die Datenerzeugung umfasst auch eine Darlegung der Vollständigkeit der Daten (im Gegensatz zum Inhalt – siehe oben): Ist davon auszugehen, dass jedenfalls ein Teil verloren geht (zB bei VoIP oder Video-Streaming fast garantiert), oder wurden umfassende Maßnahmen unternommen, um eine vollständige Protokollierung zu garantieren? Hinsichtlich fehlender Daten ist zusätzlich das Fehlermodell zu berücksichtigen. So werden bei VoIP zwar häufig bis zu mehreren Prozent der Daten fehlen, doch diese verteilen sich entweder stark (gelegentlich fehlt ein einzelnes Paket – akustisch praktisch nicht feststellbar) oder führen zu einem langen, deutlich merkbaren, Ausfall (mehrere Sekunden lange Pause). Dass exakt nur ein einziges Wort, dieses aber vollständig, verloren geht, ist dagegen äußerst unwahrscheinlich. Umgekehrt ist zB bei Systemabstürzen davon auszugehen, dass typischerweise gerade die letzten Einträge, welche auf den Absturz bzw dessen Gründe hinweisen, verloren gehen und daher zwar theoretisch als Beweise hervorragend geeignet wären, ihr Fehlen aber nicht unbedingt auf eine Manipulation hinweist.

VII. Chain of Custody

Die «Chain of Custody» soll die Identität und Integrität von Beweisen jeder Art sicherstellen[9]. In manchen Rechtssystemen ist sie essentiell für den Beweiswert, doch selbst wenn ihre Verletzung keine Wertlosigkeit von Beweisen erzeugt, ermöglicht eine Befolgung einen besseren Beweiswert. Gerade bei Daten ist dies aufgrund der leichten und unerkennbaren Veränderbarkeit besonders wichtig, wobei die Identität hier zurücktritt. Diese ist für Datenträger relevant, doch sind diese selten das Relevante an einem Beweis. Die Integrität, dh die Unverändertheit der Daten ist hingegen essentiell. Typischerweise wird diese über Hashwerte sichergestellt (siehe oben). Sollte dies fehlschlagen, zB zufällige Bitänderungen, Beschädigung des Datenträgers sodass nur mehr Teile auslesbar sind etc, so ist eine lückenlose Kette der Personen, die Zugriff auf das Speichermedium hatten, sowie dessen Art der Aufbewahrung, wichtig. Bei Daten ist daher an eine sichere Speicherung bzw Lagerung mit exakter Protokollierung aller Zugriffe zu denken. Wird dies mit der Verhinderung bzw Protokollierung von Schreibzugriffen

[9] Sichere Aufbewahrung des Beweismittels und Dokumentation jeder Weitergabe an die nächste Person der Kette bis zum aktuellen Zeitpunkt.

kombiniert, kann auch Daten mit fehlerhaften oder fehlenden Hashwerten ein hoher Beweiswert zugemessen werden.

Abgesehen von den Daten an sich gibt es einen zweiten wichtigen Aspekt: Der Hashwert selbst muss sicher dokumentiert werden. Hierbei können zB el. Signaturen (siehe oben) Einsatz finden. Dies ist allerdings keine echte Chain of Custody, da es hierbei nicht mehr darauf ankommt, wann der Hashwert an wen «übergeben» wurde oder ob das Papier des Ausdrucks das Original ist, sondern um die Bestätigung des Wertes durch eine bestimmte Person.

VIII. Dokumentation des Sicherstellungsvorgangs

Eine Dokumentation des Vorgangs der Sicherstellung der Daten ist besonders wichtig, wenn dieser nicht wiederholbar ist. Handelt es sich um aktive Systeme (zB Firmen-Server) oder lediglich transiente Daten (zB den Arbeitsspeicher eines gerade ausgeführten Prozesses), so kann zwar technisch eine erneute Datensammlung stattfinden, doch muss diese zwangsläufig ein anderes Ergebnis liefern. Ist eine Wiederholung überhaupt nicht möglich (das Gerät kann zB nicht dauerhaft mit Strom versorgt und gleichzeitig sichergestellt werden; Daten in einem Cloud-Account eines Dritten, der nicht unbegrenzt bezahlt werden kann oder dessen Dienst eingestellt wird usw), so ist eine umfassende Dokumentation die einzige Möglichkeit, den Beweiswert der erhobenen Daten sicherzustellen.

Denn dies ermöglicht es, später zumindest den Vorgang der Datensammlung nachzuvollziehen und zu bewerten: Wurden alle (für den Beweiszweck erforderlichen) Daten gesammelt, wurde dabei korrekt vorgegangen, traten dabei Fehler auf, sind die Daten zum Vorgang konsistent etc. Da bei einer Live-Erhebung von Daten diese auch verändert werden, kann aus der Dokumentation des Vorgangs weiters geschlossen werden, welche Änderungen auf die Beweismittelsicherung selbst zurückzuführen sind, bzw welche Daten hierdurch zerstört wurden. Werden zB Dateien in einem aktiven Dateisystem kopiert, so verändert dies uU (je nach Methode, Betriebssystem etc) deren Zeitstempel: Den letzten Zugriffzeitpunkt der Quelldatei, sowie Schreib- und/oder Erstellungsdatum der Kopie. Solche Änderungen sind bei der Auswertung zu berücksichtigen, was jedoch nur aufgrund der Dokumentation möglich ist.

Zusätzlich ermöglicht eine umfassende Dokumentation festzustellen, welche Geräte und Software in welcher Version eingesetzt wurden. Hieraus lässt sich auf etwaige Fehler oder Ungenauigkeiten schließen – was ggf den Beweiswert reduziert, bzw ermöglicht dies eine exakte Wiederholung.

IX. Sicherheitsmaßnahmen

Eine Darstellung der vorhandenen Sicherheitsmaßnahmen, welche die relevanten Daten bzw Systeme schützen, ergänzt die Beurteilung des Beweiswertes von Daten. Sind diese zB besonders gegen Veränderungen geschützt, so kann auf sie eher vertraut werden, als wenn diese durch jeden Benutzer beliebig verändert werden können. Sicherheitsmaßnahmen müssen nicht dazu dienen, Beweise zu erzeugen oder bewahren, sondern können auch auf anderen Gründen beruhen. Beispielsweise dient verschlüsselte Übertragung von Webseiten dazu, Dritte zu hindern, während der Übertragung Veränderungen vorzunehmen **und** den Inhalt geheimzuhalten. Um dies zu erreichen ist es gleichzeitig nötig zu verifizieren, dass die Daten auch tatsächlich vom intendierten Webserver stammen. Protokolle des Zugriffs auf solch abgesicherte Websites, deren URLs etc beweisen daher deutlich besser, dass der angezeigte Inhalt auch tatsächlich vom Server stammt, als wenn keine Sicherheitsmaßnahmen eingesetzt wurden (unverschlüsselte Übertragung, daher unentdeckte Modifikation durch jedes System auf dem Übertragungsweg möglich)[10]. Naturgemäß ist hier individuell zu prüfen, welche Auswirkung die konkret vorhandenen Sicherheitsmaßnahmen auf die jeweiligen Beweise hatten.

Generelle technisch-organisatorische Maßnahmen oder Zertifizierungen können ebenso ein hohes Sicherheitsniveau belegen, welches den Wert von Beweisen erhöht – In jeder Hinsicht wäre es grundsätzlich schwerer, Manipulationen vorzunehmen, wie wenn auf Sicherheit kein Wert gelegt wird.

In speziellen Fällen können Sicherheitsmaßnahmen Beweise sogar mit herausragendem Beweiswert ausstatten, etwa beim Einsatz von Hardware-Sicherheits-Modulen –

[10] Unverschlüsselte Webseiten können auch trivial vom lokalen Computer aus angezeigt werden und dennoch den richtigen Namen in der Browser-Adresszeile anzeigen.

bei diesen ist es nach dem Stand der Technik unmöglich, darin enthaltene kryptographische Schlüssel zu extrahieren. Auch hier ist jedoch eine genaue Bewertung notwendig: Ging es darum, ob jemand den Schlüssel in Erfahrung bringen kann (→ effektiv unmöglich) oder dass der Schlüssel zB zur Signierung eingesetzt wurde (→ erfordert nur die Möglichkeit, eine Signatur im Modul auszulösen, zB Kenntnis von Passwort/PIN etc).

Sicherheitsmaßnahmen können auch Auskunft darüber geben, wie leicht es war, das System während der Datenerstellung zu manipulieren (und anschließend die Veränderungen rückgängig zu machen), sodass gar keine späteren Modifikationen notwendig waren, da von Beginn an falsche Daten aufgezeichnet wurden.

X. Redundanz bzw unterstützende Daten

Daten existieren oft mehrfach: Bei korrekter Erhebung sollten alle übereinstimmen bzw zueinander passen. Existieren Diskrepanzen, so reduziert dies den Beweiswert, sofern die Unterschiede nicht, zB durch deren Erzeugung oder die Methode der Datenerhebung (siehe oben) erklärt werden können. Es ist daher sinnvoll, Daten «mehrfach» sicherzustellen, sofern es sich um Daten handelt aus:

- Unterschiedlichen Quellen: Dies liegt dann vor, wenn ein Datum mehrfach bzw auf unterschiedliche Weise aufgezeichnet wird, zB ein physikalischer Wert, der mit zwei Sensoren gemessen wird. (Fast) identische Werte weisen dann darauf hin, dass nicht einer der Sensoren defekt war. Automatische Aufzeichnung und zusätzlich manuelle Eingabe sind eine weitere Möglichkeit.
- Unterschiedlichen Systemen: Mehrere Systeme unter Kontrolle verschiedener Personen sollten identische bzw ähnliche Daten speichern, sofern es sich um denselben Wert handelt. Eine Nachricht sollte beim Absenden etwa genauso lang sein wie bei ihrem Empfang. Für eine unbemerkte Veränderung der Daten wäre dann eine Manipulatation auf beiden Systemen erforderlich, was jedenfalls schwieriger und aufwändiger ist als in einem einzelnen System bzw bei einheitlicher Kontrolle derselben.
- Kopien: Handelt es sich um ein komplexes System, so werden Daten sehr oft auf unterschiedliche Art bzw an mehreren Stellen gespeichert, zB in einer Systemdatei

(System-Log) und einer Programmdatei (Application-Log), einer Datenbank, und indirekt als daraus abgeleitete Werte. Diese können unterschiedlich stark gesichert sein, aber ein Angreifer hätte jedenfalls darauf zu achten, alle synchron zu ändern, wenn Manipulationen vorgenommen werden sollen.

Auch wenn nicht alle Daten vollständig identisch sind, zB wenn sie für unterschiedliche Zwecke gespeichert wurden und sich daher nur teilweise überschneiden, so bedeutet eine Übereinstimmung doch einen höheren Beweiswert. Dies umso mehr, da es bei unterschiedlichen Zwecken deutlich schwerer ist, alle «Duplikate» zu finden (da eben nicht identisch bzw bei daraus berechneten Werten eine passende Berechnung separat nachträglich durchzuführen) und entsprechend zu manipulieren.

Dies gilt offensichtlich auch umgekehrt: Wenn Redundanz bzw unterstützende Daten eigentlich vorhanden sein sollten (zB automatische Erzeugung), jedoch fehlen. Dies reduziert den Beweiswert, solange nicht nachvollziehbar erklärt wird, warum diese erwarteten Daten fehlen[11].

XI. Wiederholungsmöglichkeit

Ein grundlegendes Merkmal des Beweiswertes ist, ob die Erhebung der Daten, ebenso wie deren Auswertung, durch Dritte und/oder auf eine andere Weise wiederholt werden kann. Ist dies unmöglich, bedeutet dies gleichzeitig, dass eine vollständige Überprüfung unmöglich ist, was den Beweiswert reduziert: Es ist den Ergebnissen einfach zu trauen (oder nicht); siehe jedoch oben für Methoden, dies objektiv nachvollziehbar zu machen. Gründe für eine fehlende Wiederholungsmöglichkeit werden in anderen Abschnitten näher dargestellt. Weitere Gründe hiefür können sein, dass Unternehmen keine Grunddaten herausgeben (können) oder diese als Betriebsgeheimnis gelten, bzw

[11] Siehe das Beispiel in SCHMID (Hrsg.): Studienarbeit von cand. Wirtsch. Inf. Hanno Baur: Zur «Beweiskraft informationstechnologischer Expertise» (Stand 6/2010), CyLaw-Report XXXV, 25: Fehlen alle Logdateien weil diese (nachgewiesenermaßen zB aus Datenschutzgründen) regelmäßig gelöscht werden, beinträchtigt dies den Beweiswert nicht. Fehlt jedoch ausschließlich die eine relevante Logdatei, reduziert dies den Beweiswert der Daten, welche durch die fehlende Datei bestätigt werden könnten.

dass ihre Beschaffung, zB wegen Internationalität, nicht (rechtzeitig) möglich ist. Entsprechende Argumente sind jedoch gerade bei Daten besonders kritisch zu sehen, da eine Übermittlung technisch trivial ist (verschlüsselte Übermittlung über das Internet bis hin zu Festplatte per Kurier) und, aufgrund der Einfachheit von Kopien, der Herausgebende keinen Verlust erleidet. Relevant sind daher insb Geheimhaltung (dh es werden nur die Ergebnisse der Auswertung mitgeteilt, aber nicht die Grunddaten) bzw rechtliche Hindernisse (insb bei Internationalität[12]).

XII. Determinismus

Daten sind selten für sich allein aussagekräftig, sondern müssen zuvor ausgewertet werden. Bei dieser Auswertung können auch nicht-deterministische Verfahren zur Anwendung kommen, zB Varianten von machine-learning oder probabilistische Verfahren. Dem sind Methoden gleichzusetzen, welche zwar deterministisch sind (dh bei jeder Ausführung das gleiche Ergebnis produzieren), welche aber keine Erklärung für das Resultat liefern können. Derartige Auswertungen besitzen nur einen extrem reduzierten Beweiswert, da ihre Grenzen und Limitierungen nicht exakt bekannt sind, und sie (wie diverse Beispiele der Bilderkennung zeigen) auch völlig falsch liegen können[13] bzw uU unsichtbar zu manipulieren sind[14]. Dies bedeutet nicht, dass die Auswertung immer bis ins letzte Detail zu erklären ist, aber dass dies zumindest auf Anfrage hin gegenüber Experten möglich sein muss.

---

12 Da innerhalb der EU vermutlich der weltweit stärkste Schutz der Privatsphäre existiert (DSGVO), sollte dies nicht als Argument für eine Verweigerung der Datenherausgabe ausreichen.

13 Beispiel: Bilder von Hunden auf Gras sowie von Huskies auf Schnee werden zum Traineren des Algorithmus zu deren Unterscheidung verwendet. Dies bedeutet jedoch, dass viele Algorithmen tatsächlich den Hintergrund bewerten/erkennen und nicht das Tier. RIBEIRO/SINGH/GUESTRIN: «Why Should I Trust You?»: Explaining the Predictions of Any Classifier, 2016, CoRR, https://arxiv.org/abs/1602.04938 (16.9.2019).

14 Siehe ATHALYE/ENGSTROM/ILYAS/KWOK: Synthesizing Robust Adversarial Examples, ICML 2018, https://arxiv.org/abs/1707.07397 (16.9.2019) mit diversen Beispielen: Altar der als Afrikanischer Elefant erkannt wird, Holzfass als Guillotine, Clownfisch als Panflöte etc.

Indirekt bedeutet dies weiters, dass eine Datenauswertung nicht auf Vermutungen beruhen darf, sondern durch Experimente zu belegen ist: «Wird eine Datei aus einem Archiv entpackt, ändern sich ihre Zeitstempel nicht». Dies klingt logisch, aber ist es auch tatsächlich so? Die einfachste Lösung ist, dies mit dem verwendeten Programm auszuprobieren und die Zeitstempel vorher und nachher zu untersuchen, idealerweise in verschiedenen Varianten (zB auf einer lokalen Festplatte, einem Netzwerklaufwerk, einem Online-Speicherdienst). Erst nach einer nachvollziehbaren Erklärung warum das Ergebnis so sein soll, verbunden mit einer tatsächlichen Überprüfung der vorhergesagten Ergebnisse, kann darauf vertaut werden, dass ein Auswertungsverfahren zuverlässig ist[15].

XIII. Anzahl der Auswertungsschritte

Ebenso einen Beitrag zum Beweiswert von Daten liefert die Anzahl der erforderlichen Auswertungsschritte: Wie viele Interpretationen sind erforderlich, um zum Endergebnis zu kommen? Ein Beispiel: In einem Log ist vermerkt, dass die fragliche E-Mail versendet wurde. Dies ist ein guter Beweis dafür, dass dies tatsächlich erfolgte, denn aus dem Eintrag ergibt sich unmittelbar diese Schlussfolgerung. Wenn wir jedoch lediglich die Länge der verschickten E-Mail kennen und diese mit allen E-Mails auf dem System vergleichen, damit zu einer bestimmten E-Mail kommen, diese zu einem dadurch fesgestellten Zeitpunkt verwendet wurde, und wir dann feststellen, welcher Benutzer zu diesem Zeitpunkt angemeldet war, so nennt dies den menschlichen Absender der E-Mail. Der Wert dieser Schlusskette ist jedoch deutlich geringer, da sich bei jedem Schritt Fehler einschleichen können und es uU andere Erklärungen gibt.

Wichtig für den Beweiswert sind daher:

– Wie viele Indirektionsschritte waren erforderlich? Je mehr Zwischenschritte erforderlich sind, desto unsicherer ist das Ergebnis.

[15] Siehe dazu auch den Daubert-Standard. Dieser stammt ursprünglich aus den USA, aber dessen Inhalte, insb dass eine Methode wissenschaftlich überprüft worden sein muss um für Beweise akzeptabel zu sein, haben sich international verbreitet.

- Welche Schritte sind technische Konsequenzen bzw logische Schlussfolgerungen? Technische Konsequenzen sind durch die Hard- oder Software fix vorgegeben und wurden durch deren Untersuchung bestätigt (und sind leicht durch Experimente überprüfbar), während logische Schlussfolgerungen darauf aufbauen, was Programmierer normalerweise machen würden, was die übliche Folge einer bestimmten Aktion ist etc – was jedoch nicht mit den konkreten Tatsachen übereinstimmen muss.
- Erfolgt die Auswertung vor- oder rückwärts? Tritt ein Problem ein, so wird ein Logeintrag geschrieben. Dies bedeutet jedoch nicht, dass jeder Logeintrag auf ein Problem zurückzuführen ist. Es macht daher einen Unterschied ob geprüft wird «Wir vermuten, ein Problem trat auf. Sind die Daten mit dieser Vermutung konsistent?» oder «Wir haben einen Logeintrag. Was war der Grund hierfür?».
- Wie wurde über die einzelnen Schritte entschieden? Wurde einfach etwas ausprobiert und zufällig ein Ergebnis gefunden oder gab es einen bestimmten Grund, warum genau dieser Schritt durchgeführt wurde? Idealerweise sollte immer eine Vermutung aufgestellt werden, welche dann durch den Schritt versucht wird zu falsifizieren (wissenschaftliche Methode). In der Praxis häufig verbreitet ist jedoch auch eine positive Suche: Ein bestimmter Sachverhalt wird vermutet und dann wird geprüft, ob dieser sich durch die vorhandenen (oder noch zu erhebenden) Daten nachweisen lässt. Dies bedeutet jedoch, dass die Auswahl der Prämisse essentiell ist. Es sollte allerdings für jeden Schritt immer einen Grund geben, warum dieser erfolgte (und warum andere mögliche oder erfolgversprechende andere nicht/später) und was man davon erhoffte (zusammen mit dem tatsächlichen Ergebnis).

XIV. Alternative Darstellungsmöglichkeiten

Werden die Ergebnisse mit einem Programm für Menschen dargestellt (zB Bilder, aber auch grafische Darstellungen von Zahlungsströmen, Kommunikationsnetzwerke, bis hin zu Texten oder Tabellen etc), so hängt ihr Beweiswert auch davon ab, ob es alternative Darstellungsmöglichkeiten gibt[16]. Kann zB die Bild-Datei auch mit einem

[16] Siehe MOMSEN/HERCHER, Digitale Beweismittel im Strafprozess. Eignung, Gewinnung, Verwertung, Revisibilität (2013), 176, https://www.strafverteidigervereinigungen.org/Material/Themen/Technik%20%26%20Ueberwachung/37_momsen.html (16.9.2019).

anderen Programm dargestellt werden oder nur mit diesem einen – von dem nicht unbedingt genau bekannt ist, wie/ob es funktioniert und welche Aspekte es auslässt (zB Metadaten)? Bei nur einer Option ist die Wahrscheinlichkeit, dass diese Software fehlerhaft ist (oder dass der Fehler von einem Angreifer ausgenützt werden kann!) deutlich höher, als wenn mehrere Programme eine identische Visualisierung erzeugen. Dies kann auch als Variante der Wiederholungsmöglichkeit angesehen werden, wobei hier nicht Datenerhebung, sondern Auswertung im Vordergrund steht.

Hierzu gehört auch die Art der Interpretation: Handelt es sich lediglich um eine «Umkodierung» von Binärdaten in grafische Pixel/Texte usw oder wird eine aktive Bearbeitung vorgenommen oder ist möglich, beispielsweise durch eingebettete Makros? Dies ist das sogenannte «Secure Viewer» Problem elektronischer Signaturen, weshalb bei diesen nur die Signatur von reinen Schwarz-Weiß-Bildern empfohlen wird, um jede Manipulation (zB unsichtbare Darstellung gelben Textes auf einem Bildschirm, dessen Gelb-Leitung unterbrochen wurde, oder Makros welche je nach Datum andere Inhalte anzeigen) auszuschließen.

XV. Angreifer-Charakterisierung

Sofern sich ein möglicher Angreifer charakterisieren lässt, ist auch dies ein Maß für den Wert von Beweisen: Welche Kenntnisse und Ressourcen stehen diesem zur Verfügung und wäre es mit diesen möglich gewesen, die Daten zu verfälschen? Wer hatte welche Möglichkeiten (zB Zugriff auf den Server, die Kommunikationsverbindung, das Programm, die Datei)? Dies ist als ein Negativ-Kriterium zu sehen, dh wenn die Möglichkeiten des Angreifers nicht ausreichen, sind die Daten «normal» zu bewerten, während entsprechende Optionen den Beweiswert reduzieren. Praktisch ist dieses Kriterium aufgrund der Voraussetzungen eher von geringerer Bedeutung, da potentielle Angreifer nicht unbedingt feststehen, deren Ressourcen schwer zu bestimmen sind, und gerade im IT-Bereich auch Laien Software bedienen können, die uU sehr komplexe Aktionen ausführt – sofern sie Zugang zu dieser erreichen.

Es ist daher eher zu empfehlen, die Charakterisierung als Anregung zu sehen, nach Spuren für entsprechende Manipulationen zu suchen, und den Beweiswert nicht nur aufgrund potentieller Möglichkeiten zu reduzieren. Stattdessen kann zB eine Darstellung

erfolgen, welche Aktionen mit welchen Ressourcen notwendig gewesen wären, um die Daten (im Hinblick auf die durchgeführten Überprüfungen) zu fälschen.

Als Beispiel kann ein Windows-Server dienen: Wurde er falsch konfiguriert? Falls aktuell kein Fehler festgestellt werden kann – Wie hätte ein Fehler vertuscht werden können, zB durch Neuinstallation? Im Falle einer Neuinstallation wäre es allerdings notwendig gewesen, eine falsche Historie in den Logdateien zu erzeugen – sofern diese überprüft wurden. Hierzu ist es auch nötig, das Einspielen von Sicherheits-Updates zu exakt dem Zeitpunkt zu simulieren, an dem diese von Microsoft veröffentlicht wurden – wenn dieses Faktum geprüft wurde. Um dies zu fälschen, müsste Zugang zu den (alten) Updates bestehen, was ein vorausschauendes Archivieren oder ein Hacken der Microsoft-Server erfordert – oder ein Verändern ausnahmslos aller betroffenen Zeitstempel. Dies kann noch weitergeführt werden und ermöglicht es so, den Wert der Beweise fast beliebig zu verstärken: Es werden weitere Punkte untersucht, die ebenso hätten modfiziert werden müssen, was progressiv schwieriger durchzuführen und zu wissen ist. Voraussetzung ist allerdings, dass es überhaupt Möglichkeiten hierfür gibt, dh Redundanz oder unterstützende Daten (siehe schon oben).

D. ANWENDUNGSBEISPIEL

Eine konkrete Darlegung der erläuterten Kriterien kann an Hand der Entscheidung OLG Jena, 28.11.2018, 2 U 524/17[17] erfolgen. Hierbei ging es um einen Unterlassungsanspruch nach UWG, basierend auf einer «fehlerhaften» eBay-Seite. Von dieser Seite existierte allerdings nur mehr ein Ausdruck auf Papier.

Zuerst wurde festgestellt, dass ein Ausdruck eines Screenshots kein elektronisches Dokument sei, auch wenn es von einem solchen abstammt. Inhaltlich wäre die Beurteilung jedoch auch bei elektronischer Vorlage auf einem Datenträger identisch. Dies ist problematisch zu sehen, da bei einer Datei zusätzliche Metadaten vorliegen, was bei einem Papierausdruck nicht der Fall ist; ebenso kann der Webseiten-Text untersucht werden, der weitere Daten beinhaltet (zB JavaScript-Programme, Kommentare, ausgeblendete

[17] Siehe auch Anmerkung von ELZER, OLG Jena: Bildschirmfoto als Beweismittel, https://rsw.beck.de/aktuell/meldung/olg-jena-bildschirmfoto-als-beweismittel (16.9.2019)

Elemente etc – welche alle bei einem Ausdruck vollständig fehlen). Die Datei besitzt daher deutlich mehr «Inhalte» als der Ausdruck – und damit einen höheren Beweiswert (egal in welche Richtung, positiv wie auch negativ).

Zuerst wurde im Verfahren festgestellt, dass kein erhöhter Beweiswert vorliegt, da weder eine qualifizierte Signatur (siehe Abschnitt C.III) noch ein elektronischer Zeitstempel (C.IV) vorhanden ist. Es ist nicht einmal irgendeine Datumsangabe vorhanden, weder auf dem Ausdruck selbst (=auf dem Bildschirm) noch zB als manuelle Hinzufügung auf dem Papier. Es fehlt daher jede Dokumentation der Sicherstellung (C.VIII). Der Hinweis «Möglicherweise kein Versand nach Kiribati» auf der Webseite deutet auf einen Abruf vom Ausland aus hin, was jedoch nicht in Betracht kommt; ebenso passt der Kontakt-Link nicht zum Verkäufernamen (C.X Redundanz bzw unterstützende Daten sind widerprüchlich). Dies kann uU auf vorherige andere Nutzung und einen Cache zurückzuführen sein (C.VI – Darlegung der Datenerzeugung). Ein Teil der AGB ist unvollständig abgebildet: Eine Umrahmung blieb – unerklärterweise – teilweise leer (C.II – Es ist unklar, ob dies bei eBay tatsächlich jemals derart angezeigt wird oder ob es eine nachträgliche Veränderung war). Fehlende Informationen zur Änderungshistorie der Webseite sind jedoch kein Problem, da die festgestellten Änderungen erst nach dem Zeitpunkt der Erstellung des Screenshots erfolgten (C.X, diesmal positive Redundanz). Weiters liegt zwar eine eidesstattliche Erklärung vor, doch bezieht sich diese nur darauf, dass eine Abmahnung erfolgte, aber zB nicht warum oder aufgrund welcher Beweise (C.VII – Keine Chain of Custody). Eine Angreifer-Charakterisierung (C.XV) erfolgte nur sehr indirekt, indem darauf hingweisen wurde, dass der Gegner nicht pauschal auf die einfachen Möglichkeiten der Fälschung von Screenshots hinwies, sondern konkrete unstimmige Details (siehe oben) anführte.

Aufgrund dieser Vielzahl an Problemen gelangte das Gericht zur Auffassung, dass der Screenshot nicht zum Beweis der Vorwürfe geeignet sei.

Wie hätte ein solcher Vorgang besser bewiesen werden können? Einserseits dadurch, dass der Screenshot elektronisch gespeichert vorgelegt wird (mehr «Original»; siehe oben), andererseits durch weitere Maßnahmen:

– Darlegung, wer den Screenshot wann auf welchem Gerät erstellt hat (C.VIII), wie dabei vorgegangen wurde (welcher Browser, Cache löschen…), wo er gespeichert

war (C.II), und wer darauf Zugriff hatte (C.VII). Idealerweise erfolgt auch eine unmittelbare Dokumentation dieser Daten, zB durch Niederschrift, oder auch als elektronisches Dokument.
- Signierung (C.III), Zeitstempel (C.IV) und/oder Hashwert (C.V) der Screenshot-Datei, um Zeitpunkt der Erstellung, Ersteller bzw Unverändertheit seit damals nachzuweisen.
- Wiederholung des Vorgangs an mehreren Tagen bzw auf mehreren Geräten (C.X) und/oder mit unterschiedlichen Browsern (C.XIV).
- Vollständige Speicherung der Webseite mittels des Browsers, sodass eine (teilweise) Besichtigung auch später möglich ist (C.XI).
- Besondere Sicherheitsmaßnahmen zur Aufbewahrung der Daten (C.IX) sind zB durch Speichern in einem Softwaresystem für Anwaltskanzleien möglich.

Im konkreten Beispiel für die Parteien unmöglich sind die sichere Aufbewahrung bis zur Sicherstellung (C.I) – diese könnte jedoch über eine Darstellung der Sicherheitsmaßnahmen von eBay erfolgen. Die Darlegung der Datenerzeugung (C.VI) kann entweder durch Mitarbeiter von eBay erfolgen oder mittels Testens durch Dritte. Determinismus (C.XII) und die Anzahl der Auswertungsschritte (C.XIII) spielen hier keine Rolle, da die «Auswertung» in der unmittelbaren Besichtigung des Bildes bzw der Webseite besteht. Die Charakterisierung von Angreifern (C.XV) würde hier eher technisch zu erfolgen haben, indem dargestellt wird, welche Möglichkeiten für Manipulationen bestehen und wie aufwändig diese wären – da die Gegenpartei als ein «Angreifer» praktisch feststeht. Zusätzlich sollte diskutiert werden, ob Dritte ein Angebot im Namen (oder unter ähnlichem Account etc) hätten einstellen können.

E. ZUSAMMENFASSUNG

Eine Vielzahl an Kriterien wurde präsentiert (und anhand eines Beispieles kurz praktisch dargestellt), mit deren Hilfe der Beweiswert elektronischer Daten beurteilt werden kann. Nich in jedem Fall sind alle möglich und praktisch nie werden alle gleichzeitig vorhanden sein. Aber umgekehrt sollte es keinen Fall geben, in dem kein einziges Kriterium erfüllt wird – da Digitaldaten einfach und unbemerkt fälschbar sind. Im weiteren Sinne ist es daher sowohl bei der Gestaltung von Computersystemen wie auch bei der

Auswertung elektronischer Beweise erforderlich, auf diese Kriterien zu achten. Dies kann in einfachster Weise erfolgen (Logdateien, auf welche nur ein Administrator zugreifen kann; Stichwortartige Beschreibung des Auswertungsvorgangs), bis hin zu komplexen System (Logdateien auf dritten Systemen, mit elektronischen Signaturen und Zeitstempeln, regelmäßig auf WORM-Medien gesichert; Videomitschnitt und Bildschirmaufzeichnung des Auswertungsvorgangs). Damit kann eine zu hohe Technikgläubigkeit vermieden werden, ebenso wie ein komplettes Misstrauen gegenüber Ergebnissen von Computern. Ebenso erleichtert eine Aufschlüsselung der Gründe für den Beweiswert die Nachvollziehbarkeit einer Entscheidung für betroffene Parteien, und erleichtert bzw ermöglicht überhaupt erst eine erfolgreiche Bekämpfung des Ergebnisses bei Fehlern.