ARTICLE:

DIGITAL DATA AS HEARSAY

By **Steven W Teppler**

This article examines the proposition that all digital data is hearsay in legal proceedings within the United States of America. This article also analyzes the inadequacy, if not outright failure, of the current approaches to dealing with the hearsay exception used to offer computer-generated information into evidence. The author proposes that until the Federal Rules of Evidence are revised to reflect the ephemeral nature of digital evidence, such evidence should be considered hearsay and deemed inadmissible unless a hearsay exclusionary exception is successfully asserted. It is further proposed that if admissibility of digital evidence is sought pursuant to a hearsay exception, such evidence should be made subject to heightened reliability requirements.

Summary

Subject to certain exceptions not pertinent to this discussion (unfair prejudice, confusion of the issues and such like), all relevant evidence is generally considered admissible once a proper foundation has been laid pursuant to Rule 901 of The Federal Rules of Evidence (F.R.E.).

Once authenticated, the F.R.E. provides that hearsay evidence is generally inadmissible unless it falls under an exception. The F.R.E. defines hearsay as 'a statement, other than one made by the declarant, while testifying at the trial or hearing, offered in evidence to prove the matter asserted.'[1] The two hearsay exceptions generally applicable to digital data embrace notions of trustworthiness or reliability, and are commonly known as the Business Records Exception found in F.R.E. 803(6) and the residual hearsay exception set forth in F.R.E. Rule 807. These exceptions to the hearsay rule generally require a showing of reliability (or in the case of the business records exception, the lack of a challenge in relation to reliability).

Since digital data is inherently ephemeral and therefore not demonstrably reliable, there is a low bar to attaining admissibility by operation of a hearsay exception, that is by a literal adherence to current requirements, which tends to reflect a service that is proffered but not performed.[2] Accordingly, these approaches fall short of their intended objective because the F.R.E. (and most judicial authority) does not properly address reliability issues arising from the inherently ephemeral nature of digital data. Although this shortcoming has been documented since at least as early as the 1970's, the Federal Rules of Evidence have not been amended to demand of a party seeking to admit digital data that degree of reliability properly reflective of the frailty of digital evidence.

Until the Federal Rules of Evidence are revised to directly address the ephemeral nature of this new specie of evidence, an argument may be made that all digital data is hearsay, and that an affirmative showing of reliability must be demonstrated if admissibility is to be sought under an exception to the hearsay rule.

Although two United States Circuit Courts of Appeal have rejected the comprehensive application of the hearsay rule to all digital data, it is contended that well-established authority from at least one prominent Federal Circuit provides the constitutional basis for deeming all digital data as hearsay. Moreover, and despite the mostly orthogonal arguments made in opposition, the undisputed nature of digital data itself compels the conclusion that all digital data is hearsay. Finally, this article examines the potential implications of the application of the hearsay evidence rule to digital evidence used in both the criminal and civil context.

Until the Federal Rules of Evidence are revised to address information in digital format, the 'trustworthiness' standards set forth in F.R.E. Rule 803(6) and 807 hearsay exceptions, together with recent judicial authority, provide the proper standards for determination of authentication based on a positive showing of 'reliability', and discuss how such an emerging concept of 'reliability' should be used as the primary condition for admissibility of digital data sought be offered as evidence.

[1] *Fed. R. Evid. Rule 801(c).*

[2] *For instance, the well laid out arguments in George Paul, Foundations of Digital Evidence, (ABA, 2008) 131-149; also Stephen Mason, Electronic Evidence: Disclosure, Discovery & Admissibility (LexisNexis Butterworths, 2007) Chapter 4.*

Background

Digital data is also known as 'Electronically Stored Information' (ESI) for the purposes of the Federal Rules of Civil Procedure. Accordingly, for purposes of uniformity, the terms 'computer generated information,' ESI, and 'digital data,' in any format and however stored, are used interchangeably in this article.

The vast majority of information currently generated is digital in nature.[3] It follows that the vast majority of information offered as evidence will also be digital in nature, and this trend is reflected in the December 2006 amendments to the Federal Rules of Civil Procedure.[4] Digital data is inherently malleable, or ephemeral.[5] The inherently ephemeral nature of computer-generated data creates new issues that have a significant and detrimental effect on reliability, authentication and ultimately on the issue of admissibility. To date, these issues remain largely ignored by both the bench and the bar, and directed into unsuitable definitions or relegated to obsolescent analyses. The reason for this ignorance or misapprehension is probably the result of a basic misunderstanding of the nature of both computer-generated information and the variable nature of the computing environment by which such information is generated. The result of this general misunderstanding can be seen in the current mixture of judicial approaches to the admissibility of digital evidence.[6]

It should be made clear at the outset that the F.R.E. never refers to or directly addresses digital evidence. The Federal Rules of Evidence predate by decades the 2006 electronic discovery amendments to the Federal Rules of Civil Procedure, and so it is not surprising that the F.R.E. makes no mention of ESI. Despite the approach of the thirtieth anniversary of near ubiquity, however, the term 'computer' is notably missing from the F.R.E. Moreover, even the authentication provisions of Rule 901 refer generally to the accuracy of a 'process or system' in producing a result[7] without indicating whether the process or system is a computer, or whether the result is computer generated information.

The term 'data compilation' makes one of its rare appearances in Article VIII F.R.E., and is expressly included as a record of a regularly conducted activity under what is commonly referred to as the 'Business Records' exception to the hearsay rule. Judicial authority generally supports the proposition that computer generated information is a subset of the umbrella term 'data compilation' for purposes of analysis under the business records exception:[8]

> '[Defendant] does not dispute the well established proposition that "computer data compilations may constitute business records for purposes of Rule 803(6), ... and may be admitted at trial if a proper foundation is established." *United States v. Croft*, 750 F.2d 1354, 1364 (7th Cir.1984) (citing *United States v. Young Brothers, Inc.*, 728 F.2d 682, 694 (5th Cir.), cert. denied, 469 U.S. 881, 105 S.Ct. 246, 83 L.Ed.2d 184 (1984)).' *U.S. v. Hayes 861* F.2d 1225, *1228 (10th Cir. 1988).

> 'A business record may include data stored electronically on computers and later printed out for presentation in court, so long as the original computer data compilation was prepared pursuant to a business duty in accordance with regular business practice.' *Potamkin Cadillac Corp. v. B.R.I. Coverage Corp.*, 38 F.3d 627, 632 (2d Cir.1994). *Health Alliance Network, Inc. v. Continental Cas. Co.* 245 F.R.D. 121, *129 (S.D.N.Y., 2007).

A second appearance of the term 'data compilation' appears in F.R.E. Rule 901(b), but curiously, only from within the context of authenticating a 'Public Record' or 'Ancient Documents.'[9] A final F.R.E. reference to 'data compilation' is found in Rule 1001, which generally requires that an original is required to prove the content of a writing, recording or photograph.[10] 'Writing and recordings' are defined in pertinent part to include 'letters, words or numbers, or their equivalent, set down by handwriting, typewriting, printing, photostatting, photographing, magnetic impulse, mechanical or electronic recording, or other forms of data compilation.'[11]

[3] *Peter Lyman and Hal R. Varian, How Much Information? at http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/.*

[4] *Fed. R. Civ. P. Rules 16(b)(5); 26(a)(1)(B), 34(a), (b); 37; 45(a)(1)(C).*

[5] *Bruce H. Nearon, Jon Stanley, Steven W. Teppler and Joseph Burton, 'Life After Sarbanes-Oxley: The Merger of Information Security and Accountability', Jurimetrics Journal, Vol. 45, No. 4, Summer 2005, 379, 387.*

[6] *Compare St. Clair v. Johnny's Oyster & Shrimp, Inc., 76 F. Supp. 2d 773 (S.D. Tex, 1999), digital data taken from the internet described as 'voodoo information' with D & H Auto Parts, Inc. v. Ford Marketing Corp., 57 F.R.D. 548, *552 (E.D.N.Y.1973) 'In relying upon data processing by a machine, there should be no more necessity for oral testimony concerning the reliability of the machine operations than that of the manual procedure supplanted, whether it be bookkeeping, order preparation, or mathematical computation'.*

[7] *Fed. R. Evid. Rule 901(b)9.*

[8] *Fed. R. Evid. Rule 803(6). Curiously, the F.R.E. Rule 807 residual hearsay exception rule makes no mention of data compilations.*

[9] *Fed. R. Evid. Rule 901(b) 7, 8.*

[10] *Fed. R. Evid. Rule 1002.*

[11] *Fed. R. Evid. Rule 1002.*

F.R.E. Article X: digital data as 'source' data

With the vast majority of all information generated today originating as electronic or computer-generated data, the inescapable conclusion to be drawn is that digital data will become the main source of evidence used in modern litigation. Despite this massive shift in specie of evidence, there has been a relative paucity of judicial authority, and certainly no emergent majority view dealing with the vagaries inherent in respect of computer-generated information and the directions for its admissibility into evidence.

As early as the late 1970's, judges have written about the need to amend the Federal Rules of Evidence to reflect the unique evidentiary issues presented by the inherently ephemeral nature of computer-generated data.[12] Unfortunately, the current Federal Rules of Evidence do not directly address the unique authentication or admissibility issues raised by this massive shift from evidence in physical format to evidence in digital format.[13] It is hoped that the recent amendments to the Federal Rules of Civil Procedure will accelerate corresponding amendments to the Federal Rules of Evidence.[14] Until such a time, attorneys and judges will continue to deal with inconsistent, and at times contradictory evaluative admissibility frameworks for digital evidence.

Adding to this unwieldy and inconsistent framework is a general lack of understanding by attorneys and judges of what constitutes computer-generated information, and what constitutes 'source information'. The source data of all computer-generated information is binary in nature, and the data processed, viewed, printed out, or stored is composed of ordered sets of zeroes and ones.[15] This binary data are acted upon (processed) by other ordered sets of binary data comprising the operating system and other data processing software applications to produce what are commonly referred to as a data files.[16] 'Source' data are, therefore always comprised of zeroes and ones that are then processed, or rendered, by the operating system and various applications to produce files. These files are generally further processed by other applications to produce images that can be viewed on a screen, or can be viewed by printing the data out on to paper.[17] Nevertheless, the source data for either an image viewed on a screen or a computer-generated paper printout are the binaries, or the ordered sets of zeroes and ones, that comprise the true, or source data, used to produce the screen image or paper printout. The data (or information) actually read or perceived by a human reader (or members of a jury) should therefore be considered the last 'view' in a set of a 'views of views' and not the 'source' or origination data.[18] In other words, while a person might read, hear or see computer-generated data, it is impossible to read, hear, or see source computer-generated source or origination data. In order to perceive source computer data as native data, it is necessary to interpret the language in which that data is written (such as 'C' or 'Visual Basic'). In order to interpret the language in which data is written, it is necessary in turn to understand the language in which it is written. The ultimate aim in understanding or examining computer-generated information is to understand the assertions, or speech, of the computer programmers (all of which are human) who by object code or source code provide the instructions to computers to make conditional statements.

Admissibility generally

The procedural schema in the United States 'requires the parties to present trial evidence pursuant to rules that make it clear when proof has been formally proffered before it is introduced and then may be considered by the trier of fact in resolving fact issues.

[12] *2 McCormick on Evidence §294 (6th edition, 2006), Commonwealth v. Klinghoffer, 564 A.2d 1240 (Pa. 1989); in a noted 1976 dissent, Judge Van Graafeiland presciently pointed to the need to amend the rules of evidence to address the admissibility issues presented by computer-generated information; Perma Research and Development v. Singer Co., Van Graafeiland, J dissenting, 542 F.2d 111, 124-26 (2nd Cir.1976): It is unfortunate that more than three decades later, no such amendments have been adopted, and the current inconsistent approach to authentication and admissibility is the direct result of that failure to amend.*

[13] *2 McCormick on Evidence §294 (6th edition 2006).*

[14] *Fed. R. Civ. P. Rules 16(b)(5); 26(a)(1)(B), 34(a), (b); 37; 45(a)(1)(C).*

[15] *Bruce H. Nearon, Jon Stanley, Steven W. Teppler, and Joseph Burton, 'Life After Sarbanes-Oxley: The Merger of Information Security and Accountability'.*

[16] *Bruce H. Nearon, Jon Stanley, Steven W. Teppler, and Joseph Burton, 'Life After Sarbanes-Oxley: The Merger of Information Security and Accountability' at 388.*

[17] *Bruce H. Nearon, Jon Stanley, Steven W. Teppler, and Joseph Burton, 'Life After Sarbanes-Oxley: The Merger of Information Security and Accountability' at 388.*

[18] *There is much confusion as to the term 'original' as it applies to computer-generated data. The phrase 'first instantiation' (which implies 'origin') rather than 'original' is used with good reason, and exemplifies one of the challenges in adapting the application of the Federal Rules of Evidence to computer-generated information. The commonly used definition for original is incompatible with the concept of 'initial' 'first' or 'earliest' with 'only.' This definition has no inherent value in respect of digital evidence. 'Original' digital data files can be reproduced in exact bit for bit copies. Unlike paper 'originals' there may never be 'only one' original. Data files may in fact, be 'duplicate originals' created at different times. First instantiation, or origin, however, refers to the characteristics of the source of the data, the environment (including controls) and provenance of the initial creation of digital data. Thus, the adoption and substitution of the term, 'first instantiation' for 'original' is suggested as more appropriate. It should also be noted that the adoption of this term also permits a disambiguation of the term 'time' for digital data creation. While 'first instantiation' can have only one time reference as it relates to data creation, 'original' data can be created at many different times.*

The proponent needs to know how to introduce evidence, the opponent must know when to object, and the judge needs to know when to rule. The rules of practice concerning presentation of evidence, offers of proof, and objections all are designed to secure this result."[19] To this end, the F.R.E. provides the contextual groundwork (further interpreted by case law) in accordance with which counsel may offer evidence, or challenge, impeach, or rebut such evidence. The F.R.E., together with case law precedent, provides guidelines for a court in determining evidentiary rulings.

The provisions of the F.R.E. lend themselves to a flow chart of actions that must be taken by a party submitting digital evidence, and decisions to be made by a judge, before any such admission into evidence.

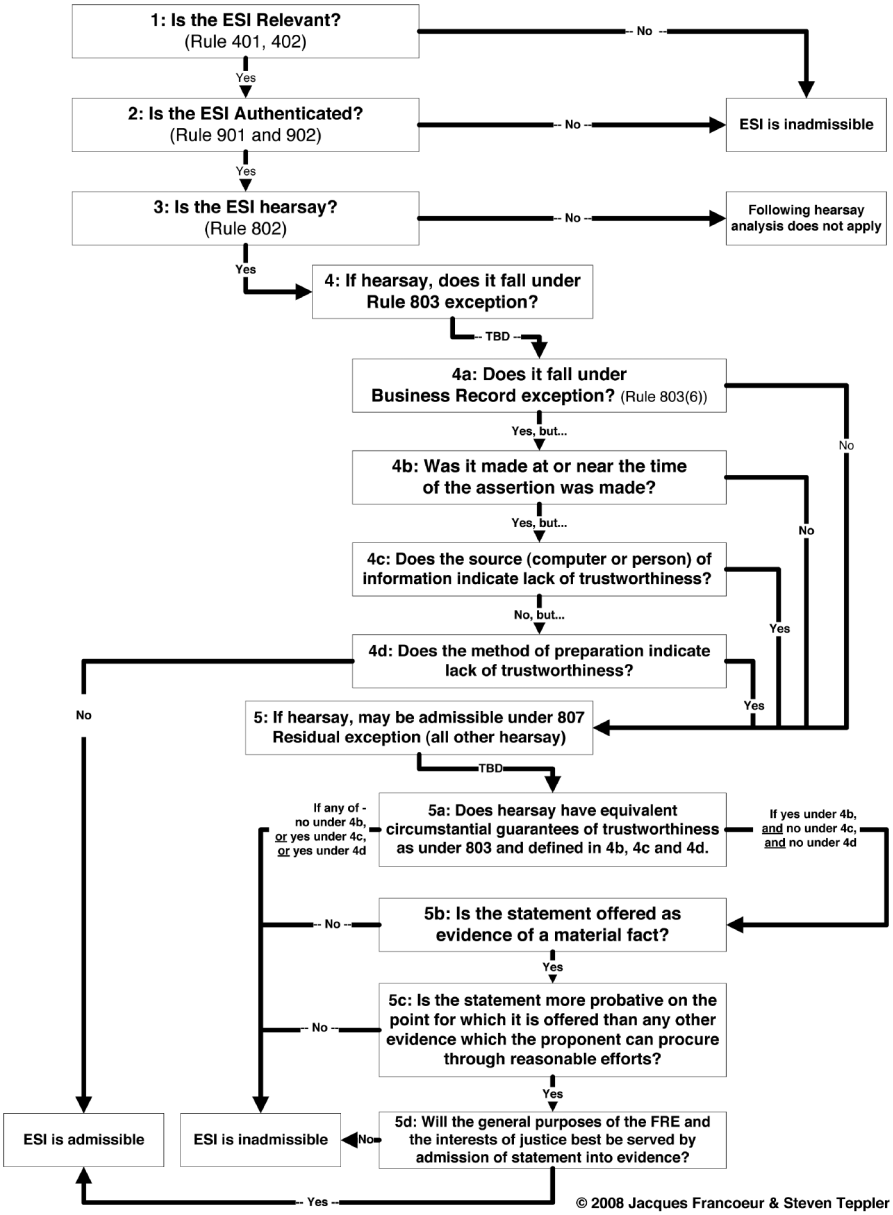

Figure 1: **ESI Admissibility Decision Tree**

[19] *1 McCormick on Evidence §51 (6th edition).*

The decision points of this flow chart are not fixed, and, subject to existing precedent, they provide a judge with the discretion to determine the admissibility of an item of evidence. Moreover, attorneys are aware that evidence, whether a thing, a record, a photograph, or testimony, is not admitted automatically into trial for scrutiny by a jury or judge. For reasons not pertinent to this discussion (such as privilege and prejudice) computer-generated information sought to be admitted (and otherwise admissible) may be excluded (or not permitted to be used at trial).[20] It should be understood that whilst any point reached along the F.R.E. flow chart discussed in this article may be favourably met, admission is not necessarily guaranteed by laying a proper foundation for authentication, or by the applicability of a hearsay exception.

Generally, therefore, all relevant evidence that is not privileged is admissible.[21] Once the initial hurdles of relevancy, privilege, prejudice, and such like have been met by a party offering the evidence, evidence must be authenticated by some means that satisfy the requirements of 901(a) F.R.E. that evidence 'is what it's proponent claims', or, as more commonly stated, that evidence 'is what it purports to be.'[22]

'Traditional' authentication

In order for evidence to be admissible, it must be identified or authenticated by extrinsic evidence in a manner that complies with F.R.E. Rule 901(a). Examples of methods of authentication are set forth in F.R.E. Rule 901(b). Such methods include the testimony of a witness or witnesses with knowledge, expert opinion, distinctive characteristics 'and the like,' or the efficacy of a particular method or process in producing a particular result.[23]

'Traditional' hearsay

While F.R.E. Rule 901 addresses authentication as a pre-condition to admissibility, F.R.E. 801 refers to the exclusion of hearsay evidence even if the party offering the evidence lays a proper foundation for authentication. Accordingly, Article VIII of the Federal Rules of Evidence effectively imposes a post-authentication requirement that a hearsay determination be made as a second pre-condition to admissibility. In order to be admissible, therefore, the evidence offered must first be authenticated or it is excluded. Even if authenticated, the evidence is excluded if deemed hearsay, unless the evidence falls under an exception to the hearsay rule.[24] The result, which is common to the operations of Rule 901 (authentication) and Rule 801 (hearsay rule), is to permit or preclude the admissibility of evidence at trial. The authentication provisions of F.R.E. Rule 901 and the hearsay exclusionary provisions of F.R.E. 801 may therefore be considered to occupy equal status in respect of admissibility. Finally, the Federal Rules of Evidence permit post-admission introduction before a jury of relevant evidence pertaining to 'weight or credibility.'[25]

Once authenticated, evidence may be deemed hearsay and inadmissible, or it may be deemed hearsay but falling under an exception to the hearsay exclusionary rule, in which case the evidence maintains its admissible status.

Hearsay, which is defined as an out-of-court statement made by a declarant at a trial or hearing, offered in evidence to prove the truth of the matter asserted, is generally not admissible.[26] There are, however, certain exceptions to the hearsay rule, where, under certain conditions, a court is permitted (but not required) to admit evidence that would otherwise constitute inadmissible hearsay.[27] One major exception is provided for by F.R.E. Rule 803(6), and is typically referred to as the Business Records Exception. The residual hearsay exception provided by F.R.E. Rule 807 permits admissibility of other types of hearsay based upon equivalent showings of circumstantial trustworthiness (such as those enumerated in Rules 803 and 804).[28]

Lack of uniformity in the judicial approach

There is no uniformity of approach in lower court decisions towards the issue of authentication and admissibility of computer-generated information offered as evidence for trial. The issue is complicated by the absence of any United States Supreme Court guidance as to whether digital data is inadmissible hearsay, or not. This lack of Supreme Court guidance has not escaped judicial notice.[29] Some judges appear to view all computer-generated information as hearsay, perhaps saved from exclusion by qualifying under the business records exception.[30] Other judges do not consider certain types of computer-generated data as hearsay,

[20] *Fed. R. Evid. 104 (a), 402 (relevance); 403 (prejudice, waste of time); 501 (privilege).*
[21] *Fed. R. Evid. §§401, 402.*
[22] *Fed. R. Evid. 901(a)*
[23] *Fed R. Evid. Rules §§901(b)(1), (3), (4) and (9).*
[24] *Fed. R. Evid. Rules 803, 807.*
[25] *Fed. R. Evid. Rule 104(e).*
[26] *Fed. R. Evid. Rule 801(c).*
[27] *Fed. R. Evid. Rule 803.*
[28] *Fed. R. Evid. Rule 807.*
[29] *Hawkins v. Cavalli 2006 WL 2724145 (N.D. Cal 2006).*
[30] *St. Clair v. Johnny's Oyster & Shrimp, Inc., 76 F. Supp. 2d 773 (S.D. Tex, 1999).*

By categorizing computer-generated information only as a subset of business records, judges have thus been able to avoid the central issues that are uniquely inherent to the authentication of computer-generated information.

and only require an F.R.E. Rule 901(b)(9) showing that the evidence is an accurate result from a system or process. In more recent decisions, however, judges have tended to consider a higher degree of evidential reliability, even for laying a foundation under F.R.E. Rule 901.[31]

The approach to electronic evidence posted on the internet by one judge, seems to indicate a marked disinclination to admit computer-generated information by labelling such data 'voodoo information' incapable of finding a basis for admission under the 'most liberal' interpretation of the hearsay exception rules.[32] The Court in the *St. Clair* case places much emphasis on its understanding (some of it presumably apocryphal) of computer-generated data:

> Anyone can put anything on the Internet. No web-site is monitored for accuracy and nothing contained therein is under oath or even subject to independent verification absent underlying documentation. Moreover, the Court holds no illusions that hackers can adulterate the content on any web-site from any location at any time. For these reasons, any evidence procured off the Internet is adequate for almost nothing, even under the most liberal interpretation of the hearsay exception rules found in Fed.R.Civ.P. 807.[33]

Other judicial authority accords a greater degree of presumptive trustworthiness or reliability to computer-generated information and is friendlier to its admission as evidence. The United States District Court for the Central District of California, relying on a 9th Circuit precedent, specifically eschews the *St. Clair* approach in favour of admitting print-outs of computer logs from a web site.[34] Although the court in *Perfect 10* acknowledges a reduced evidentiary standard in preliminary injunction motions, it nevertheless ruled certain print-outs of web pages admissible after considering the declaration of the party offering the print-outs together with the circumstantial authenticity of the content (internet domain address and the date of the print-outs).[35] Nevertheless, neither the *St. Claire* nor the *Perfect 10* decisions provide any substantive basis for concluding that computer generated information is, or is not, hearsay.

Efforts to date to categorize computer-generated information as hearsay or non-hearsay for purposes of admission into evidence at trial have been far from clear – indeed, contradictory at times. To date, the Supreme Court has not offered an opinion on the issue.[36] It is generally agreed that most judges simply (and conveniently) place digital data into the hearsay category as business records, and impose the requisite corresponding contemporaneity, because statements under the F.R.E Rule 803(6) business records exception to the hearsay rule.[37] By categorizing computer-generated information only as a subset of business records, judges have thus been able to avoid the central issues that are uniquely inherent to the authentication of computer-generated information. The Pennsylvania Supreme Court recently acknowledged that:

> Judicial decisions to date have largely skirted the edge of the problem because they have been concerned mainly with computerized records made in the regular course of business. [Citations omitted]. Routinely prepared records, admitted pursuant to business records acts such as 28 U.S.C. § 1732 are well recognized exceptions to the hearsay rule, because their regular use in the business of the company insures a high degree of accuracy. Proof of day-to-day business reliance upon computerized records should therefore make less onerous the

[31] *In re Vee Vinhnee, 336 B.R. 437 (9th Cir. BAP 2005); In re Vargas, 396 B.R. 511 (C.D. Cal. 2008); Lorraine v Markel American Life Ins Co, 241 F.R.D. 534 (D. Md. 2007; State v. Swinton, 847 A.2d 921 (Conn. 2004); Rodd v. Raritan Radiological Associates, 860 A. 2d 1003 (N.J. Sup. A.D. 2004).*

[32] *St. Clair v. Johnny's Oyster & Shrimp, Inc., 76 F. Supp. 2d 773 (S.D. Tex, 1999).*

[33] *St. Clair, 76 F. Supp. 2d 773 at 774-775. It is probable that the judge's reference to Rule 807 is actually to Federal Rules of Evidence, and meant to cite the Federal Rules of Evidence Rule 807, and not the Federal Rules of Civil Procedure.*

[34] *Perfect 10, Inc. v. Cybernet Ventures, Inc. 213 F.Supp.2d 1146 (C.D. Cal. 2002).*

[35] *Perfect 10, 213 F. Supp. 2d at 1154.*

[36] *Hawkins v. Cavalli 2006 WL 2724145 (N.D. Cal 2006).*

[37] *Hawkins v. Cavalli.*

> burden of laying a proper foundation for their admission. *United States v. Russo*, supra, 480 F.2d [1228] at 1239-40.[38]

The *Klinghoffer* court considered that computer-generated information that was not categorized as a business record as hearsay, but (unlike the court in *St. Clair*), admitted the evidence on the condition of meeting the 'circumstantial guarantees of trustworthiness' set forth in the residual hearsay provisions of F.R.E. Rule 807:

> Where, however, a computer is programmed to produce information specifically for purposes of litigation, an entirely different picture is presented. Its product, which is hearsay and conclusory, is not admissible under 28 U.S.C. §1732 or similar state statutes. [Citations omitted]. Under such circumstances, a court should not permit a witness to state the results of a computer's operations without having the program available for the scrutiny of opposing counsel and his use on cross-examination. *United States v. Dioguardi*, 428 F.2d 1033 (2nd Cir.), cert. denied, 400 U.S. 825, 91 S.Ct. 50, 27 L.Ed.2d 54 (1970). Moreover, such availability should be made known sufficiently in advance of trial so that the adverse party will have an opportunity to examine and test the inputs, program and outputs prior to trial. *United States v. Russo*, supra, 480 F.2d at 1241.[39]

Indeed, some state judges have made the requirements for authenticating a business record interchangeable with those for laying a foundation for its admissibility under the hearsay exception.[40]

The implications arising from these findings of interchangeability appear to illustrate the poorly articulated need to incorporate a requirement of showing trustworthiness or reliability (typically a finding made from within the context of a hearsay determination), into the authentication process.

It is clear that computer-generated information that is not a business record might consist of a digital photograph of an accident scene taken by a bystander, a computer-generated document containing a home inventory for insurance purposes, or a non-business related e-mail containing allegedly defamatory matter.

None of these examples can be easily (if at all) included in the business records category, and it is not surprising that there is no authority directly addressing these examples and evaluating whether they are hearsay or not. Indeed, it appears that the drafters' intent, although not specifically mentioned, was to make F.R.E. Rule 801 a limiting definition (and a limiting evidentiary exclusion rule) such that if a specie of evidence did not fit clearly into one of the definitions of hearsay, it was not to be considered hearsay:

> The definition [of hearsay set forth in Rule 801] does not in terms say that everything not included within the definition is not hearsay, but that was the intended effect of the rule, according to the Advisory's Committee's Note.). *U.S. v. Hamilton*, —-F.3d—-, 2005 WL 1519112, citing John W. Strong, McCormick on Evidence, § 246, at 97 (5th edition 1999).[41]

With the vast amounts of digital information generated each year, much of it non-business records, it is clear that a well-articulated approach to computer generated information that is hearsay that falls into F.R.E. Rule 807 will need to be developed.

Hearsay, digital data and the 'declarant'

An increase of what at least one judge has called a 'lack of understanding' of computer-generated evidence is a new and critical complication that arises out of attempts to define computer generated information.[42] This complication involves semantics, specifically those relating to the concept of hearsay. Hearsay is defined as a 'statement, other than one made by the declarant while testifying at the trial or hearing, offered in evidence to prove the truth of the matter asserted.'[43] A statement is defined in part as an oral or written assertion intended to be an assertion.[44] A declarant is defined as a person who makes a statement.[45] The Federal Rules of Evidence appear to state, therefore, that hearsay does not exist without a declarant, and that the pre-condition to being a declarant is that a declarant must be a person. The semantic problems arise from the meaning and application of the term declarant as it appears in the various hearsay provisions of the Federal Rules of Evidence. A literal interpretation of F.R.E. Rule 803 is that a declarant may not be

[38] *Com. v. Klinghoffer, 564 A. 2d 1240, 1242-1243 (Pa 1989).*
[39] *Klinghoffer, 564 A. 2d at 567.*
[40] *F.D.I.C. v. Carabetta, 739 A.2d 301 (Conn. App. 1999) ' Requirements for authenticating a business record are identical to those for laying a foundation for its admissibility under the hearsay exception for business records'.*
[41] *It is perhaps with good reason that the Supreme Court of Connecticut noted that the divergent views on computer-generated evidence arise in large part from the lack of understanding by those at the bar and the bench. State v. Swinton, 847 A. 2d 921, footnote 24 (Conn. 2004), at 938.*
[42] *State v. Swinton, 847 A. 2d 921, footnote 24 (Conn. 2004), at 938.*
[43] *Fed. R. Evid. Rule 801(c).*
[44] *Fed. R. Evid. Rule 801(a).*
[45] *Fed. R. Evid. Rule 801(b).*

computer data or a computer program because neither data nor computer program is a person.[46] Indeed, two United States Circuits, and at least one District Judge have held that where a computer generates data without the assistance of a person, there is neither a 'statement' nor a 'declarant' and therefore no hearsay.[47] Some courts have distinguished between computer 'generated' and computer 'stored' information in making a hearsay determination.[48] That line of authority is emblematic of the result of a lack of understanding of how computers work, as all computer information is always first generated. There can be no storage of computer information without generation occurring first. Computer-generated information may then be stored, transmitted, or even deleted, but it must exist before it is stored, and in order to exist it must be generated. This issue is related to the distinction between 'original' data, and origination, source, or first instantiation, of computer generated information. Accordingly, an analysis in relation to 'generated' and 'stored' data is fiction, and creates a distinction without a difference, although some might wish the matter was otherwise.[49]

Digital data is hearsay

There is a plausible argument that can be made in support of the proposition that all digital data constitutes some type of hearsay. Certain assumptions must first be made. First, computer generated information of any type, whether output, operating system or application files or data, and even the metadata, are statements made by a computer programmer or like person. These statements, or assertions, are conditional statements, which in essence provide instructions to a computer that, given a certain set of conditions, the computer is told to make a statement on behalf of the computer programmer. That statement may be another instruction, or it may be computer generated information output by the computer.[50] The computer only generates information it is instructed to make on behalf of the person instructing it to make a statement.[51] Contrary to popular opinion, computers do not make mistakes, nor do they generate any information not instructed by a human programmer to make. If a mistake is made, it is not the computer that makes a mistake, but the result of a mistaken statement (i.e., an instruction or assertion) that a computer is told by the programmer to make – whoever the programmer may be (that is, a third person may cause malicious software to be downloaded on to a computer, and the computer will thus take instructions from this software).

The *Hamilton* case provides a good starting example of how a court can get it wrong, even by drawing a conclusion that is not supported by logic. Here, the judge determined that a file header cannot be a statement made a person who transmits that file to another computer over the internet. The *Hamilton* court accordingly ruled that there is no hearsay because there is no person making a declaration as required by F.R.E. 801(b). However, the commands contained in computer programs to create a file header, to transmit a file, to receive a file, to create a log of file creation, transmission or receipt activities, and to enter or not enter information into a log file, are all statements and may be considered to be a declaration of a person, that person being a programmer instructing a computer to make such statement in his or her stead. The instructions generally provide for the following analysis:

[46] Stevenson v. State, 920 S.W.2d 342 (Tex.App.-Dallas 1996). Computer not a declarant and information generated by a computer held not to be hearsay.

[47] United States v. Hamilton, 413 F.3d 1138, 1142 (10th Cir.2005) (file header information accompanying pornographic images uploaded to the internet held not to be hearsay); see also Hawkins v. Cavalli, 2006 WL 2724145 at p. 12 (N.D. Cal 2006) header generated by a facsimile machine was not hearsay because "nothing 'said' by a machine is hearsay").

[48] Hawkins v. Cavalli.

[49] The author has experienced first-hand attempts to delineate between ESI 'generated', and ESI 'stored' in a litigation matter pending at the publication date of this article. The author's firm represents the plaintiff, and requested from the defendant electronically stored information, in native data format, with all associated metadata, and as generated by defendant in the conduct of its everyday activities. The defendant produced documents in TIFF, rather than in native data format, claiming that while it might have 'generated' such information in 'live' or native format, it 'stored' such information only as TIFF format files. The difference between generated and stored here is significant. The 'generated' ESI here would have provided searchable content and metadata. The TIFF files produced were not searchable, and contained no metadata. In this instance, the 'first instantiation' of data could only be the data as generated, and not as ultimately stored. If this matter does not settle, the author intends to raise the issue before the court.

[50] This is demonstrated in the case of State of Connecticut v Julie Amero (Docket number CR-04-93292; Superior Court, New London Judicial District at Norwich, GA 21; 3, 4 and 5 January 2007), where the police officer for the prosecution insisted that the colour of a hyper-link proved that the accused had clicked on to a pornographic web site because it was red, when in fact the web designer entered code to the web page to make it red when viewed. For an exhaustive analysis of this case, see Stephen Mason, general editor, International Electronic Evidence, (British Institute of International and Comparative Law, 2008), xxxvi-lxxv.

[51] The traditional approach to hearsay evidence has more concerned with the elimination of second hand evidence provided by a witness who for some reason is not available to be cross examined in court. This approach fails utterly when faced with the inherent traits unique to digital evidence. Whilst it is true that digital evidence is ultimately generated by a computer, it is also the result of the speech, or declaration, of at least one computer programmer, speaking in a particular language, and translated by the computer into human readable output. Although the computer is not human, the information it generates represents the declaration of its programmers, as to what human readable output, or statement, should be generated.

> When a certain condition or conditions are met, I (the computer programmer or system administrator) want you (the computer) to say 'this' and nothing else, on my behalf.

This means that a computer and computer program will only produce information within the purview of the instructions contained in the source code of the application, or program, and the application or program will only produce information intended to be created by the declaration of the creator of that application or program. An argument can be made, therefore, that there is, and must always be, a person-declarant for any computer-generated information. To find that computers autonomously generate information independent of direct human instruction, (as does the 10th Circuit in *Hamilton* and the District Court for the Central District of California in *Hawkins*) resembles anthropomorphism, and would impart sentience into computing devices that simply (and at least at present) does not exist.[52]

Moreover, the statement made by a programmer to a computer that instructs *the computer to make another statement, such as a file header or other metadata, illustrates the computer programmer's desires and intent to make his or her statement through that computer's processes*. It is not, as so presciently stated by Judge Van Graafeiland, merely a 'calculation' made a machine with a 'giant memory'.[53] For instance, the file header contains specific information, including a statement made by programmer that he or she desires to convey if certain conditions are met, including a statement of time. Note that ultimately, programmers, administrators and human users of a computer are making statements. Persons make these statements, and these statements made by these persons can easily be deemed as declarations falling within the purview of the hearsay rule. To date, no authority expressly adopts this position.[54] If, however, the objective is to provide for uniformity and consistency in relation to the authentication of digital data, the treatment of computer-generated information generally as hearsay would be a major step in reaching this aspiration.

Determining what is hearsay

Judicial authority appears to divide computer-generated information into three neat categories for the purpose of distinguishing what is hearsay. The first category refers to the creation of computer-generated information input into a computer solely by a person. The second category refers to that class of computer-generated information input into a computer in part by a person, and in part by a computer application. The third category refers to computer information generation created without direct human input or assistance.[55] A person creating a memorandum using a word processing application may exemplify the first category. The second category is exemplified by a person creating a form for a computer to arrange and complete. An example of the third category of computer-generated information exists where a computer creates a record of a transaction with another computer. These categories will be examined from the perspective of the traditional approach, and will consider the complications and contradictions either created or left unresolved by that approach.

First category: the memorandum 'created' by a human

The content of a memorandum is created by a person and is generally considered hearsay whether or not it is also considered a business record. If the memorandum is a business record, the provisions of F.R.E. 803(6) must be satisfied.[56] F.R.E. 803(6) requires that either the author of the memorandum must give evidence to provide corroborative testimony, or a 'custodian or other qualified witness' must testify that the 'data compilation' was 'made at or near the time by, a person

[52] *The late Alan Turing is considered by many to be the father of modern binary computing, and he described a 'test' for computer independence of thought, or sentience. The Turing Test is a proposal for a test of a machine's capability to demonstrate thought. Described by Professor Alan Turing in his paper 'Computing machinery and intelligence' Mind (1950) 59, 433-460, it proceeds as follows: a human judge engages in a natural language conversation with two other parties, one a human and the other a machine; if the judge cannot reliably tell which is which, then the machine is said to pass the test. It is assumed that both the human and the machine try to appear human. In order to keep the test setting simple and universal (to explicitly test the linguistic capability of the machine instead of its ability to render words into audio), the conversation is usually limited to a text-only channel such as a teletype machine as Turing suggested or, more recently IRC or instant messaging; http://www.loebner.net/Prizef/TuringArticle.html.*

[53] *Perma Research and Development v. Singer Co., Van Graafeiland, J, dissenting, 542 F.2d 111, 124 (2d Cir. 1976).*

[54] *Such analyses are most likely to be found in dissenting opinions, and even then little consideration is given to the analysis. The dissenting opinion in an unpublished Virginia case considers the issue with the intensity of a Klieg light, but ultimately disregards the categorization of computer-generated information into 'hearsay' and 'non-hearsay': '...[i]t is unlikely that computer-generated evidence will be offered into evidence for some purpose other than 'to prove the truth of a matter asserted,' and thus is hearsay'), Watlington v. Commonwealth, 2000 WL 1672871 (Benton, J, dissenting and citing Randy Snyder, 'Note, Assuring the Competency of Computer-Generated Evidence', 9 Computer Law Journal 103, 104 (1989). [italics added]*

[55] *For a similar analysis, see Stephen Mason, Electronic Evidence: Disclosure, Discovery & Admissibility, xiii.*

[56] *The memorandum is both an F.R.E. 803(6) memorandum and a 'data compilation.' The difference is that a memorandum has some semantic meaning ascribed to it, transforming it into 'information.' For the purpose of this example, however, the terms are used interchangeably.*

with knowledge, if kept in the course of a regularly conducted business activity, and if was the regular practice of that business activity to make the ... data compilation.' Notably, F.R.E. Rule 803(6) also requires that there be no indications of a lack of trustworthiness as a precondition of admissibility.

If the memorandum is considered not to be a business record, another traditional approach might still deem the contents of the memorandum hearsay, and therefore it will be necessary to comply with the pre-condition regarding admissibility under the residual hearsay requirements set out in F.R.E. 807. These requirements include, amongst other things, a showing of the 'equivalent circumstantial guarantees of trustworthiness' required by F.R.E. Rules 803 or 804. F.R.E. 807 therefore appears to incorporate, by reference, the 'no lack of trustworthiness' standard set forth in the business records hearsay exception provisions of F.R.E. 803(6).

There are significant problems with this analysis. All computer-generated information has metadata, or data about data, generated in association with the generation of the content itself. The question that then arises is whether the data compilation comprising the memorandum includes the content of the memorandum, and the metadata associated with that memorandum. It must be correct that the additional data is included with the content.[57] That metadata, which is also computer-generated information, can contain a plethora of information, including source data, time and date information, a digital signature, routing information, date of creation, the last time it was viewed, modifications, the approval of a purported person who reviewed the content, and even the application and version of the application with which the content was created.[58] It is asserted by some, that the generation of this data, is made without the input or assistance from a person. In accordance with decided authority and F.R.E. 801(b), this information could not be considered hearsay, even if it otherwise might be considered a business record. Thus, while the content of the memorandum might be hearsay (whether or not a business record), the associated metadata responsible for all aspects of its existence and format inexplicably is not. If the content is a person or declarant, and metadata is anything but a person or declarant, it is suggested that a two-step authentication process for such computer-generated data ought to be considered. The content of the memorandum, which is hearsay, would first require determination under the provisions of F.R.E. Rule 803(6) or 807. The metadata associated with the memorandum, however, would only require authentication under the provisions of F.R.E. 901(b)(1) or 901(b)(9). Under this analysis, the memorandum created by a person and input into a computer could never be considered to be created only by a person and therefore purely hearsay under either F.R.E. 803(6) or 807. Not surprisingly, the same analysis may be used where a person creates a form to be filled out by other people using various forms of software.

In other words, while the content of the memorandum would be considered hearsay, and subject to analysis as to whether it was hearsay and therefore to be excluded, or an exception and therefore admitted, the metadata associated with the content would only need to be authenticated to be admitted. Since the reliability of metadata in some instances may be of greater evidential significance than the content (such as where the metadata, but not the content has been altered) issues of reliability, when considering the position on the status of hearsay, would attach to both the content and the metadata. In reality, therefore, computer-generated information in categories one and two are one and the same, and should be treated in an identical manner.

The element of time significantly complicates any hearsay analysis. At issue will be what time is referred to as it relates to the memorandum. It could be the time that the document was created by the purported author of the memorandum; the time stated in the content of

[57] *An even more intriguing possibility exists where a person digitally signs an entire data compilation, including metadata. The digital signature is a representation of a statement by the purported signer, and the metadata by definition forms a part of that statement, even though first instantiated by an 'automated' computer process.*

[58] *Metadata is not only evidence about evidence, but is evidence itself. Log files, master file tables, e-mail headers and the like are all evidence of digital events that occur within a computer, and these digital events may, by themselves be used to prove an assertion or claim. For example, the time of an event associated with a memorandum may appear in at least two areas outside the memorandum that a human is able to view. These times should not differ, but may well do so in the event of time-based data manipulation. Without access to such metadata, a party would not have the ability to test the consistency of the asserted time of relevancy (if not the reliability of the time itself). A second example supporting the production of metadata in evidence exists in the case of 'hybrid documents' or documents of one format embedded within documents of another format. For instance, it is easy to bring together a Microsoft Excel document into a Microsoft Word document. In such cases, the Excel® spreadsheet could clearly be considered metadata to the Word® document. The Excel® document, potentially containing relevant evidence, would be rendered totally invisible and undetectable to the reader perusing the document using Word.® If a producing party converted the hybrid Word® document to PDF format, the format conversion process would strip all the Excel® information. The production of all relevant metadata is therefore critical to the efficacy of the discovery process itself.*

the memorandum (which may differ from the time the document is recorded as being created); the time that the memorandum was created according to the metadata information (i.e., file properties or file header). In addition, other questions that might be posed include the time typed into the memorandum, and whether this constitutes a 'declaration' by a 'person' of the 'time' of the declaration. A further issue is the time value contained in the metadata, and whether it is a statement by the person who told the computer to state a specific time on his or her behalf. Arguably, it may be necessary to reconcile the 'time' contained in the content, or 'hearsay' portion of the data compilation as being admissible as a part of a 'declaration' by a 'person' with a different 'time' statement (and statement it is) contained in metadata, which may arguably only be admitted under the F.R.E. 901(b)(9) 'accuracy of result' rule.[59]

If these two 'times' differ substantially, and if reliability is the new watchword for admissibility of computer-generated evidence, there are a number of possible permutations: (1) the computer-generated content with the more 'reliable' time is admissible, and that the computer-generated information content considered 'less reliable' is excluded, (2) that the entire data compilation, including content and metadata, must be excluded, or (3) that the entire data compilation is admissible. It is respectfully submitted that none of these options can be preferred, because the current criteria for categorizing and evaluating computer-generated information are contradictory, cumbersome, and ill-suited to accomplish the task. For example, the first choice would mean a court excludes metadata and admit content, or admit metadata but exclude content, both of which would defeat any possibility of establishing the provenance of the computer-generated information offered as evidence. Excluding or admitting the entire data compilation might obviously serve to further the purpose of one party, but it certainly would detract from the integrity of the evidentiary process in particular, and the efficacy of trial proceedings in general.

Characterizing all computer generated information as hearsay, and imposing a 'reliability' requirement as an exception to the exclusionary rule, would avoid these artificially created distinctions.

Second category: digital data generated in part with human assistance

As discussed above, no computer data can be created or generated by a human without some associated data or metadata, generated by the computer itself. Accordingly, and in this way, computer-generated information described in the first and second categories are identical.

Third category: digital data generated without a human being

In this category, metadata created during the generation of computer information, such as a file header (or data about data) created during an upload of a image file to a remote computer, has been held not to be hearsay because, using a strict application of the hearsay rule, there was no 'person' making a declaration. In *U.S. v. Hamilton*, 412 F3d 1138, 2005 WL 1519112 (10th Cir. 2005) the court held:

> The district court in this case correctly concluded that the header information that accompanied each pornographic image was not hearsay. Of primary importance to this ruling is the uncontroverted fact that the header information was automatically generated by the computer hosting the newsgroup each time Hamilton uploaded a pornographic image to the newsgroup. In other words, the header information was generated instantaneously by the computer without the assistance or input of a person. As concluded by the district court, this uncontroverted fact clearly places the header information outside of Rule 801(c)'s definition of 'hearsay.' In particular, there was neither a 'statement' nor a 'declarant' involved here within the meaning of Rule 801.

In this instance, a cogent argument appears to be made to the effect that the computer-generated information created by a remote computer during the process by which the remote computer receives computer-generated information transmitted to it from another computer is not a statement of a person, and therefore not hearsay. In this instance, both metadata information (file header, IP address) as well as the content created by the remote computer might be considered not to be hearsay and subject only to the 'accuracy of result'

[59] *As discussed above, a third possibility is for a court to consider metadata as hearsay, as the declaration of a person in the position of a computer system or network administrator, a computer programmer, and the like. In that case, the admissibility of both the memorandum (if deemed hearsay) and its associated metadata would be predicated upon complying with the appropriate requirements of admissibility provided under F.R.E. 803(6) or F.R.E. 807.*

requirement of F.R.E. 901(b)(9). The argument against this logic and in favour of determining the data transmitted to be hearsay, is that the receiving computer is carrying out the stated intent or declaration of the system or network administrator, or a programmer, to carry out some request that the receiving computer was told by the sending computer, which in turn was requested by a statement or declaration of the person or sender.

It is suggested that all computer generated information is hearsay of some sort, and the issues raised by the creation of the artificial distinctions between human generated computer information and non-human generated computer generated information illustrate that ultimately, these categories are merely distinctions without a difference.

Constitutional issues – digital data as speech

In a recent decision from the Second Circuit Court of Appeals,[60] the court determined that computer language, including object code as well as source code, to be 'speech' for purposes of the First Amendment:

> Having concluded that computer code conveying information is 'speech' *450 within the meaning of the First Amendment, we next consider, to a limited extent, the scope of the protection that code enjoys. *Universal City Studios, Inc. v. Corley* 273 F.3d 429, *449 -450 (2d Cir. 2001)
>
> But the fact that a program has the capacity to direct the functioning of a computer does not mean that it lacks the additional capacity to convey information, and it is the conveying of information that renders instructions 'speech' for purposes of the First Amendment. The information*448 conveyed by most 'instructions' is how to perform a task.' Universal City Studios, Inc. v. Corley 273 F.3d 429, *447-448 (2d Cir. 2001).
>
> Programmers use snippets of code to convey their ideas for new programs; economists and other creators of computer models publish the code of their models in order to demonstrate the models' vigor. Brief of Amici Curiae Dr. Harold Abelson et al. at 17; Brief of Amici Curiae Steven Bellovin et al. at 12-13; see also *Bernstein v. United States Department of Justice*, 176 F.3d 1132, 1141 (9th Cir.) (concluding that computer source code is speech because it is 'the preferred means' of communication among computer programmers and cryptographers), reh'g in banc granted and opinion withdrawn, 192 F.3d 1308 (9th Cir.1999). *Universal City Studios, Inc. v. Corley* 273 F.3d 429, *448 (2d Cir. 2001).
>
> Reinforcing the conclusion that software programs qualify as 'speech' for First Amendment purposes-even though they instruct computers-is the accelerated blurring of the line between 'source code' and conventional 'speech.' There already exist programs capable of translating English descriptions of a program into source code. Trial Tr. at 1101-02 (Testimony of Professor Andrew Appel). These programs are functionally indistinguishable from the compilers that routinely translate source code into object code. These new programs (still apparently rudimentary) hold the potential for turning 'prose' instructions on how to write a computer program into the program itself. Even if there were an argument for exempting the latter from First Amendment protection, the former are clearly protected for the reasons set forth in the text. As technology becomes more sophisticated, instructions to other humans will increasingly be executable by computers as well. *Universal City Studios, Inc. v. Corley* 273 F.3d 429, *448 (2d Cir. 2001).
>
> ... but ... code, because it uses a notational system comprehensible by humans, is communication that qualifies as speech. *Universal City Studios, Inc. v. Corley* 273 F.3d 429, *449 (2d Cir. 2001)

Even object code, source code, as well as developer's remarks from uncompiled code have been held to comprise 'speech' for purpose of the First Amendment.[61] Accordingly, if computer generated information is held 'speech' for purpose of the First Amendment, it has to be questioned as to why it has deprecated to the status of 'non-speech' for hearsay purposes. This dichotomy has serious implications in criminal proceedings.

[60] *Universal City Studios, Inc., v. Corley 273 F.3d 429 (2d Cir. 2001) 447-450.*

[61] *The defendants asserted at oral argument that DeCSS, or some versions of it, contain programmer's comments, ' which are non-executable appendages to lines of executable code.' Tradescape.com v. Shivaram, 77 F.Supp.2d 408, 418 (S.D.N.Y. 1999). Such comments are protected by the First Amendment: Universal City Studios, Inc. v. Reimerdes 82 F.Supp.2d 211, *220 (S.D.N.Y. 2000).*

Crawford v Washington – testimonial hearsay and sixth amendment rights

The decision in *Crawford v Washington*[62] generally holds that a Sixth Amendment right to cross-examination will arise where testimonial hearsay is used to establish an element of a crime or used to convict. Consider the application of this to computer-generated information used to convict or to establish an element of a crime. When considered, for example, in respect of the output of a blood alcohol testing appliance or other electronically stored evidence, the issue of whether ESI is hearsay takes on new and increased significance. If such computer-generated evidence (the blood alcohol testing device is a computer) is testimonial hearsay, a defendant will be entitled to cross examine the source code and object code in order to establish his or her innocence.

Computer code and output has the ability to 'speak' for someone, and at times this 'someone' can be a coder or programmer. In an exhibit to a software patent issued by the US Patent and Trademark Office, the patentee included his uncompiled source code as an exhibit. In a criminal matter, as a defendant, it might be necessary to know how the computer code in a device that was instrumental in providing evidence that a crime was committed, might have so failed. Under the doctrine in *Crawford v Washington*, the defendant would be guaranteed the right under the Sixth Amendment to cross examine the 'code' or speech of the programmer, and perhaps even the programmer. If, however, the computer code, or computer generated information, was deemed not to be hearsay, the right to examine either the computer source code or the programmer might not be guaranteed.

The Supreme Court recently expanded the application of the *Crawford* doctrine and appears to be edging toward an understanding that a computer might indeed be the declarant uttering testimonial hearsay, thereby enabling the defendant the right of cross examination under the Sixth Amendment Confrontation Clause. In *Melendez-Diaz v. Massachusetts*, —- S. Ct. —- 2009 WL 1789468 (2009), the court determined that a drug testing examiner's certificate (considered equivalent to an affidavit) was both accusatory and testimonial, thus permitting cross examination. The court reasoned that such certificates were created with the sole intent to be used as evidence at trial, and that under Massachusetts law, the sole purpose of the certificate was to provide prima facie evidence of composition, quality and net weight. This evidence was clearly both testimonial and accusatory:

> The certificates here are affidavits, which fall within the "core class of testimonial statements" covered by the Confrontation Clause, id., at 51, 124 S.Ct. 1354. They asserted that the substance found in petitioner's possession was, as the prosecution claimed, cocaine of a certain weight-the precise testimony the analysts would be expected to provide if called at trial. Not only were the certificates made, as Crawford required for testimonial statements, "under circumstances which would lead an objective witness reasonably to believe that the statement would be available for use at a later trial," id., at 52, 124 S.Ct. 1354, but under the relevant Massachusetts law their sole purpose was to provide prima facie evidence of the substance's composition, quality, and net weight. Petitioner was entitled to "be confronted with" the persons giving this testimony at trial. Id., at 54, 124 S.Ct. 1354.'

The court further held that such certificates are the functional equivalent of testimony given in court:

> The fact in question is that the substance found in the possession of Melendez-Diaz and his co-defendants was, as the prosecution claimed, cocaine-the precise testimony the analysts would be expected to provide if called at trial. The "certificates" are functionally identical to live, in-court testimony, doing "precisely what a witness does on direct examination." *Davis v. Washington*, 547 U.S. 813, 830, 126 S.Ct. 2266, 165 L.Ed.2d 224 (2006).

At the heart of the court's extension of the *Crawford* doctrine is the notion of reliability that can be tested. The court found that the aim of the Confrontation Clause is to ensure evidentiary reliability, but confirms that the guarantee is procedural rather than substantive. Moreover, from a procedural context, the Confrontation Clause not only requires reliability, but a method to assess, or test that reliability:

> To be sure, the Clause's ultimate goal is to ensure reliability of evidence, but it is a procedural rather than a substantive guarantee. It commands, not that evidence be reliable, but that reliability be assessed in a particular manner: by testing in the crucible of cross-examination. ... Dispensing with confrontation because testimony is obviously reliable is akin to dispensing with jury trial because a defendant is

[62] *541 U.S. 36 (2004).*

It appears that perhaps in future decisions, the reliability of computer-generated information (testable accuracy and trustworthiness) may be made a precondition for admissibility, rather than a factor to be accorded post-admission weight by a trier of fact.

> obviously guilty. This is not what the Sixth Amendment prescribes. 541 U.S., at 61-62, 124 S.Ct. 1354.

Footnote five to the *Melendez-Diaz* decision provides a tantalizing hint, and perhaps only a hint, that the court might entertain an appeal based on the Confrontation Clause in connection with the reports (not maintenance reports) created by Breath-a-lyzer (blood alcohol testing) appliances. The argument might be that the output of the appliance (a computer) is ultimately both accusatory and testimonial; that the code, or language used to create the accusatory output will be considered testimonial hearsay, and accordingly that a defendant that is accused of having a higher-than-legal blood alcohol levels based on such appliances, will be afforded the right to examine the code under the Sixth Amendment confrontation clause. A future decision will be interesting, because the report of the breath-a-lyzer itself will need to be deemed a form of hearsay as a predicate for any Confrontation Clause analysis:

> Respondent and the dissent may be right that there are other ways-and in some cases better ways-to challenge or verify the results of a forensic test.FN5 But the Constitution guarantees one way: confrontation. We do not have license to suspend the Confrontation Clause when a preferable trial strategy is available. (p. 8)
>
> FN5. Though surely not always. Some forensic analyses, such as autopsies and breathalyzer tests, cannot be repeated, and the specimens used for other analyses have often been lost or degraded.

It appears that perhaps in future decisions, the reliability of computer-generated information (testable accuracy and trustworthiness) may be made a precondition for admissibility, rather than a factor to be accorded post-admission weight by a trier of fact. In the absence of a new evidence rule directed to the admissibility of digital evidence, the characterization of digital information as hearsay would at least require a demonstration of some degree of testable reliability as a precondition for admissibility.

Computers that accuse

An example of the confusion arising from the failure to determine whether digital data is hearsay or not, is perhaps best exemplified by recent opinions in criminal cases. These decisions focus on the admissibility of information generated by what are commonly known as 'breath-a-lyzer' machines, or computers that measure an automobile driver's breath-alcohol levels. A court in Texas has held that:

> The Intoxilyzer instrument self-generates data. [It] cannot be a declarant. Because the Intoxilyzer is not a declarant, the data it generates is not a statement and cannot be hearsay.[63]

A few recent cases have undertaken a different analysis. A Florida court has required that under the Florida full information law, the source code used in a breath-a-lyzer machine must be produced for examination by the state.[64] The court stated in pertinent part that:

> An instrument or machine that can be used by the State to establish the guilt of an accused subjecting them to mandatory fines, mandatory loss of driving privileges, and loss of freedom (sometime mandatory) should be made available to the defense for open inspection ... [The disclosure of] full information should include the software that runs the instrument. To construe the statute otherwise, is tantamount to granting the state authority to use confidential information (i.e., the software code) to establish the guilt of a criminal defendant ... The software is an integral part of the intoxilyzer. Unless the defense can

[63] *James v. State, 2000 WL 1665126 at *2 (Tex.App.-Dallas 2000) citing Stevenson v. State.*

[64] *F.S.A. Chapter 316, §316.1932(1(f)(4).*

> see how the intoxilyzer breathalyzer works ..., it remains as stated by the Court in *Muldowny* and more recently by Judge Ralph E. Erikkson as being nothing more than a 'mystical machine' used to establish an accused's guilt. *State v. Lentz*, 12 Fla. L. Weekly Supp. 806(a) (18th Judicial Circuit, Seminole County, April 29, 2005).[65]

The court did not address the hearsay issue, but the analysis clearly indicates that computer code is used to establish guilt. The device does not keep any samples of the breath provided by a suspect driver, and the only evidence is the information processed by the appliance. It is inescapable that the computer information generated by the Intoxilyzer therefore accuses (or exonerates) a defendant, and this information is 'spoken' by the code contained in the device. The code conveys information, and that information is the programmer's statement, or declaration. For example, a recent patent issued by the United States Patent Office included some uncompiled source code that contained the comment 'If this fails, we are f*d.'[66] It should be pointed out that an officer who administers the Intoxilyzer test does not determine or asses whether a suspect driver has an blood alcohol level above the limit set by law. Accordingly, the only 'testimony' for blood alcohol level can come from the Intoxilyzer itself, or more specifically, the code that speaks to that information.[67]

The issue as to whether computer-generated information is or is not hearsay may eventually be resolved under the standard articulated by the United States Supreme Court in *Crawford v. Washington*.[68] In that opinion, the Supreme Court held generally that the use of testimonial hearsay automatically invokes a defendant's Sixth Amendment rights to confrontation. A future case might present a series of facts that includes the use of computer-generated information, not generally considered hearsay (such as the output of an Intoxilyzer) but which may nonetheless be considered testimonial. If computer-generated information is held to be testimonial in nature, it would not take a quantum leap in analysis to find that such testimony is hearsay, even when uttered by a computer.

Reliability of digital data

More recent judicial authority appears to indicate a trend away from considering the need to authenticate digital data, whether determined to be hearsay or not, in favour of a more general and flexible concern for reliability.[69] When evidence codes were first developed, there was a preference for personal testimony over documentary testimony because personal testimony was considered more reliable. Over time, admissibility rules have been developed for the introduction of documentary evidence, with reliability as the touchstone to admission.

The modern requirement for reliability appears to be the result of the merger of the 'accuracy of result' test embodied in F.R.E. Rule 901(b)(9), the 'trustworthiness' test enumerated in F.R.E. Rule 803(6), with the 'circumstantial guarantees' test enumerated in F.R.E. Rule 807. Hints of this merger have appeared as early as 1987: 'The principal precondition to admission of documents as business records pursuant to Fed.R.Evid. 803(6) are that the records have sufficient indicia of trustworthiness to be considered reliable.'[70] Other courts have, in more recent decisions, included the concept of reliability into determining admissibility where digital data, whether considered hearsay or not, is being offered into evidence.

The basis for determining the reliability of computer-generated information differs greatly from that of physical evidence. Although the provenance of the evidence must still be established, the requirements in respect of digital data are not the same as physical evidence. Simply put, it is not old wine in new bottles. On the one hand, it is possible to observe the process by which a human controls and applies pen to paper, and forensic tests can be performed to assist in and corroborate witness testimony in connection with a determination as to the authenticity of the document. On the other hand, it is not certain (and certainly without access to and interpretation, or translation of the source or origination data and code) how a computer is programmed to speak for its programmers or content creators. This lack of knowledge means the reliability of the data may not be certain.

Digital data comprises a different specie of evidence.

[65] *State of Florida v. Jack Irish, et al., Case No. 2006-CT-02109 SC (County Court, 12th Judicial Circuit, FL 2006).*

[66] *United States Patent Number 5,619,571 Page 30.*

[67] *But see U.S. v. Washington, 498 F.3d 225 (4th Cir. 2007) holding that a toxicology laboratory testing machine data output was not the out of court statement of the laboratory technician, and appears to eschew the notion that computer generated information is testimonial hearsay, and therefore not subject to Crawford's Sixth Amendment confrontation rights. Note that this decision holds merely that the testing machine's output is not the statement of the laboratory technician, and does not address whether the machine's statement is the statement of the program (and programmer) that generated the data.*

[68] *541 U.S. 36 (2004).*

[69] *For example, see Lorraine v Markel American Ins. Co., 241 F.R.D. 534 (D. Md. 2007); In re Vee Vinhnee, 336 B.R. 437 (9th Cir. BAP 2005).*

[70] *Potamkin Cadillac Corp. v. B.R.I. Coverage Corp., 38 F.3d 627, 632 (2nd Cir.1994); Saks Intern., Inc. v. M/V Export Champion, 817 F.2d 1011, 1013 (2nd Cir. 1987).*

It is designed to be ephemeral, although note that there is a difference between 'wiping' or expunging data, and changing digital data. To thoroughly expunge data from a single computer requires more that merely downloading and running a wiping utility. There are many locations on the hard disc that might either contain a pointer to an earlier version of data, or a copy of data that might otherwise be considered irretrievable. For instance, the operation of an automatic file backup in a word processing application may save one or a number of recent versions of a document. Moreover, evidence of the existence of digital data (and what happened to it during its life cycle) may be found in a master file table or other logging operation that takes place without the knowledge of the user. In addition, a simple erasure or deletion typically does not expunge data, but only removes the 'pointer' to the data, and the data itself remains accessible unless and until overwritten (in whole or in part) by newly generated data. In a networked system, a user may think she is expunging data at a workstation, only to find that the network server automatically copies, archives or backs up all data generated by the workstation. That said, the distinction between expunging data and the ephemeral nature of data may best be explained by presuming that ephemeral digital data relates to the difficulty of proving persistent data integrity (some technologists might describe this as proving 'statefulness'), rather than whether it does or does not exist. This characteristic should be considered when determining reliability. Digital data is almost totally dependent upon corroborative testimony that may have little, if anything to do with the authenticity of the content sought to be admitted. The criteria to ascertain admissibility of computer-generated information must, it is suggested, require a demonstration of heightened reliability. It addition, it must do so in a manner that does not merely mirror the techniques for evaluating physical evidence.

Whether couched in terms such as 'trustworthiness,' 'accuracy,' or 'truthfulness,' the concept of reliability is central for laying a foundation leading to the admissibility of evidence. To the detriment of modern jurisprudence, and within the realm of digital evidence, however, the evolution of the concept of reliability as a precondition to admissibility has not kept up with the revolution in information technology. Nevertheless, there has been some slow but steady judicial recognition of the reliability issues that are unique and inherent to digital data. In a seminal 2004 opinion, the Supreme Court of Connecticut announced its new approach to computer-generated evidence, declaring reliability as an essential pre-condition of admissibility.[71] The *Swinton* opinion held in pertinent part that a trial court improperly admitted into evidence computer enhanced photographs of bite marks and images that purported to represent the defendant's dental structure as lacking a proper foundation. The approach of the court in *Swinton* was to assess the admissibility itself (rather than the weight) of computer enhanced evidence.

It is significant that the *Swinton* court itself grappled with the concept of computer-generated evidence, and noted that there 'is no universal definition of that term.'[72] The members of the *Swinton* court also recognizes the unique evidentiary issues presented by computer-generated evidence and

> d[id] not agree with the state's proposition that the enhanced photographs in the present case are like any other photographs admitted into evidence, and we determine that, to the extent that a computer was both the process and the tool used to enable the enhanced photographs to be admitted as evidence, we consider these exhibits, for the purposes of this analysis, to be computer generated.[73] [internal footnotes omitted]

It was also noted that that the appearance of computer-generated evidence at trials in Connecticut was limited and typically involved business records.[74] In a manner strikingly reminiscent of Judge Van Graafeilands dissenting comments in *Perma Research*, the members of the *Swinton* court appear to bemoan the paucity of understanding by both attorneys and the members of the judiciary about the nature and issues presented by digital data, and suggested that this lack of understanding has contributed in turn to the scarcity of relevant authority:

> Commentators have attempted to explain this lack of case law involving basic foundational challenges to this sort of evidence. "Although computer systems raise serious reliability issues, the reported cases do not adequately reflect this reality." R. Garcia, "'Garbage In, Gospel Out': Criminal Discovery, Computer Reliability, and the Constitution," 38 UCLA L. Rev. 1043, 1087 (1991). Why do the reported cases fail to adequately expose the serious reliability issues raised by computerized information? Many people, including defense attorneys, prosecutors, judges, and juries, do not understand computers.[75]

[71] *State v. Swinton, 847 A.2d 921 (Conn. 2004).*
[72] *Swinton, 847 A. 2d at 937.*
[73] *Swinton, 847 A. 2d at 938.*
[74] *Swinton, 847 A. 2d at 938.*
[75] *Swinton, 847 A. 2d footnote 25 at 940.*

The *Swinton* reliability test for admissibility of computer-generated evidence has been accorded increasing authority, and has been relied upon and extended by other courts in considering the admissibility of computer-generated (rather than enhanced) exhibits. In a recent New Jersey decision the court stated:

> In our view, the use of a computer-generated exhibit requires a more detailed foundation than that for just photographs or photo enlargements. The latter "must be proved to be faithful representations of the subject at the time in question. Fundamentally, photographs are deemed to be pictorial communications of a qualified witness." *State v. Smith*, 27 N.J. 433, 448, 142 A.2d 890, 899 (1958). However, considering the reliability problems arising from computer-generated exhibits and the processes by which they are created, see *State v. Swinton*, 268 Conn. 781, 847 A.2d 921, 941-43 (2004), there must be "testimony by a person with some degree of computer expertise, who has sufficient knowledge to be examined and cross-examined about the functioning of the computer".[76]

In *In re Homestore.com Inc. Securities Litigation*, the court found that web page print-outs bearing a URL (Uniform Resource Locator) address and date stamp were improperly authenticated by declaration, and did not '... bear the indicia of reliability demanded for other self-authenticating documents under Fed. R. Evid. 902. To be authenticated, some statement or affidavit from someone with knowledge is required; for example, Homestore's web master or someone else with personal knowledge would be sufficient.'[77]

More recently, a bankruptcy panel for the 9th Circuit Court of Appeals upheld a bankruptcy court's refusal to admit the unopposed offer of a computer-generated print-out consisting of an American Express cardholder's transactions.[78] Adopting what appears to be the blending of the business records hearsay exception F.R.E. Rule 803(6) with the 'accurate results' standard provided by F.R.E. Rule 901(b)(1) and (9) and the 'circumstantial guarantees of trustworthiness' set forth in F.R.E. Rule 807, the court affirmed the lower court's finding that 'the electronic nature of the records necessitated, in addition to the basic foundation for a business record, *an additional authentication foundation* regarding the computer and software utilized in order to assure the continuing *accuracy* of the records.'[79]

Conclusion

It has long been accepted that computers are not 'calculators with a giant memory.'[80] The prescience of Judge Van Graafeiland's dissenting comments in Perma Research has been borne out by the ensuing decades of ill-informed and often contradictory judicial authority. Judge Van Graafeiland stated that:

> as courts are driven willy-nilly into the magic world of computerization, it is of utmost importance that appropriate standards be set for the introduction of computerized evidence ... Although the computer has tremendous potential for improving our system of justice by generating more meaningful evidence than was previously available, it presents a real danger of being the vehicle of introducing erroneous, misleading, or unreliable evidence. The possibility of an undetected error in computer-generated evidence is a function of many factors: the underlying data may be hearsay; errors may be introduced in any one of several stages of processing; the computer might be erroneously programmed, programmed to permit an error to go undetected, or programmed to introduce error into the data; and the computer may inaccurately display the data or display it in a biased manner. Because of the complexities of examining the creation of computer-generated evidence and the deceptively neat package in which the computer can display its work product, courts and practitioners must exercise more care with computer-generated evidence than with evidence generated by more traditional means. Roberts, 'A Practitioner's Primer on Computer-Generated Evidence', 41 U.Chi.L.Rev. 254, 255-56 (1974). There are those knowledgeable in the field of computerization who believe that new evidentiary rules will be required to channel and control the use of this new medium. Freed, Computer Records and the Law Retrospect and Prospect, 15 Jurimetrics J. 207, 208 (1975).[81]

Judge Van Graafeiland's dissenting comments are so timely that they might have been written yesterday.

The proposed characterization of all computer generated information as hearsay is supported by both

[76] *Rodd v. Raritan Radiological Associates, 860 A. 2d 1003, 1011-1012 (N.J. Sup. A.D. 2004).*
[77] *In re Homestore.com Inc. Securities Litigation, 347 F.Supp.2d 769, 782-783 (C.D.Cal. 2004).*
[78] *In re Vee Vinhnee, 336 B.R. 437 (9th Cir. BAP 2005).*
[79] *In re Vee Vinhnee, at 444-445. [Emphasis Added]*
[80] *Perma Research and Development v. Singer Co., Van Graafeiland, J, dissenting, 542 F.2d 111, 124 (C.A.N.Y. 1976).*
[81] *Perma Research and Development v. Singer Co., Van Graafeiland, J, dissenting, 542 F.2d (C.A.N.Y. 1976) at 124-125.*

an examination of computer language, as well as long-standing authority that treats computer code as 'speech' for purposes of the First Amendment. There is no supportable or substantial reason to change the meaning of computer-generated information from 'speech' for protecting First Amendment rights, to 'non-speech' in determining whether it is hearsay, especially in those cases where hearsay would necessarily invoke rights under the Sixth Amendment Confrontation in criminal matters under the *Crawford v. Washington* doctrine.

In the absence of the adoption of a new evidence rule addressing the admissibility of computer-generated evidence, treating all computer generated information as hearsay would eliminate low-quality evidentiary admissibility, and provide a better 'reliability' standard that is now beginning to emerge for computer-generated information. It is strongly suggested that the acceptance of computer-generated information as hearsay, and the adoption of a generalized requirement of 'reliability' would provide an effective and consistent approach to the admissibility of digital data. Further, the universal adoption of a general reliability standard would curtail current attempts to categorize digital data as hearsay or not, thereby eliminating what is in essence a distinction without a difference. Moreover, the adoption of a standard of reliability would represent a significant step forward in the development of a unified, flexible set of criteria to establish admissibility for digital data.[82] In the future, computer-generated information that is (1) generated by human input, (2) by hybrid human and computer input, (3) by the computer only, becomes merely digital data. By deeming computer-generated information as hearsay, such pseudo-categorizations of computer-generated information would be eliminated, and the requirement of reliability could be universally imposed and uniformly considered. In addition, the confusion between 'computer stored' and 'computer enhanced' data may disappear, thereby removing the current illogical, contradictory, and ultimately unworkable distinction that surrounds digital data. Eventually, it is hoped, the evidentiary rules will be revised to accurately reflect the need for a flexible requirement of testable reliability as a pre-condition for admissibility of digital data.

Steven W. Teppler is Senior Counsel KamberEdelson, LLC in New York, specializing in digital data protection, electronic discovery, and authentication and admissibility issues. He is a Co-Chair of the American Bar Association's eDiscovery and Digital Evidence Committee, and a contributing author to Foundations of Digital Evidence (American Bar Association, 2008).

steppler@kamberedelson.com
http://www.kamberedelson.com

[82] *A single hearsay rule applicable to the testing of both witnesses and computer-generated information may well prove too complex or unwieldy. Indeed, what may work best is a new, separate evidence rule made expressly applicable to computer-generated evidence, and requiring a threshold showing of testable accuracy, reliability and trustworthiness as a pre-requisite to admissibility.*

