Notes

SEMANTIC SEARCHES

ATHUL K. ACHARYA†

ABSTRACT

Courts and commentators have struggled with the problem of cabining digital searches while still allowing law enforcement sufficient latitude to be efficient and effective. This Note examines current proposals, such as requiring search protocols or abandoning the plain view doctrine, before proposing a solution of its own: revisiting the constitutional requirement of particularity in the warrant. Focusing on particularity is not new; the problem is describing, ex ante, where to search within a corpus of seized data. The language of files and folders is both inadequate and incoherent for this task, but in rejecting it, courts have largely given up on particularly describing where in the data to search.

Data is information, and information has meaning—semantics. Computers are increasingly able to sort and segregate data according to the human meaning it represents. Accordingly, magistrate judges can describe, ex ante in natural language, the type of data that examiners may search based on the evidence sought. Forensic examiners can then use automated tools to retrieve information responsive to that semantic description without searching the entirety of the data. Thus, the privacy of suspects, guilty and innocent, can be protected without giving up the plain view doctrine or compromising effective law enforcement.

† Duke University School of Law, J.D. expected 2014; Purdue University, M.S. 2008; University of Rochester, B.S. 2006. I would like to thank my advisor, Professor Lisa K. Griffin, as well as Ethan Blanton, for their comments and feedback. Thanks also to Conor Reardon, Luke Ricci, Tim Hultzman, Nathan Williams, Shifali Baliga, Ethan Carroll, James Waters, and the rest of the *Duke Law Journal* for improving this Note with their unflagging editorial expertise.

INTRODUCTION

The advent of electronic storage media (ESM) and electronically stored information (ESI)[1] has confounded the law of search and seizure. Can data be "searched"?[2] When does that happen?[3] What is a "particular" warrant or a "reasonable" search?[4] What are the limits on the purview of plain view?[5] The widespread use of ESM and ESI has resulted in difficult questions going to the heart of the balance between law-enforcement and privacy interests. Courts have charted a haphazard course through this minefield,[6] generally—though not always—paying lip service to privacy concerns while allowing law enforcement unfettered and unprecedented discretion in the execution of searches and seizures.[7]

---

1. This Note occasionally uses the terms *hard drive* and *data* as generic stand-ins for ESM and ESI, respectively.

2. *Compare* Orin S. Kerr, *Searches and Seizures in a Digital World*, 119 HARV. L. REV. 531, 551–54 (2005) (proposing that a Fourth Amendment search occurs when ESI is exposed to human observation), *with* Josh Goldfoot, *The Physical Computer and the Fourth Amendment*, 163 BERKELEY J. CRIM. L. 112, 113 (2011) (denying that an examination of ESM is a search at all).

3. *See* Orin S. Kerr, *Ex Ante Regulation of Computer Search and Seizure*, 96 VA. L. REV. 1241, 1248 (2010) (observing that computer searches entail a physical search and seizure followed by an electronic search). For one court's explanation of this two-stage practice, see *infra* notes 42–47 and accompanying text.

4. *See* U.S. CONST. amend. IV (requiring that searches not be "unreasonable" and that warrants "particularly describ[e]" their objects). *Compare, e.g.*, United States v. Hill (*Hill II*), 459 F.3d 966, 978 (9th Cir. 2006) (holding that the requirement that searches must be reasonable is sufficient to constrain computer searches without more), *with id.* at 974 (observing that every file on seized ESM must necessarily be examined (quoting United States v. Hill (*Hill I*), 322 F. Supp. 2d 1081, 1088–89 (C.D. Cal. 2004))).

5. Bryan K. Weir, *It's (Not So) Plain To See: The Circuit Split on the Plain View Doctrine in Digital Searches*, 21 GEO. MASON U. C.R. L.J. 83, 92–93 (2010) (observing that spatial and temporal constraints on the scope of physical searches do not apply to computer searches).

6. *Compare, e.g.*, United States v. Carey, 172 F.3d 1268, 1275–76 (10th Cir. 1999) (suppressing the results of a search that exceeded the scope of the warrant and suggesting a number of techniques to minimize exposure of irrelevant information), *with* United States v. Burgess, 576 F.3d 1078, 1092 (10th Cir. 2009) (refusing to suppress the results of a very similar search that used none of those techniques).

7. *See, e.g.*, United States v. Rosa, 626 F.3d 56, 62, 66 (2d Cir. 2010) (finding that a warrant "lacked meaningful parameters on an otherwise limitless search of Rosa's electronic media," but was not "so defective that an officer [would] lack a reasonable basis for relying upon it"); United States v. Adjani, 452 F.3d 1140, 1149 (9th Cir. 2006) (purporting to "understand the heightened specificity concerns in the computer context" and yet refusing to require any such specificity). *But see* United States v. Comprehensive Drug Testing, Inc., 621 F.3d 1162, 1170, 1174, 1175 (9th Cir. 2010) (en banc) (per curiam) (invalidating a search, albeit on narrow grounds); *id.* at 1178–80 (Kozinski, C.J., concurring) (proposing extensive ex ante regulations).

As always in search and seizure law, the tradeoff is between ensuring effective law enforcement and vindicating legitimate privacy concerns. The question is how to maintain the balance reached in the physical realm when conducting search and seizure in the electronic domain.[8] It has become clear that physical rules cannot be transplanted, unchanged, to the world of digital search and seizure without upsetting that balance.[9] In response, magistrate judges have recently begun requiring a *search protocol* in the warrant—prescribing, ex ante, *how* a computer search is to be executed[10]—but at least one prominent scholar has attacked this practice on both constitutional and normative grounds.[11]

This Note proposes a way to regulate computer searches that is both firmly grounded in the Constitution and normatively attractive. As a starting point, this Note assumes the necessity of broad overseizure of ESI[12] and adopts Professor Orin Kerr's definition of a search of data: exposure to human observation, by way of a display, printer, or other output device.[13] This Note proposes that a warrant to search ESI should be limited to a *semantic zone*—a nontechnical description of the type of content an agent may lawfully search, based on the evidence that the government has probable cause to search for. As a brief preliminary example, if the government has established probable cause to search for pay-owe sheets related to drug trafficking, the appropriate semantic zone to search would be *text document and spreadsheet data*.[14] This Note further proposes that in ESI cases, magistrate judges perform a new, supervisory function, conducted through their traditional role of vetting warrants.

---

8. *See* Orin S. Kerr, *An Equilibrium-Adjustment Theory of the Fourth Amendment*, 125 HARV. L. REV. 476, 480 (2011) ("When new tools and new practices threaten to expand or contract police power in a significant way, courts adjust the level of Fourth Amendment protection to try to restore the prior equilibrium.").

9. *See Comprehensive Drug Testing*, 621 F.3d at 1176 (noting the "serious risk that every warrant for electronic information will become, in effect, a general warrant, rendering the Fourth Amendment irrelevant").

10. *See infra* notes 104–08 and accompanying text.

11. *See infra* notes 109–19 and accompanying text.

12. *See infra* Part I.B. For a cogent discussion of what it means to "seize" ESI, see generally Orin S. Kerr, *Fourth Amendment Seizures of Computer Data*, 119 YALE L.J. 700 (2010).

13. Kerr, *supra* note 2, at 551–54.

14. This may seem to be a broad zone, but it would exclude, for example, the suspect's videos and pictures. For a discussion of a case in which this would have made a difference, see *infra* notes 70–78 and accompanying text.

The remainder of this Note proceeds as follows. Part I starts with a sketch of Fourth Amendment search doctrine concentrating on the particularity requirement, followed by a discussion of the two-stage nature of computer searches. Part II introduces two paradigmatic problems arising from the search of seized ESI and critiques the solutions proposed so far. Part III argues that the root of the problem is the lack of particularity in searches of ESI. It then evaluates current conceptions of ESI and argues that they are inadequate to articulate particularity limits before proposing a new perspective—the semantic perspective—which gives rise to semantic zones. Finally, Part IV explores the application of those rules in various real and hypothetical test cases.

I. LEGAL BACKGROUND

A. Searches Under the Fourth Amendment

The Fourth Amendment regulates government searches and seizures.[15] A Fourth Amendment search occurs when the government violates a personally held and objectively reasonable expectation of privacy.[16] The general rule is that a government search is "*per se* unreasonable under the Fourth Amendment" unless pursuant to either a warrant or one of "a few specifically established and well-delineated exceptions."[17] To obtain a warrant, the government must have probable cause to believe that specific evidence of a specific crime will be found in a specific place.[18] Probable cause is a "fair probability" under the totality of the circumstances that the search will discover evidence of a crime.[19] The government's determination

---

15. The text of the Amendment reads:

> The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.

U.S. CONST. amend. IV.

16. Kyllo v. United States, 533 U.S. 27, 33 (2001). The Supreme Court has recently made clear that in addition to this test, a search occurs when the government "obtains information by physically intruding on a constitutionally protected area." United States v. Jones, 132 S. Ct. 945, 950 n.3 (2012).

17. Horton v. California, 496 U.S. 128, 133 & n.4 (1990) (quoting United States v. Ross, 456 U.S. 798, 825 (1982)).

18. Maryland v. Garrison, 480 U.S. 79, 84 (1987).

19. Illinois v. Gates, 462 U.S. 213, 238 (1983). The Supreme Court has consistently refused to further articulate the contours of the probable-cause standard. *See, e.g.*, Maryland v. Pringle,

of probable cause must be vetted by a "neutral and detached magistrate," who may, if he agrees that there is probable cause, issue a warrant for the search.[20]

A warrant must "*particularly* describ[e] the place to be searched, and the persons or things to be seized."[21] The Framers of the Fourth Amendment were concerned with preventing the practice of general warrants used in England, as well as the related writs of assistance in the Colonies, which authorized the Crown's customs officers to rummage through the homes of colonists and seize prohibited or uncustomed goods.[22] The requirement that both the *object* of the search and the *place* to be searched be particularly described in the warrant is the primary safeguard against general searches.[23]

The particularity requirement is a function of what the police know at the time they seek the warrant. With regard to places, the leading case is *Maryland v. Garrison*,[24] in which police had probable cause to search one third-floor apartment but, unaware that the floor comprised multiple apartments, requested a warrant for the entire floor.[25] The Court held that "if the officers had known, or even if they should have known, that there were two separate dwelling units on the third floor . . . they would have been *obligated* to exclude respondent's apartment from the scope of the requested warrant."[26] Similarly, the object of the search may be described by a "generic classification[]"—for example, "currency"—but "only when a more

540 U.S. 366, 371 (2003) ("The probable-cause standard is incapable of precise definition or quantification into percentages because it deals with probabilities and depends on the totality of the circumstances."); *Gates*, 462 U.S. at 238 ("The task of the issuing magistrate is simply to make a practical, common-sense decision . . . .").

20. Johnson v. United States, 333 U.S. 10, 13–14 (1948).

21. U.S. CONST. amend. IV (emphasis added).

22. United States v. Chadwick, 433 U.S. 1, 7–8 (1976); United States v. Marron, 275 U.S. 192, 195 (1927). For a canonical account of the colonies' experience with writs of assistance, see generally NELSON B. LASSON, THE HISTORY AND DEVELOPMENT OF THE FOURTH AMENDMENT TO THE UNITED STATES CONSTITUTION 51–78 (1937).

23. Andresen v. Maryland, 427 U.S. 463, 480 (1976).

24. Maryland v. Garrison, 480 U.S. 79 (1987).

25. *Id.* at 85.

26. *Id.* (emphasis added). The Court upheld the ensuing search because the officers had not, in fact, known, and acted reasonably upon realizing that the floor contained two apartments. *Id.*

precise description is not possible," as when police do not know every relevant serial number.[27]

Once a search is underway, the object of the search constrains where law enforcement may *reasonably* search.[28] As the Seventh Circuit memorably put it, "[i]f you are looking for an adult elephant, searching for it in a chest of drawers is not reasonable."[29] Thus, even in the absence of a pinpoint description of where to search within the particular place, the description of the *thing* to be seized—itself particularized by probable cause—limits the *scope* of the search.

The plain view doctrine is an important exception to the warrant requirement for seizures. An officer may seize an object without a warrant as long as three conditions are met: (1) the object has lawfully come into the officer's view, (2) "its incriminating character [is] 'immediately apparent,'" and (3) the officer has "a lawful right of access to the object."[30] In *Horton v. California*,[31] for example, the warrant only authorized a search for the proceeds of a robbery, "including three specifically described rings."[32] The Court held that seizure of the weapons used in the robbery did not violate the Fourth Amendment because they were discovered in the course of a lawful search, and it was immediately apparent that they were the weapons used in the robbery.[33]

---

27. United States v. Bright, 630 F.2d 804, 812 (5th Cir. 1980). The canonical test for "whether a description [of the thing to be seized] is sufficiently precise" is set out in *United States v. Spilotro*, 800 F.2d 959 (9th Cir. 1986):

> (1) whether probable cause exists to seize all items of a particular type described in the warrant; (2) whether the warrant sets out objective standards by which executing officers can differentiate items subject to seizure from those which are not; and (3) whether the government was able to describe the items more particularly in light of the information available to it at the time the warrant was issued.

Id. at 963 (citations omitted).

28. *See* United States v. Ross, 456 U.S. 798, 824 (1982) ("The scope of a warrantless search . . . is defined by the object of the search and the places in which there is probable cause to believe that it may be found."). The search at issue in *Ross* was warrantless, but the Court reasoned that the scope of a warrantless search of an automobile, justified by probable cause, is identical to the scope of the search a magistrate could have authorized in a particular warrant, justified by probable cause. *Id.* at 823. Thus, the rule for searches generally is that the object of a search defines its permissible scope. *Id.* at 824.

29. Platteville Area Apartment Ass'n v. City of Platteville, 179 F.3d 574, 579 (7th Cir. 1999).

30. Horton v. California, 496 U.S. 128, 136–37 & n.7 (1990) (quoting Coolidge v. New Hampshire, 403 U.S. 443, 466 (1971) (plurality opinion)).

31. Horton v. California, 496 U.S. 128 (1990).

32. *Id.* at 131.

33. *Id.* at 142.

Enforcement of the Fourth Amendment is primarily manifested in the exclusionary rule[34]—a "judicially created remedy"[35] under which defendants may move to suppress evidence obtained in violation of their Fourth Amendment rights. The primary rationale for the exclusionary rule is that suppressing illegally obtained evidence should deter the police from violating the Fourth Amendment, "'by removing the incentive to disregard it.'"[36]

B. Searches of ESI

The story of ESI searches begins at the "dawn of the information age,"[37] with *United States v. Tamura*,[38] a case about paper records "so intermingled that they [could not] feasibly [have been] sorted on site."[39] The *Tamura* court held that in such a case, officers may seize all the documents as long as they seal them "pending approval by a magistrate of a further search."[40] A few years earlier, the Supreme Court had held that, in searches of intermingled records, it was unavoidable (and therefore permissible) to examine "some innocuous documents . . . at least cursorily, in order to determine whether they are, in fact, among those papers [sought]."[41] Thus, the stage had been set for broad overseizure and modern searches of ESI.

Twenty-four years later, the Ninth Circuit supplied, in *United States v. Hill*,[42] one of the clearest articulations of the broad overseizure practice. The defendant, accused of possessing child pornography, argued that the warrant was overbroad in authorizing seizure of all ESM instead of only that which actually contained child pornography.[43] The court held that the "significant burden" of

34. *See* Weeks v. United States, 232 U.S. 383, 398 (1914) (announcing the exclusionary rule, albeit in different terms); *see also* Mapp v. Ohio, 367 U.S. 643, 650, 660 (1961) (incorporating the exclusionary rule against the states).

35. United States v. Calandra, 414 U.S. 338, 348 (1974).

36. *Mapp*, 367 U.S. at 656 (quoting Elkins v. United States, 364 U.S. 206, 217 (1960)).

37. *See* United States v. Comprehensive Drug Testing, Inc., 621 F.3d 1162, 1169 (9th Cir. 2010) (en banc) (per curiam) (discussing the lineage of ESI search doctrine).

38. United States v. Tamura, 694 F.2d 591 (9th Cir. 1982).

39. *Id.* at 595.

40. *Id.* at 595–96. The *Tamura* court further held that when the need for broad overseizure is known ahead of time, officers should inform the magistrate, who should authorize such seizure only upon determining that "no other practical alternative exists." *Id.*

41. Andresen v. Maryland, 427 U.S. 463, 482 n.11 (1976).

42. United States v. Hill (*Hill II*), 459 F.3d 966 (9th Cir. 2006).

43. *Id.* at 973.

carrying properly equipped computers with trained personnel[44] and the dangers and difficulties of on-site inspection[45] meant that, as long as the affidavit gave a "reasonable explanation," "blanket removal of all computer storage media for later examination" was quite reasonable.[46] Thus, investigations involving computer evidence occur in two stages: first, there is a search of a physical place, followed by an en masse seizure of ESI or ESM; second, there is a search of the seized data for incriminating evidence.[47]

The natural corollary question is how to regulate the subsequent search of the seized ESI. Until recently, there was no ex ante regulation of computer searches; courts, reasoning that "'[t]here is no way to know what is in a file without examining its contents,'" refused to require restrictions in the warrant.[48] Typically, the caveat would be affixed that, despite the need for an open-ended warrant, the reasonableness of the search would be reviewed ex post[49]—but at that later stage, the same logic has served to justify almost any search.[50]

---

44. *See id.* at 974 (quoting United States v. Hill (*Hill I*), 322 F. Supp. 2d 1081, 1088–89 (C.D. Cal. 2004)) (noting the variety of operating systems, file systems, and media types).

45. *Id.* at 974–75 (citing the risk of "compromis[ing] the integrity of the evidence by attempting to access the data at the scene" and the "many hours and perhaps days" that examining every file might take (quoting *Hill I*, 322 F. Supp. 2d at 1089)).

46. *Id.* at 976. As Josh Goldfoot, Senior Counsel of the Computer Crime and Intellectual Property Section at the U.S. Department of Justice, points out, the reasonable explanation requirement has merely led to a rise in boilerplate language in warrant affidavits. *See* Goldfoot, *supra* note 2, at 136–37.

47. Kerr, *supra* note 2, at 547. Kerr offers the following metaphor: "data acquisition refers to collecting the hay, and data reduction involves looking through the haystack for the needle." *Id.*

48. *Hill II*, 459 F.3d at 978 (quoting *Hill I*, 322 F. Supp. 2d at 1090).

49. *Id.*

50. *See, e.g.*, United States v. Richards, 659 F.3d 527, 541 (6th Cir. 2011) ("[S]earching the entire server was necessary . . . because individuals often mislabel directory files, the server might contain related websites, and the unallocated server space might contain materials pertaining to those websites."); United States v. Stabile, 633 F.3d 219, 239 (3d Cir. 2011) (finding reasonable a search of a filesharing directory, under suspicion of child pornography, even though the scope of the warrant was "limited to evidence of financial crimes" because "criminals can easily alter file names and file extensions to conceal contraband"); United States v. Williams, 592 F.3d 511, 522 (4th Cir. 2010) ("[A] computer search must, by implication, authorize at least a cursory review of each file on the computer . . . ."); United States v. Burgess, 576 F.3d 1078, 1094 (10th Cir. 2009) ("[I]n the end, there may be no practical substitute for actually looking in many (perhaps all) folders and . . . documents . . . ."); *Hill II*, 459 F.3d at 978 ("There is no way to know what is in a file without examining its contents . . . ." (quoting *Hill I*, 322 F. Supp. 2d at 1090) (quotation marks omitted)). Of course, warrants that fail even to link the search to a particular crime will generally be invalidated. *See, e.g.*, United States v. Rosa, 626 F.3d 56, 61–62 (2d Cir. 2010) (invalidating a warrant that "failed to state with any level of particularity the specific criminal activity alleged or the type of digital evidence to be sought");

Since then, attempts have been made—with varying degrees of success—to rein in computer searches both ex ante and ex post, but it is worth examining, first, some of the problems that arose in this context.

II. TWO PROBLEMS AND SOME SOLUTIONS

Courts and commentators have tended to focus their ire on the problem of the plain view doctrine—its potential for licensing overbroad, general searches in the context of ESI.[51] If law enforcement may lawfully view every file, evidence of *any* crime discovered will be admissible, as long as its incriminating character is immediately apparent. But less visible and arguably more perturbing is the predicament of the innocent suspect. In such a case, the officer executing the search is, in effect, in the position of proving a negative—that the seized data does *not* contain incriminating evidence—and often has the power to search every file seized to satisfy that condition.[52]

A rigorous solution must address the root of the matter by cabining the permissible scope of ESI searches. This Part examines both problems in detail before exploring attempts to solve them, including scrutinizing the officer's subjective intent, abolishing the plain view doctrine altogether, and requiring a search protocol in the warrant.

A. The Plain-View Problem

The admissibility of any evidence discovered in plain view during an authorized search revitalizes concerns about the very "general searches" the particularity requirement was supposed to avert.[53] There is no shortage of "low-level offenses" for which probable cause

---

United States v. Riccardi, 405 F.3d 852, 863 (10th Cir. 2005) (finding a warrant that "permitted the officers to search for anything—from child pornography to tax returns to private correspondence"—to be overbroad).

51. *See infra* note 95 and accompanying text. *See generally* Weir, *supra* note 5 (analyzing the circuit split concerning the plain view doctrine).

52. *See Hill II*, 459 F.3d at 974 ("To be certain that the medium in question does *not* contain any [incriminating] material, the officers would have to examine every one of what may be thousands of files on a disk . . . ."). As discussed in Part II.C.3, magistrates have increasingly been imposing a search protocol ex ante, but that practice has come under constitutional attack from at least one prominent scholar. *See infra* notes 109–19 and accompanying text.

53. *See* Coolidge v. New Hampshire, 403 U.S. 443, 469–70 (1971) (plurality opinion) (noting the tension between the plain view doctrine and the prohibition on general warrants).

can be "relatively easy to establish," making it very easy to obtain a warrant to search an "unpopular or politically powerless" target.[54] The fear is that, having obtained such a pretextual warrant, police could conduct a general search, seizing evidence of any crime, whether it were in the warrant or not.[55]

In the physical world, the Supreme Court initially addressed such concerns by requiring a subjective test of the officer's intent: evidence unrelated to the justification for the search was inadmissible unless it came into plain view inadvertently.[56] But in *Horton v. California*,[57] the Court changed course and eliminated the subjective-intent element.[58] "Scrupulous adherence" to the particularity requirement, the Court held, was sufficient to obviate the possibility of general searches. The Court reasoned that as long as the warrant particularly described the place to be searched and the evidence sought, the officer was constrained by the corresponding permissible scope of the search.[59] Thus, even if the warrant was obtained by pretext, the Court reasoned that the officer could hardly use it to conduct a general search.[60]

Even in the physical realm, this reasoning only goes so far: a sufficiently small pretextual "thing" can authorize a thoroughly comprehensive search.[61] In the digital realm, courts routinely accept the assertion—ex post, at least—that *any* file may contain the evidence sought,[62] fully vitiating the logic of *Horton*.[63] The

54. Kerr, *supra* note 2, at 567 (citing William J. Stuntz, *The Pathological Politics of Criminal Law*, 100 MICH. L. REV. 505, 512–18 (2001)).

55. *Id.*

56. *Coolidge*, 403 U.S at 469. Despite not commanding a majority, Justice Stewart announced a rule that, within nine years, was largely accepted by lower courts. Linda Novak, Note, *The Precedential Value of Supreme Court Plurality Decisions*, 80 COLUM. L. REV. 756, 774 (1980).

57. For a summary of the facts, see *supra* notes 32–33 and accompanying text.

58. Horton v. California, 496 U.S. 128, 139–41 (1990).

59. *Id.* at 140 & n.10.

60. *See id.*

61. *See* Weir, *supra* note 5, at 93 (observing that "a warrant to search a house for stolen diamonds . . . allows the police to search everything in the house because of the diamond's small size").

62. For examples from the Third, Fourth, Sixth, Ninth, and Tenth Circuits, see *supra* note 50. Of course, those same courts maintain that searches must be "limited in scope by the terms of the warrant's authorization," United States v. Phillips, 588 F.3d 218, 223 (4th Cir. 2009), but this limitation has been given little content.

invasiveness of physical searches, however, is still bounded by the amount of time, money, and manpower that law enforcement can bring to bear.[64] Searches of seized ESI, however, encounter significantly attenuated limits: a single agent can conduct the entire search, and he may search it at his convenience, over the course of months.[65] Additionally, the agent has powerful tools at his disposal to organize, classify, and explore the data,[66] in contrast to the methodical way in which a physical search must proceed. Furthermore, physical searches often take place with the suspect or a third party present,[67] introducing an observer who can testify as to whether the search was reasonable, whereas computer searches generally occur off-site, where the only prying eyes are those of the police. Thus, in the digital realm, both doctrinal and practical hurdles to pretextual searches are substantially removed; the plain view doctrine allows "an end-run around the particularity requirement" of the Fourth Amendment.[68]

B. Innocent Suspects

The plain-view problem in the context of ESI is highly visible to the judiciary because there is evidence of a crime and a defendant to challenge the search.[69] But the lack of limits on searches of a suspect's

63. For a discussion of one court's attempt to rein this in, see *infra* Part II.C.1. For a discussion of search protocols, which significantly alleviate the problem but may be unconstitutional, see *infra* Part II.C.3.

64. Kerr, *supra* note 2, at 569 (observing that when searching a physical location, "[a] search team must be organized and trained; the location must be controlled during the execution of the search").

65. Weir, *supra* note 5, at 93 ("Instead of many officers searching a house in haste, a single analyst can peruse a hard drive extensively and at his leisure, providing a low-cost search without time constraints.").

66. *See* Robyn Burrows, Note, *Judicial Confusion and the Digital Drug Dog Sniff: Pragmatic Solutions Permitting Warrantless Hashing of Known Illegal Files*, 19 GEO. MASON L. REV. 255, 260 (2011) ("[Forensic] programs 'index' the imaged hard drive by organizing files into a searchable format. Using [forensic software], an examiner can perform keyword searches, recover deleted material, flag encrypted files, and analyze altered files."); Wayne Jekot, *Computer Forensics, Search Strategies, and the Particularity Requirement*, PITT. J. TECH. L. & POL'Y, Spring 2007, at 1, 45, *available at* http://tlp.law.pitt.edu/ojs/index.php/tlp/article/view/29/29 (discussing forensic software packages).

67. *See, e.g.*, Burrows, *supra* note 66, at 284–85 (noting that the presence of the individual being searched is a check on drug-dog sniffs).

68. Weir, *supra* note 5, at 87.

69. *See* Anthony G. Amsterdam, *The Supreme Court and the Rights of Suspects in Criminal Cases*, 45 N.Y.U. L. REV. 785, 787 (1970) (noting that the Supreme Court "review[s] the conduct of police almost exclusively in criminal cases where the defendant is the asserted victim of police misconduct").

data should be more troubling when the suspect is actually innocent, because the natural stopping point for the search is the last file on the computer. The Tenth Circuit case of *United States v. Walser*[70] helps illustrate the issue.

In *Walser*, the police had probable cause to search the defendant Walser's hotel room for "evidence of the possession of controlled substances," and received a warrant to that effect.[71] The officer who executed the search found a computer, and began exploring its contents in the room.[72] He found and opened some JPEG picture files, apparently suspecting that they might contain "images of drug use."[73] Upon discovering that they were, in fact, adult pornography, he seized the computer "in anticipation of conducting a more thorough search" off-site.[74]

Once off-site, the officer navigated to the documents directory of a spreadsheet program, ostensibly continuing the search for "address books, spreadsheets, [and] databases" that might provide evidence of drug transactions.[75] Out of "approximately ninety files and four sub-folders," however, he chose to open a movie file and discovered that it contained child pornography.[76] Walser argued that opening the movie exceeded the scope of a warrant for records of drug transactions; the government countered that a computer search warrant authorized the search of every file.[77]

The court was unwilling to announce quite so broad a rule. Seeking to avoid suppression, it emphasized the "restraint" the officer showed by subsequently requesting a second warrant rather than continuing to "rummage"—but implicitly, it held that opening a movie file was within the scope of a warrant for evidence of drug transactions.[78] This is a stretch, but even assuming that the officer

70. United States v. Walser, 275 F.3d 981 (10th Cir. 2001).

71. *Id.* at 983. The magistrate issuing the warrant was not actually aware of the presence of a computer, *id.* at 983 n.1, but the warrant authorized, among other things, searching for "records, and/or receipts, written or electronically stored, income tax records, checking and savings records, records that show or tend to show ownership or control of the premises and other property used to facilitate the distribution and delivery [of] controlled substances," *id.* at 984 (alteration in original).

72. *Id.* at 984.

73. *Id.*

74. *Id.*

75. *Id.*

76. *Id.* at 984–85.

77. *Id.* at 987.

78. *See id.*

expected to find that Walser had conveniently taped himself participating in a drug deal—and was not merely curious about Walser's pornography tastes—it should be troubling when one considers: What if Walser had been innocent?[79]

When a suspect is innocent, there is *no* contraband to be found. However, having seized *all* of his data, there is no check to prevent zealous law enforcement from expanding an initially fruitless search into increasingly less pertinent areas, in search of evidence that is not there. If privacy is to have any meaning, this cannot be reasonable,[80] and in theory, law enforcement is disincentivized from doing so by the threat of suppression.[81] But, as *Walser* shows, in the electronic domain that threat is simply not forthcoming. Ex post review is highly deferential when an expanding search does eventually uncover evidence of some crime.[82] On the other hand, as long as the suspect is innocent, he will never know the extent of the intrusion to challenge it. The predicament of the innocent suspect—that his privacy is regulated by rules developed in the context of guilty suspects—is a familiar one, but it is profoundly magnified in the digital domain by the power of broad overseizure.

C. Solutions

The problem with the plain view doctrine in the digital realm—that it removes the penalty for a general rummaging search—is but a symptom of the larger malaise: the police have the ability and, generally, the authority to rifle through an entire hard drive in search of evidence. As demonstrated by the problem of the innocent suspect, this is an unacceptable incursion on privacy interests[83]—even when a fishing expedition is not the motive. This state of affairs has not escaped judicial and critical notice. This Section describes and critiques the three solutions that have been proposed. The first two

79. He was not: marijuana and related paraphernalia were found in the room. *Id.* at 984. It is unclear whether he was ever prosecuted on that score.

80. *Cf.* Akhil Reed Amar, *The Future of Constitutional Criminal Procedure*, 33 AM. CRIM. L. REV. 1123, 1133 (1996) ("The Constitution seeks to protect the innocent. The guilty, in general, receive procedural protection only as an incidental and unavoidable byproduct of protecting the innocent *because* of their innocence.").

81. *See supra* notes 34–36 and accompanying text.

82. For further examples, see *supra* note 50 and accompanying text.

83. *See* Raphael Winick, *Searches and Seizures of Computers and Computer Data*, 8 HARV. J.L. & TECH. 75, 81 (1994) ("The variety of information commonly stored on a computer, and the enormous and ever-expanding storage capacity of even simple home computers, justifies the highest expectation of privacy.").

tackle only the plain view doctrine; the third attempts to address the core privacy concern, but runs into constitutional and normative difficulties.

1. *Inadvertence Requirement*. An early response that focused on the plain-view problem was to effectively reinstate the subjective-intent element for evidence discovered in plain view during searches of ESI. The oft-cited precedent in this line of decisions is *United States v. Carey*.[84] In *Carey*, the defendant was being investigated for possession and sale of cocaine.[85] The police obtained a warrant to search his computer for information relating to drug transactions.[86] The agent was unable to find any relevant files through keyword searches, but did notice some image files.[87] He opened the first one, discovered it contained child pornography, "developed probable cause to believe the same kind of material was present on the other image files," and proceeded to open several others without obtaining a new warrant.[88] The court held that his discovery of the first file was inadvertent and therefore admissible, but, because he then "temporarily abandoned" the authorized search for evidence of drug trafficking and embarked on an unauthorized search for child pornography, the court suppressed the latter files.[89]

In reinstating the inadvertence requirement for plain-view digital evidence, the court recognized that, in the digital context, the reasoning of *Horton* fails—the plain view doctrine can indeed "be used to extend a general exploratory search from one object to another until something incriminating at last emerges."[90] However, *Horton*'s criticism of the inadvertence requirement remains apt for three main reasons. First, focusing on the "subjective state of mind of the officer" is still a poor way to achieve evenhanded law

84. United States v. Carey, 172 F.3d 1268 (10th Cir. 1999); *see also* United States v. Mann, 592 F.3d 779, 784 (7th Cir. 2010) (giving some weight to the officer's inadvertence in affirming the lower court's admission of evidence under the plain view doctrine); United States v. Schlingloff, No 11-40073, 2012 WL 4378148, at *4 (C.D. Ill. Oct. 23, 2012) (interpreting *Mann* as requiring inadvertence).

85. *Carey*, 172 F.3d at 1270.

86. *Id.*

87. *Id.* at 1271.

88. *Id.*

89. *Id.* at 1283 & n.4.

90. *See id.* at 1272 (quoting Coolidge v. New Hampshire, 403 U.S. 443, 466 (1971) (plurality opinion)).

enforcement,[91] and the scope of a valid search is better defined by the object of the search.[92] Second, *Horton* is still good law, and *Carey* is hard to square with it.[93] Finally, the plain-view problem is the symptom but not the disease—even with an inadvertence requirement, the police may still scour all of a suspect's data as long as they demonstrate their intention not to rummage by getting a second warrant when they find something unexpected.[94]

2. *Abolishing the Plain View Doctrine.* A second approach, proposed by Chief Judge Kozinski of the Ninth Circuit and some scholars, is to eliminate the plain view doctrine entirely for electronic searches.[95] This is strong medicine that cuts deeply into the societal interest in crime control, and needs correspondingly strong justification.[96] The idea seems to be that if the only evidence usable in a prosecution is that related to the justification for the search, then a fishing expedition cannot possibly yield any dividends. This justification is unrealistic, may even be counterproductive, and still fails to address the base privacy concerns.

First, given the reality of broad overseizure and off-site search, the cost of an exploratory search is quite low,[97] but the returns can be enormous. Even if unrelated evidence cannot be used directly, the police are now aware of it. That bell cannot be unrung. As with all

---

91. Horton v. California, 496 U.S. 128, 138 (1990). For example, given two searches that are identical from the point of view of the suspect, the admissibility of plain-view evidence would depend on the inherently unreliable determination of the officer's subjective state of mind. In addition, a subjective-intent inquiry would have the perverse effect of excluding evidence which the police had *some* expectation of finding, but not on evidence sufficient to constitute probable cause.

92. *Id.* at 139–40.

93. Indeed, the Tenth Circuit itself may be retreating from this doctrine. *See* United States v. Burgess, 576 F.3d 1078, 1092 (10th Cir. 2009) (suggesting that *Carey* be limited to its facts).

94. *See supra* notes 71–79 and accompanying text.

95. *E.g.*, United States v. Comprehensive Drug Testing, Inc., 621 F.3d 1162, 1180 (9th Cir. 2010) (en banc) (Kozinski, C.J., concurring) ("Magistrate judges should insist that the government waive reliance upon the plain view doctrine in digital evidence cases."); Kerr, *supra* note 2, at 576–84 (arguing that "the best way to neutralize dragnet searches is to rethink the plain view exception in the context of digital evidence"); Weir, *supra* note 5, at 113 ("[C]ourts should act as the Ninth Circuit did and abolish the doctrine's application to digital searches.").

96. *See* Lily R. Robinton, *Courting Chaos: Conflicting Guidance from Courts Highlights the Need for Clearer Rules To Govern the Search and Seizure of Digital Evidence*, 12 YALE J.L. & TECH. 311, 344 (2010) ("Abolishing the plain view doctrine with respect to digital searches may create risks to society that outweigh those created by governmental intrusion into individual privacy.").

97. *See supra* notes 64–68.

applications of the exclusionary rule, if the purpose of the search is something other than eventual prosecution—such as embarrassment or harassment—excluding plain-view evidence would have no effect on police incentives.[98] Alternately, the police may use that evidence to further an investigation, as long as the evidence eventually introduced in court is sufficiently attenuated or might inevitably have been discovered.[99] Thus, a pretextual warrant can pay significant dividends.

Second, abolishing the plain view doctrine may actually be counterproductive. The advantages for law enforcement of the plain view doctrine are obvious, but it is good for privacy interests, too. The plain view doctrine incentivizes police to constrain their search: the reward for "scrupulous adherence" to the permitted scope of search is the admission of any evidence not in the warrant but found in plain view.[100] If the police are "undeterred by a potential loss of plain-view evidence," there will be no concern about invasiveness to balance the zealous pursuit of evidence.[101] Indeed, for this reason, abolishing the plain view doctrine may be *worse* for innocent suspects.

Third, regardless of the motives of the police—whether the warrant is legitimate or pretextual—abolishing the plain view doctrine does nothing to address the basic problem of protecting legitimate expectations of privacy after broad overseizure. This is also why a related proposal, that data irrelevant to the investigation should be sequestered by a third party, preferably independent of the government,[102] is not useful—the privacy violation occurs regardless of who is employing the violator.[103] To be fair, however, Chief Judge Kozinski's proposal included a very privacy-protective doctrinal modification: search protocols.

---

98. *See* William J. Stuntz, *Warrants and Fourth Amendment Remedies*, 77 VA. L. REV. 881, 918 n.79 (1991) ("When the motivation for the police conduct is not evidence-gathering, the exclusionary rule imposes no cost on the police officer, and consequently cannot deter misconduct.").

99. Murray v. United States, 487 U.S. 533, 536–37, 539 (1988).

100. *See* Horton v. California, 496 U.S. 128, 139–41 (1990). The question of the permissible scope of a search of ESI is, of course, fundamental and contested. Under current rules, *see supra* Part I.B, the scope is everything seized.

101. Weir, *supra* note 5, at 105.

102. United States v. Comprehensive Drug Testing, Inc., 621 F.3d 1162, 1180 (9th Cir. 2010) (en banc) (Kozinski, C.J., concurring).

103. When combined with a search protocol, as in Chief Judge Kozinski's proposal, such a rule makes some sense, as it "keep[s] as many eyes off non-seizable information as possible." *See* Weir, *supra* note 5, at 104.

3. *Search Protocols*. In 1994, Professor Raphael Winick, noting that searches of ESI are likely to involve "large quantities of personal information . . . intermingled with relevant information," presciently proposed that warrants include a search protocol: "an outline of the methods that [investigators] will use to sort through the information."[104] *Carey*, though best known for reintroducing the inadvertence requirement, also quoted extensively from Winick's work,[105] and in the past decade, federal magistrate judges around the country have increasingly required that a search protocol be attached to a computer warrant.[106] In 2010, Chief Judge Kozinski urged magistrate judges to "insert[] a protocol for preventing agents involved in the investigation from examining . . . data other than that for which probable cause is shown."[107]

At first glance, search protocols make a lot of sense. Because they preclude forensic officers from trawling through the entirety of the data looking for evidence, search protocols prevent the plain view doctrine from effectively authorizing general searches in the digital world, much as the particularity requirement prevents the plain view doctrine from doing so in the physical world.[108] For the same reason, search protocols also protect the innocent suspect from the exposure

104. Winick, *supra* note 83, at 107–08.

105. United States v. Carey, 172 F.3d 1268, 1275–76 (10th Cir. 1999).

106. *See, e.g.*, United States v. Ganias, No. 3:08CR00224(AWT), 2011 WL 2532396, at *7 (D. Conn. June 24, 2011), *appeal docketed*, No. 12-240 (2d Cir. Jan. 20, 2012); *In re* U.S.'s Application for Search Warrant To Seize and Search Elec. Devices from Edward Cunnius, 770 F. Supp. 2d 1138, 1152–53 (W.D. Wash. 2011); *In re* Search of 3817 W. West End, 321 F. Supp. 2d 953, 961 (N.D. Ill. 2004); *In re* Search Warrant, 2012 VT 102, ¶ 7. The search protocol in *Ganias* is typical, and authorized the following techniques:

> (a) surveying various file "directories" and the individual files they contain (analogous to looking at the outside of a file cabinet for the markings it contains and opening a drawer believed to contain pertinent files);
>
> (b) "opening" or cursorily reading the first few "pages" of such files in order to determine their precise contents;
>
> (c) "scanning" storage areas to discover and possibly recover recently deleted files;
>
> (d) "scanning" storage areas for deliberately hidden files; or
>
> (e) performing key word searches through all electronic storage areas to determine whether occurrences of language contained in such storage areas exist that are intimately related to the subject matter of the investigation.

Ganias, 2011 WL 2532396, at *7.

107. *Comprehensive Drug Testing*, 621 F.3d at 1179.

108. Chief Judge Kozinski proposed to include search protocols *as well as* to eliminate the plain view doctrine and require third-party segregation of data. *Id.* at 1180; *see* Weir, *supra* note 5, at 103–05. Such an approach would arguably confer *stronger* privacy protection in the digital realm than in the physical realm.

of all of his data to government eyes. Unlike the other attempted solutions, search protocols address the privacy problem directly.

Kerr has argued, however, that specifying a search protocol in the warrant is neither constitutionally permissible nor good policy.[109] His constitutional argument relies on four cases, in nonelectronic but plausibly analogous settings, holding, respectively, that warrants need not *necessarily* specify a method of execution;[110] that agents may disregard an express knock-and-announce requirement—that is, a protocol for the search—so long as their actions are reasonable;[111] that the particularity requirement extends only to the place to be searched and the property to be seized;[112] and that a magistrate judge, having issued a warrant for probable cause and with particularity, may not further involve himself in its execution.[113] Taken together, these cases at least cast doubt on the putative binding effect of a search protocol; Kerr believes that they "point to the conclusion that the Fourth Amendment does not permit ex ante restrictions on the execution of computer warrants."[114]

Kerr further argues that search protocols, despite their laudable goal of protecting individual Fourth Amendment interests,[115] are in fact poor policy. He argues that search protocols are essentially ex ante attempts to regulate reasonableness,[116] which, in the absence of concrete facts, must be little more than guesses as to what *will* be reasonable.[117] In fact, he reasons, search protocols can inhibit the development of rules for reasonable computer searches ex post, by

109. Kerr, *supra* note 3, at 1246.

110. *See id.* at 1264–66 (discussing *Dalia v. United States*, 441 U.S. 238 (1979)).

111. *See id.* at 1268–71 (discussing *Richards v. Wisconsin*, 520 U.S. 385 (1997)).

112. *See id.* at 1267–68 (discussing *United States v. Grubbs*, 547 U.S. 90 (2006)).

113. *See id.* at 1261–64 (discussing *Lo-Ji Sales v. New York*, 442 U.S. 319 (1979)).

114. *Id.* at 1271. However, the only court to examine this argument thus far has found that these cases support not the proposition that search warrants are unconstitutional, but "two more modest conclusions: that ex ante evaluation by a judicial officer cannot wholly supplant ex post assessment of law enforcement conduct and that hard and fast rules about what a warrant must and must not include are generally frowned upon." *In re* Search Warrant, 2012 VT 102, ¶ 24 n.11. For another very narrow reading of these cases, see generally Paul Ohm, *Massive Hard Drives, General Warrants, and the Power of Magistrate Judges*, 97 VA. L. REV. IN BRIEF 1 (2011), http://www.virginialawreview.org/inbrief/2011/03/20/ohm.pdf.

115. *See* Kerr, *supra* note 3, at 1247 (stressing that his argument is "about means rather than ends").

116. *Id.* at 1277.

117. *See id.* at 1279 ("[The reasonableness] standard requires courts to 'slosh [their] way through the factbound morass' . . . ." (second alteration in original) (quoting Scott v. Harris, 550 U.S. 372, 383 (2007))).

focusing litigation on compliance with the protocol rather than directly on the reasonableness of the search.[118] Thus, even if search protocols are constitutionally permissible, he asserts, magistrate judges should avoid them.[119]

III. PARTICULARITY AND PERSPECTIVE

Kerr has proposed abandoning ex ante limitations on computer searches altogether, leaving the problem of protecting privacy to the requirement that searches be reasonable.[120] But that requirement has thus far been wholly vacuous.[121] Proponents of search protocols have the right idea—enunciating ex ante bounds to prevent the privacy violation before it occurs. The problem is that search protocols are the wrong implementation—they regulate *how* a search is executed, rather than *where* and for *what*. The Constitution explicitly contemplates a kind of ex ante restriction, which computer searches have been entirely lacking: particularity.[122]

This Part argues that the key to vindicating privacy interests in the new digital reality lies in the Fourth Amendment requirement that warrants specify a particular place to be searched. The first Section explores the particular need for particularity as an ex ante check on police action in the digital realm. The second Section evaluates different ways of conceptualizing data, with particular attention to conceptualizing data as *places*. The final Section proposes a new way to think about data—the *semantic* perspective, which conceives of data as consisting of domains of *meaning*. It then examines the rules that emerge from this perspective, explores the advantages of this approach, and attempts to anticipate some objections.

---

118. *Id.* at 1289.

119. *Id.*

120. *See id.* at 1247 ("[L]imitations . . . on the execution of computer warrants . . . should be developed and identified in ex post challenges.").

121. *See supra* Part I.B. Notably, Kerr's proposed rules for reasonable searches would eliminate the plain view doctrine. Kerr, *supra* note 3, at 1287.

122. *See* Ohm, *supra* note 114, at 4 ("[Search protocols] are designed to cure the *manifest lack of probable cause and particularity* in almost every computer case."); *see also In re* Search of 3817 W. West End, 321 F. Supp. 2d 953, 958–59 (N.D. Ill. 2004) (noting the special particularity concerns for computer searches because of broad overseizure).

A. Particularity is Particularly Necessary

The constitutional requirement of particularity in the warrant, an ex ante limit on the scope of a search, has doctrinal and normative functions that are particularly important in the context of computer warrants. Doctrinally, it explicitly requires ex ante limits customized to the particular case at hand in addition to the general requirement of reasonableness. Normatively, it guides the reasonable execution of searches by informing and alerting all the relevant actors in a given search—the police, the judiciary, and the suspect—to the particular privacy interests at stake. Because police can seize all of a suspect's data and search it later, both are crucial.

1. *Particularity and Doctrine*. The Fourth Amendment, "even more than its fellows, . . . was the product of particular events that closely preceded the Constitution and the Bill of Rights."[123] The primary thrust of the Amendment was prohibiting general searches, implemented by the requirement of particularity in the warrant.[124] Indeed, in contrast to the modern focus on the requirement that searches be reasonable, state constitution precursors to the Fourth Amendment "seem[] to show that the general principle [that searches must be reasonable] was stated merely as a basis for the minor premise condemning general warrants and that the abuse attempted to be prevented was that of general warrants only."[125] And yet, as one scholar has noted, "[c]omputer search warrants are the closest things to general warrants we have confronted in the history of the Republic."[126] Computer warrants specify neither *what* to seize nor *where* to search.[127]

Rules of reasonableness, developed ex post at trial or on review, are a general sort of rule—they apply to *any* search that presents the

123. TELFORD TAYLOR, TWO STUDIES IN CONSTITUTIONAL INTERPRETATION 19 (1969).

124. *See supra* notes 21–23 and accompanying text.

125. LASSON, *supra* note 22, at 81 n.10; *see also id.* at 79 n.3 (quoting the Virginia Bill of Rights' precursor to the Fourth Amendment, VA. DECLARATION OF RIGHTS of 1776, art. X, which only prohibited general warrants and made no mention of reasonableness); *id.* at 81 n.12 (quoting the Maryland clause, MD. DECLARATION OF RIGHTS OF 1776, art. XXVI, which also prohibits general warrants without mentioning reasonableness); *id.* at 101–03 (noting that the House of Representatives never voted on the current phrasing of the Fourth Amendment, and the version that was voted upon did not prohibit unreasonable search and seizure per se).

126. Ohm, *supra* note 114, at 11.

127. *See supra* Part I.B.

predicate fact-pattern during its execution.[128] The particularity requirement is different: it generates rules that are *tailored* to the case at hand and cabin the search *before* it is executed. The point of the particularity requirement is to define the outer bounds of the search using the information available ex ante, that is, the facts establishing probable cause.[129] If that information can support a more particular description of the place or the thing for which there is probable cause, then the warrant must be more particular.[130] Particularity, probable cause, and the magistrate constitute the essential ex ante trifecta of the Fourth Amendment—a "neutral and detached" third party, not "engaged in the often competitive enterprise of ferreting out crime,"[131] ensuring that there is "a fair probability that contraband or evidence of a crime will be found *in a particular place*."[132]

Normally, "scrupulous adherence" to the particularity requirement in the warrant—and the corresponding permissible scope of the search—justifies the plain view doctrine.[133] This is why doctrinal changes that aim to "solve" the plain-view problem are operating at the wrong level: the root of the problem is that particularity in computer search warrants has thus far been missing.[134] It is not sufficient that computer warrants merely state the evidence sought or the crime committed. Given the reality of broad overseizure and the absence of conventional checks in an off-site search of ESI,[135] that leaves entirely too much to the discretion of the searching agent, who may explore as much of the data as is necessary

128. *See* Kerr, *supra* note 3, at 1279 ("For each set of facts, the courts articulate what the officers can do and cannot do as they execute the warrant.").

129. *See* Maryland v. Garrison, 480 U.S. 79, 84 (1987) (holding that a particular warrant must "limit[] the authorization to search to the specific areas and things for which there is probable cause to search").

130. *See supra* notes 24–27 and accompanying text; *see also* United States v. Ross, 456 U.S. 798, 824 (1982) ("[P]robable cause to believe that a stolen lawnmower may be found in a garage will not support a warrant to search an upstairs bedroom . . . ."); Marron v. United States, 275 U.S. 192, 196 (1927) ("As to what is to be taken, nothing is left to the discretion of the officer executing the warrant.").

131. United States v. Johnson, 333 U.S. 10, 14 (1948).

132. Illinois v. Gates, 462 U.S. 213, 238 (1983) (emphasis added).

133. *See supra* notes 53–60 and accompanying text.

134. *See In re* Search of 3817 W. West End, 321 F. Supp. 2d 953, 962–63 (N.D. Ill. 2004) ("[W]hat the government seeks is a license to roam through everything in the computer without limitation and without standards. Such a request fails to satisfy the particularity requirement of the Fourth Amendment . . . .").

135. *See supra* notes 64–68 and accompanying text.

to find the evidence sought.[136] But if the *thing* to be seized cannot be specified with any particularity, computer warrants must limit the *place* within ESI that the agent may search. Otherwise, the plain view doctrine will become a license for general searches and fishing expeditions in the digital realm.[137]

2. *Normative Function*. The Supreme Court has explained that in physical searches, particularity helps the suspect understand the justification for the search and its corresponding bounds.[138] In the context of digital evidence, in which broad overseizure is the norm, the police and the judiciary, too, must be aware of the justifiable extent of the search. Accordingly, in addition to the doctrinal need for particularity in the warrant, particularity plays a normative role that is critical in the digital realm.

First, particularity alerts law enforcement to the fact that although a broad array of information has been seized, the investigation has a substantially narrower scope, and the suspect has not relinquished his legitimate privacy interests in information outside that scope.[139] Short of eliminating broad overseizure, this is the single greatest protection that can be afforded the innocent suspect, who will almost never be able to vindicate his privacy rights before a judge.[140] Rather than weigh an abstract notion of privacy against the need to expand the search to find the expected evidence, the officer will have an independent, objective judgment as to how far, concretely, the privacy violation is justified.[141]

Second, particularity is an aid to the judiciary ex post in evaluating whether the scope of the search was truly *reasonable*.

---

136. *Cf.* LASSON, *supra* note 22, at 54 ("The writ [of assistance] empowered the officer and his deputies and servants to search, at their will, wherever they suspected uncustomed goods to be, and to break open any receptacle or package falling under their suspecting eye.").

137. Accordingly, search protocols, for all their flaws, can be viewed as an attempt to regulate the *place* that can be searched. *See* Ohm, *supra* note 114, at 9–10 (noting that the Fourth Amendment requires that the place to be searched be particularly described and arguing that search protocols satisfy this requirement).

138. Groh v. Ramirez, 540 U.S. 551, 561 (2004).

139. *See* Ohm, *supra* note 114, at 5 (observing that probable cause typically only exists for a small portion of the seized data); *id.* at 7–8 (noting that computers store increasing amounts of increasingly sensitive data).

140. *See supra* note 69 and accompanying text.

141. *See, e.g.*, United States v. Ganias, No. 3:08CR00224(AWT), 2011 WL 2532396, at *3 (D. Conn. June 24, 2011), *appeal docketed*, No. 12-240 (2d Cir. Jan. 20, 2012) (discussing a complicated chain of custody for Ganias's data, wherein each recipient was notified of the appropriate zone of search authorized by the magistrate).

When, as is usually the case for ex post review, evidence of *some* crime was found, the objective judgment of the proper scope ex ante helps to balance the natural ex post bias in favor of law enforcement.[142] It provides a critical backdrop against which to evaluate the steps law enforcement took[143]—or, alternately, vindicates those steps by virtue of additional warrant requests that serve as a coarse public record.[144]

Third, "[a] particular warrant also 'assures the individual whose property is searched or seized of the lawful authority of the executing officer, his need to search, and the limits of his power to search.'"[145] Specifically for computer warrants, particularity helps assure the suspect that despite the broad overseizure of his data, his every word and deed is not open to examination by the authorities. Particularity limits, supported by probable cause, vetted ex ante by a magistrate, are a cornerstone of the lawful use of police power; when the actual seizure is far broader than the "things" for which there is probable cause, the importance of particularity in searching within those things is correspondingly much greater. The obvious question, then, is what well-formed, constitutional, ex ante particularity limits look like in the electronic world.

B. Perspectives on ESI

To talk coherently about limits, it is necessary first to decide what ESI looks like—how it should be perceived. It takes little imagination to conceive of data as comprising a place, or rather,

142. *See* Kerr, *supra* note 3, at 1291 ("[E]x ante review of probable cause and particularity ensures that the assessment of the government's interest is unbiased by the eventual discovery of evidence or contraband in the place to be searched." (citing Stuntz, *supra* note 98, at 916, 934)).

143. *See, e.g.*, United States v. Comprehensive Drug Testing, Inc., 621 F.3d 1162, 1172 (9th Cir. 2010) (en banc) (per curiam) (marking the government's "callous disregard" of a warrant's search protocol and limited scope).

144. *See, e.g.*, United States v. Burgess, 576 F.3d 1078, 1094–95 (10th Cir. 2009) ("Hughes immediately closed the gallery view when he observed a possible criminal violation outside the scope of the warrant's search authorization and did not renew the search until he obtained a new warrant."); *Ganias*, 2011 WL 2532396, at *6 (observing, using a search protocol as a proxy for place-particularity, that the agents "viewed only data that had been extracted according[]" to the search protocol, and that, when the agents needed to expand their search, they applied for a second warrant).

145. Groh v. Ramirez, 540 U.S. 551, 561 (2004) (quoting United States v. Chadwick, 433 U.S. 1, 9 (1977)).

multiple places, not all of which must be searched.[146] Most people have asked themselves, at some point: "*Where* did I put that file?" The harder question is: What do those places look like? The choice of perspective ripples throughout the application of Fourth Amendment doctrine, impacting the definition of a search,[147] the zone of a search,[148] and, importantly, the language available for a magistrate to articulate limits ex ante. There are essentially three perspectives that have been propounded thus far: the filesystem perspective, the exposure-based perspective, and the physical perspective.

1. *The Filesystem Perspective*. The most intuitive way to think about computer data is in terms of how the filesystem and operating system structure it.[149] Under this conception, the physical hard drive is a container; it contains folders and files, each of which is a separate container that may contain more containers.[150] This perspective is attractive because it comports with the physical metaphors that computers use to represent data: "files," "folders," "the desktop," and so on.[151] In addition to the intuitive familiarity of this perspective, it is attractive because it appears to allow traditional Fourth Amendment doctrine to transfer with minimal changes to the digital realm.[152] This promise has not been borne out.

The basic idea is that if data is a series of nested containers, then existing Fourth Amendment doctrine concerning closed containers should apply in the digital realm without more.[153] Thus, a search

146. *See* Jekot, *supra* note 66, at 35 ("Computer storage devices do not contain just one place; they hold multifarious data, such as metadata and user and system files and folders, in numerous small spaces, including bits and bytes and slack and unallocated spaces.").

147. *See supra* note 13 and accompanying text.

148. *See* Kerr, *supra* note 2, at 554 ("The zone of a search determines the extent to which a particular search in a space eliminates privacy protection elsewhere in that space."). This is distinct from the scope of a search. *See supra* notes 28–29 and accompanying text.

149. This has variously been called "the subcontainer perspective," Goldfoot, *supra* note 2, at 118–20, and the "virtual file" approach, Kerr, *supra* note 2, at 554–57.

150. Goldfoot, *supra* note 2, at 119.

151. *See* ANDY RATHBONE, WINDOWS XP FOR DUMMIES 20 (2d ed. 2004) ("You can create files and folders right on your new electronic desktop . . . .").

152. *See* Goldfoot, *supra* note 2, at 123–24 (noting efforts "to render the existing physical rules abstract, and then use them to govern forensic examiners' work"); Ohm, *supra* note 114, at 5 (describing the common Fourth Amendment analogy from computers to filing cabinets).

153. *See* Kerr, *supra* note 2, at 555 ("If you analogize a computer hard drive to a suitcase, each file is like its own individual zippered pocket in the suitcase.").

occurs when a container (a folder or a file) is opened;[154] the zone of the search is the entire container, so having opened a file, for example, the police may view as much of it as they would like;[155] and—herein lies the rub—the only language available to limit a search ex ante is the language of files and folders. Of course, the police are not usually familiar with the suspect's file organization or naming scheme ex ante, and accordingly warrants simply elide ex ante limits altogether.[156] But the structure of the filesystem has little to do with the probable cause for the search—viewed through the lens of files and folders, there is probable cause for *every* subcontainer.[157] Thus, the filesystem perspective is the source of the plain-view problem.[158]

Furthermore, many files are *themselves* containers, such as database files,[159] "page files,"[160] and Windows Registry files,[161] so they have high evidentiary value *because* they are repositories for multifarious data from different sources.[162] For instance, if a website is

154. *See* United States v. Carey, 172 F.3d 1268, 1273 (10th Cir. 1999) (holding that opening files can expand the scope of a search).

155. *See* Kerr, *supra* note 2, at 556 (observing that the filesystem perspective would permit an officer to expose every page of a hundred-page open document).

156. The argument for eliding limits is similar to the one for broad overseizure. *See, e.g.*, United States v. Hill (*Hill II*), 459 F.3d 966, 978 (9th Cir. 2006) ("Forcing police to limit their searches to files that the suspect has labeled in a particular way would be much like saying police may not seize a plastic bag containing a powdery white substance if it is labeled 'flour' or 'talcum powder.'"); United States v. Hay, 231 F.3d 630, 637 (9th Cir. 2000) ("The government knew that Evans had sent 19 images directly to Hay's computer, but had no way of knowing where the images were stored.").

157. *Hay*, 231 F.3d at 635 ("[T]he inquiry . . . is whether there was reasonable cause to believe the 19 files from Evans's computer were located *somewhere* in Hay's computer . . . ." (emphasis added)).

158. *See* Goldfoot, *supra* note 2, at 149 (observing that the filesystem perspective, with the plain view doctrine and without search protocols, turns ESM into "a single unit, in practice incapable of meaningful subdivision").

159. *See, e.g.*, *14.5. The MyISAM Storage Engine*, MYSQL 5.5 REFERENCE MANUAL, http://dev.mysql.com/doc/refman/5.5/en/myisam-storage-engine.html (last visited Sept. 9, 2013) (explaining that MySQL, a popular database program, using MyISAM, a popular storage engine, stores the database as three files, one of which contains all the data, which is subdivided into tables).

160. *See* MARK E. RUSSINOVICH & DAVID A. SOLOMON, WINDOWS INTERNALS PART 1, at 15 (6th ed. 2012) ("Because most systems have much less physical memory than [they need], the memory manager transfers . . . some of the memory contents . . . to disk [to] free[] physical memory so that it can be used for other processes or for the operating system itself.").

161. Goldfoot, *supra* note 2, at 127.

162. *See id.* at 129–30 ("One file can mix a drop of responsive data into a sea of unresponsive material—just as a hard drive can.").

stored as a database and there is probable cause to search the activity of one user, the filesystem perspective does not differentiate between searching his data and searching the entire database; as long as the database is one file,[163] probable cause to search some of the file is indistinguishable from probable cause to search all of the file.[164]

Finally, ESM includes data that is not stored in files and folders at all, but either as *part* of the filesystem or *outside* of it. The first type of data is known as metadata,[165] and includes the file name, owner, and creation and access times.[166] The second is unallocated space, which will often include remnants of deleted files.[167] Both of these can be highly valuable to the forensic examiner, but the filesystem perspective simply breaks down when considering them.[168] Thus, the filesystem perspective is not merely inadequate but in fact incoherent for describing zones of privacy on ESM.

2. *The Exposure-Based Perspective*. Kerr has proposed an exposure-based conception of searches of ESI that rectifies some—but not all—of the problems with the filesystem perspective. Under this conception, exposure of data to human observation is a search, and the zone of a computer search is coterminous with the extent of the information exposed.[169] Thus, every action that exposes new information—such as scrolling down a spreadsheet or querying a table from a database—is a new search.[170] This approach has the advantage of capturing searches that expose metadata or deleted files, as well as properly treating files that are themselves subcontainers as "containing distinct zones of privacy."[171] However, it does not solve

163. *See id.* at 130 ("A SQL database, holding all of a dynamic web site's data, might be a single file.").

164. *See supra* note 155 and accompanying text.

165. *See* DOMINIC GIAMPAOLO, PRACTICAL FILESYSTEM DESIGN 10 (1999) ("[Metadata is] information about the file that is not in the stream of bytes that make up the file.").

166. *Id.*; *see* Goldfoot, *supra* note 2, at 128–29.

167. Goldfoot, *supra* note 2, at 128.

168. *Id.*

169. Kerr, *supra* note 2, at 556–57.

170. *Id.* The distinction between the zone of the search and the scope of the warrant becomes important here; *most* of the time, a warrant that authorizes searching part of a file will authorize searching all of it. *Id.* at 557.

171. *See* Goldfoot, *supra* note 2, at 116; *see also* United States v. Comprehensive Drug Testing, Inc., 621 F.3d 1162, 1180–81 (9th Cir. 2010) (en banc) (Bea, J., concurring in part and dissenting in part) (criticizing officers for not being more selective in the portions of a spreadsheet they viewed).

the particularity problem—it provides no new vocabulary to describe zones of privacy ex ante.

3. *The Physical Perspective*. Goldfoot has recently proffered a third perspective, under which a hard drive is viewed as nothing more than another physical object, like a blood sample or a pair of jeans.[172] It is not *searched* at all; once lawfully seized, it is merely "examined."[173] Under this perspective, concerns like zone of search and particularity simply fade away. ESI is treated not as distinct from ESM,[174] but merely as physical properties of ESM that happen to reveal information.[175] Thus, the examiner need not "worry[] about whether his next mouse click will violate the Bill of Rights."[176]

The physical perspective is conceptually perhaps the cleanest perspective on ESI thus far, but that cleanliness comes at the cost of abandoning any restrictions on searches of seized ESM. Goldfoot contends that calls for such restrictions are essentially policy arguments rather than arguments about what level of protection is constitutionally required.[177] But unlike the private information that jeans can carry, like the wear level on a right-side pocket,[178] much of the data ESM carries are not epiphenomena of existence; they are intentionally created works, intentionally stored there.[179] A single one-by-one-half-inch USB stick could store its owner's journal, rolodex, calendar, to-do list, shopping list, "bucket" list—indeed, lists of every shape and form—library card, entire libraries, music, films, receipts, correspondence, accounts and finances, photo albums—all sorts of "papers" and "effects"[180]—and the physical perspective would

172. *See* Goldfoot, *supra* note 2, at 150 (providing an admirable exegesis of the variety of information that can be deduced from a pair of jeans).

173. *Id.* at 157.

174. As Goldfoot acknowledges, when the suspect's ESM is imaged on-site, as this Note has assumed throughout, the physical perspective is less clear about what constitutes a seizure. *Id.* at 158–60.

175. *Id.* at 155.

176. *Id.* at 157.

177. *See id.* at 160 ("At some point, the debate between the subcontainer and physical perspectives becomes a public policy debate."); *id.* at 166 (arguing that the "increase[d] . . . threats to public safety" posed by the growing use of computers "might warrant a change in law enforcement's favor").

178. *See id.* at 150 ("A worn right pocket suggests [the owner] favors that hand.").

179. *See* Ohm, *supra* note 114, at 8 ("Our computers track what we read, buy, where we go, and increasingly, *what we think*." (emphasis added)).

180. *See* U.S. CONST. amend. IV (providing protection for "papers, and effects, against unreasonable searches and seizures"); Winick, *supra* note 83, at 81 ("The intangible nature of

present the entire corpus to law enforcement for their perusal, on probable cause for any pretextual crime.[181] This is not a policy argument; the policy choice is enshrined in the Fourth Amendment.[182]

Even for inadvertently created data, which might be likened to the information that can be gleaned from a pair of jeans, the sheer scale of the potential privacy invasion makes ESM qualitatively different without more.[183] That is particularly true when one considers that ESM often stores both intentionally and inadvertently created data of other parties not suspected of crime—perhaps hundreds of other parties[184]—a fact that is becoming increasingly salient with the rise in cloud computing, wherein one's "email messages, word processing documents, voice mail messages, and business data [are commingled] on shared servers alongside the data of innumerable strangers."[185] Finally, the physical perspective treats computers, cell phones, and so on, as simply self-contained pieces of plastic and silicon—but they are networked machines, and searching one can reveal information well outside its physical boundaries.[186] The physical perspective is thus no more coherent than the filesystem perspective.

C. The Semantic Perspective

On the one hand, the physical perspective conceives of ESM as one monolithic zone of privacy, wholly forfeited once seized, despite widely held expectations of privacy that society is almost certainly

computer data does not affect the analysis, since the Court has long recognized that the Fourth Amendment protects 'intangible as well as tangible evidence.'" (quoting Warden v. Hayden, 387 U.S. 294, 305 (1967))).

181. *See supra* note 55 and accompanying text.

182. *Cf.* District of Columbia v. Heller, 554 U.S. 570, 636 (2008) ("[T]he enshrinement of constitutional rights necessarily takes certain policy choices off the table.").

183. Consider, for example, inadvertently created records like chat logs, which store verbatim every careless word typed to a friend. *Chat History*, GOOGLE CHAT HELP, http://support.google.com/chat/bin/answer.py?hl=en&answer=161925 (last visited Sept. 9, 2013).

184. It is telling that the Ninth Circuit's highly privacy-protective framework was conceived in the context of a system that stored sensitive data concerning "hundreds of players in Major League Baseball (and a great many other people)," for only ten of whom the government had probable cause to search. *See* United States v. Comprehensive Drug Testing, Inc., 621 F.3d 1162, 1166 (9th Cir. 2010) (en banc) (per curiam).

185. Ohm, *supra* note 114, at 7 (quoting *Comprehensive Drug Testing*, 621 F.3d at 1176) (quotation marks omitted).

186. *See, e.g.*, Washington v. Roden, 279 P.3d 461, 463 (Wash. Ct. App. 2012) (discussing a police officer's impersonation of a suspect by using a seized cell phone to send text messages).

"prepared to recognize as reasonable."[187] On the other hand, the filesystem perspective recognizes that seized data holds multiple zones of privacy, but it attempts to describe those zones using a poor proxy—files and folders. This Section proposes a new perspective, the semantic perspective, that does away with the proxy. When the domain of search is information, the particular place within it must be described semantically.

1. *Semantic Zones*. The semantic perspective is directly responsive to the problem with broad overseizure, in that *all* of one's content is seized, but only some *types* of content could reasonably contain the evidence sought.[188] Accordingly, a semantic zone is the set of areas on a hard drive that responds to a particular semantic description—in other words, those areas that contain a particular type of content. Thus, semantic descriptions are articulated not in structural or technical terms, but in natural language, as descriptions of content. For example, *image-related data*, as applied to a particular hard drive, describes a semantic zone: it is the set of areas on the drive that contain image-related content, including image files, metadata for image files, application data for image-editing software, fragments of image files in unallocated space, and so on.[189] All of these things, in terms of human meaning, are image-related data.

The *meaningfulness* of a semantic description[190] can vary broadly, and semantic zones can be nested and overlap. For example, *guilt-related data* is a conceivable semantic zone—the set of areas on the hard drive containing evidence of a suspect's guilt—but finding that set of areas requires a very deep understanding of the meaning of the data, either via human inspection or a perfect forensic tool.[191] On the other hand, *image-related data* is much less meaningful and accordingly much more amenable to automated extraction.[192] The meaningfulness of a given semantic description thus has important

187. *See* Katz v. United States, 389 U.S. 347, 361 (Harlan, J., concurring) (quotation marks omitted).

188. *See supra* notes 179–81 and accompanying text.

189. A search continues to be defined as exposure of data to human observation under this perspective, as such exposure almost always involves exposing the *meaning* of the data.

190. "Meaningfulness," here, refers to the extent to which the description refers to the underlying meaning of the data.

191. *See* Kerr, *supra* note 2, at 570 (discussing a hypothetical "Perfect Tool" that could "magically locate evidence described in a warrant").

192. For examples of tools that can automatically extract data responsive to a semantic description, see *infra* notes 199–205 and accompanying text.

implications for its workability, but for now, the key point is that semantic descriptions describe zones on a hard drive that, in the language of the Fourth Amendment, can operate as a "place to be searched."[193]

2. *The New Rules of Particularity*. A computer warrant that meets the particularity requirement should describe one or more semantic zones for which there is probable cause. For example, when the probable cause is for child pornography, the warrant might authorize a search for *image-related data*; evidence of tax fraud would support a warrant for *spreadsheet data*; evidence of unauthorized access, *source code and shell scripts*; and so on.

The police should not search—that is, expose to human observation—data outside the authorized semantic zones.[194] The suspect retains a legitimate expectation of privacy in those semantic zones not related to the investigation. Thus, for example, a warrant for *spreadsheet data* will support searching Registry entries concerning Microsoft Excel, but not necessarily adjacent entries, and certainly not fragments of image files in unallocated space or the creation time of a movie file. On the other hand, because semantic zones constrain computer warrants and searches—just as particular places do in the physical realm—in a manner tailored to the information available ex ante, the plain view doctrine remains viable in the digital realm.[195]

There is an inherent tension in these rules: How are the police to know what parts of the hard drive correspond to a given semantic description without looking at them?[196] The definition of a Fourth Amendment search in the digital realm is exposure to *human* observation.[197] The corollary is that law enforcement is free to employ unlimited *automatic* tools to analyze ESI without running afoul of the

193. *See* U.S. CONST. amend. IV.

194. This is not dissimilar to the rule governing the reasonable scope of physical searches, but here, by virtue of the semantic zone, some authority is shifted to the magistrate. *See infra* notes 213–15 and accompanying text.

195. For a practical example, see *infra* notes 263–67 and accompanying text. In general, when a warrant specifies a particular semantic zone, as long as the officer "scrupulous[ly] adhere[s]" to that limitation, the plain view doctrine cannot be used to conduct a general search. *See supra* note 133 and accompanying text.

196. *Cf., e.g.*, United States v. Williams, 592 F.3d 511, 521 (4th Cir. 2010) ("[T]he warrant impliedly authorized officers to open each file on the computer and view its contents, at least cursorily, to determine whether the file fell within the scope of the warrant's authorization.").

197. *See supra* note 13 and accompanying text.

Fourth Amendment.[198] There is an abundance of such tools—tools that sort and categorize data without human intervention—that law enforcement can use: keyword search,[199] file-header "magic tests" that determine a file's format regardless of filename,[200] natural language search,[201] hash matching,[202] image signature recognition,[203] optical character recognition (OCR),[204] and many more.[205] The ability of computers to automatically segregate data by the type of content it represents is only improving.[206] Law enforcement can use such tools to

198. An important question, outside the scope of this Note, is whether law enforcement can run certain automatic analyses on all seized data regardless of why it was seized; for example, whether they can check every hard drive they seize for known child pornography. *See infra* note 202. The Supreme Court has said that "some quantum of *individualized suspicion* is usually a prerequisite to a constitutional search or seizure," United States v. Martinez-Fuerte, 428 U.S. 543, 560 (1976) (emphasis added), but also that "[o]fficial conduct that does not compromise any legitimate interest in privacy is not a search," Illinois v. Caballes, 543 U.S. 405, 408 (2005) (quotation marks omitted), and that "governmental conduct that only reveals the possession of contraband compromises no legitimate privacy interest," *id.* at 408 (quotation marks omitted). *See generally* Burrows, *supra* note 66 (analogizing child-pornography dragnets to dog-sniff searches per *Caballes*). This Note, however, concentrates on the use of automatic searches to *limit* privacy violations outside the parameters of probable cause.

199. *See* Winick, *supra* note 83, at 108 (advocating keyword searches to limit the scope of a warrant).

200. *See FILE(1)*, FREEBSD GENERAL COMMANDS MANUAL (Oct. 9, 2008), http://www.freebsd.org/cgi/man.cgi?query=file&manpath=FreeBSD+9.0-RELEASE (last visited Sept. 9, 2013) (using a magic number in the header of a file to divine a file's format, even if it has been misnamed).

201. *E.g.*, *About Wolfram|Alpha*, WOLFRAM|ALPHA, http://www.wolframalpha.com/about.html (last visited Sept. 9, 2013) (accepting search queries in "[f]ree-form natural language input"); *Siri*, APPLE, http://www.apple.com/ios/siri (last visited Sept. 9, 2013) ("Ask Siri to do things just by talking the way you talk."); *cf.* Goldfoot, *supra* note 2, at 138 (asserting that automated techniques cannot catch "unanticipated wording, an egregious misspelling, an unexpected foreign language, [or] recently invented slang").

202. *Forensic Toolkit User Guide*, ACCESS DATA 25 (Oct. 2, 2012), https://ad-pdf.s3.amazonaws.com/FTK4-1_UG.pdf (describing the "Known File Filter," which "compare[s] file hashes in a case against a database of hashes" to either eliminate irrelevant files or pinpoint known contraband, like child pornography).

203. *E.g.*, *Frequently Asked Questions*, TINEYE, http://www.tineye.com/faq#how (last visited Sept. 9, 2013) (using a "unique and compact digital signature or 'fingerprint'" to match images).

204. *See Forensic Toolkit User Guide*, *supra* note 202, at 72 ("The Optical Character Recognition (OCR) process lets you extract text that is contained in graphics files. The text is then indexed so that it can be[] searched[] and bookmarked.").

205. *See, e.g.*, *id.* at 73 (detailing an Explicit Image Detection feature that scores files according to their likelihood of containing "possibly illicit content").

206. *See, e.g.*, Samy Bengio, *Large-Scale Visual Semantic Extraction*, *in* NAT'L ACAD. ENG'RS, FRONTIERS OF ENGINEERING: REPORTS ON LEADING-EDGE ENGINEERING FROM THE 2011 SYMPOSIUM 61 (2012) (presenting an algorithm to describe an image in natural language from a dictionary of one-hundred thousand or more terms); Hao Tang, Vivek Kwatra, Mehmet Emre Sargin & Ullas Gargi, *Detecting Highlights in Sports Videos: Cricket as a Test*

determine, without executing a Fourth Amendment search, the areas of ESM that are responsive to a given semantic description.[207]

The fact that semantic zones are not continuous should be of no concern; neither are files. The filesystem breaks files up into blocks, which can be placed anywhere on the disk—a "file" is simply an abstraction provided by the filesystem to present the data to programs and to the user as a single continuous piece of data.[208] The filesystem perspective is one superstructure for organizing the arbitrary block-level layout of the data.[209] The semantic perspective is simply another superstructure, but one that has the fortuitous property of being able to describe zones of privacy—semantic zones—ex ante.

Semantic descriptions should be construed narrowly,[210] but law enforcement can broaden the search by applying for a second

---

Case, *in* PROCEEDINGS OF THE 2011 IEEE INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO 1, 1 (2011) (presenting an algorithm to automatically "detect[] highlights in sports videos" in a sport-neutral way using a variety of machine-learning techniques (emphasis omitted)). *See generally* Ajay J. Joshi, Fatih Porikli & Nikolaos P. Papanikolopoulos, *Scalable Active Learning for Multiclass Image Classification*, 34 IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE 2259 (2012) (presenting techniques to train large image-classification systems with minimal training samples).

207. This should not be read as an argument in the form of "technological solutionism." *See generally* EVGENY MOROZOV, TO SAVE EVERYTHING, CLICK HERE: THE FOLLY OF TECHNOLOGICAL SOLUTIONISM (2013). Some semantic descriptions will be too "meaningful" for computers ever to effectively isolate responsive data, such as *guilt-related data*. *See supra* notes 190–93 and accompanying text. Magistrates should take care to stay abreast of developments in technology that can impact how meaningful a semantic description can workably be. Of course, if the authorized semantic zone turns out to be unworkably fine-grained, the investigating officer can always explain this in an affidavit requesting a broader warrant, *see infra* Part IV.A.3, or the officer may be protected by the plain view doctrine, *see infra* Part IV.B. The key point, once a minimum level of technology has been reached—which it has, *see supra* notes 199–206—is that semantic zones provide an operable and permissible approach to particularizing computer search warrants, regardless of the technology used or available.

208. *See* GIAMPAOLO, *supra* note 165, at 11 ("A file appears as a continuous stream of bytes at higher levels, but the blocks that contain the file data may not be contiguous on disk.").

209. As a practical matter, the semantic zone *filenames* should always be implicitly authorized, to allow forensic tools to present the responsive areas in a meaningful way. Similarly, if the facts of the case require exposure of data stored in a file-level subcontainer, such as a database, the structural information in that file (for databases, the *schema*) should also be authorized to allow automated queries to meaningfully return narrow portions of the file. *See, e.g.*, HECTOR GARCIA-MOLINA, JEFFREY D. ULLMAN & JENNIFER WIDOM, DATABASE SYSTEM IMPLEMENTATION 2, 14 (2000) (explaining that a schema is "a description of the structure of the data in a database").

210. For example, a warrant authorizing *image data* ought not to be construed to include *video data*, although it might include animated GIF files.

warrant. On the one hand, narrow construction incentivizes law enforcement to use the best tools available to sequester nonresponsive private data, deploying their powerful forensic tools to protect, rather than compromise, a suspect's privacy.[211] On the other hand, if information outside the authorized semantic zone is relevant to the investigation, if ESM cannot be carved up as finely as the magistrate envisioned, or if searching within the authorized zone reveals evidence of another crime, the police should simply apply for a second warrant, supported by probable cause, to expand the search to a new (and still particular) semantic zone.[212]

The corollary of this rule is a more ongoing, supervisory role for the magistrate—a role necessary in the presence of broad overseizure. In the physical realm, the police must make real-time decisions about the scope of a search, checked for reasonableness ex post; that level of discretion is both necessitated and justified by the real-time nature of the search. In the digital realm, the converse is true. On the one hand, the off-site nature of a search of seized ESI makes it *possible* to impose greater limits on the scope ex ante, because those limits can be tweaked as the facts of the case develop.[213] On the other hand, to guard against abuse of the awesome power of broad overseizure, it becomes *necessary* to interpose "a neutral and detached magistrate" between "zealous officers" and a decision to expand the zone of

211. Semantic zones effectively require law enforcement to use the closest-available approximations to Professor Kerr's "Perfect Tool." *See* Kerr, *supra* note 2, at 570. Accordingly, it is the *semantic* breadth of the description, rather than the technical, filesystem-level breadth, which should be construed narrowly. By contrast, the technical dimension of semantic zones should be construed broadly—if a given file is responsive to a semantic zone description, so is its metadata, the configuration or temporary cache data for the application that created it, and so on, thus avoiding some of the pitfalls of the filesystem perspective. *See supra* notes 165–68 and accompanying text.

212. Goldfoot complains about "empty formality" second-warrant requirements that generate affidavits such as: "I saw child pornography on that hard drive; therefore, I submit there is probable cause to believe there is child pornography on that hard drive." Goldfoot, *supra* note 2, at 145. This type of second warrant should be rare with semantic zones, because if one image falls within the authorized semantic zone, the rest likely do as well. *See infra* notes 264–67 and accompanying text.

213. Whereas "exigent circumstances" justify eliding the warrant requirement and concomitant objective review of probable cause, Horton v. California, 496 U.S. 128, 137 n.7 (1990), off-site searches present positively *leisurely* circumstances. *See In re* Search of 3817 W. West End, 321 F. Supp. 2d 953, 961 (N.D. Ill. 2004) ("[W]hen the government wishes to search a computer hard drive in the controlled environment of a laboratory, it is not confronted with a rapidly evolving and sometimes dangerous situation that must be addressed on the spot.").

search.[214] Effectively, some of the real-time discretion in whether to search for "an adult elephant . . . in a chest of drawers"[215] is withdrawn—no longer checked for reasonableness ex post, but limited in the warrant ex ante, subject to expansion with probable cause. Thus, the off-site nature of the search is converted from a liability for privacy interests into an asset.

3. *Objections*. At the outset, it must be noted that *any* limit on police investigations will mean that some crimes go undetected. That is the price of having a Fourth Amendment—that is, of barring general searches and protecting innocent suspects. The semantic perspective draws the boundary of a search, as is traditional, around data for which law enforcement has probable cause.[216] In any case, as long as the suspect is not innocent, the semantic-zones approach largely mimics the computer forensic methods already in use by law enforcement; the first stroke of an investigation is often to separate the relevant from the irrelevant.[217] Additionally, the ability to obtain a new warrant if probable cause is established for an additional semantic zone ensures that law enforcement is hardly hamstrung. Semantic zones, therefore, protect innocent suspects from a thoroughgoing search of their data far more than they prevent the discovery of evidence.

Goldfoot, however, asserts that automated techniques are insufficient to segregate data for which there is probable cause.[218] His argument is mostly targeted toward simplistic techniques such as

214. Johnson v. United States, 333 U.S. 10, 13–14 (1948). On the distinction between the second-warrant applications discussed here and the magistrate's participation in *Lo-Ji Sales, Inc. v. New York*, 442 U.S. 319 (1979), see *infra* note 230.

215. Platteville Area Apartment Ass'n v. City of Platteville, 179 F.3d 574, 579 (7th Cir. 1999). There are significant operational similarities between the semantic-zones approach and the reasonable scope of a physical search. The key observation in both situations is that law enforcement can get some idea of what an area might reasonably contain *before* searching it, and that regulates their discretion to search it. *See supra* notes 28–29 and accompanying text.

216. *See* Arizona v. Hicks, 480 U.S. 321, 329 (1987) (asserting that crime control and privacy should be balanced using "the textual and traditional standard of probable cause").

217. *Cf.* Burrows, *supra* note 66, at 260–61 (noting that once a forensic software loads an image of a hard drive, it indexes the data along various axes including file type, keyword, and so on); *Forensic Toolkit User Guide*, *supra* note 202, at 76–77 (describing different ways to refine a search of digital evidence). For example, in *United States v. Ganias*, the "first attempted search . . . yielded too many results for a practicable review," leading the agents to "narrow the search of the data" for reasons entirely unrelated to privacy concerns. United States v. Ganias, No. 3:08CR00224(AWT), 2011 WL 2532396, at *4 (D. Conn. June 24, 2011), *appeal docketed*, No. 12-240 (2d Cir. Jan. 20, 2012).

218. Goldfoot, *supra* note 2, at 137–38.

keyword and filename searches, but modern forensic tools are far more sophisticated, increasingly capable of identifying the meaning represented on a given area of ESM.[219] For semantic zones to work, all that is necessary is that a minimum level of technology be available to automatically classify data.[220] Of course, magistrates should keep in mind the changing capability of computers to understand human meaning when writing semantic descriptions,[221] but any miscalculations on that score can be dealt with by applying for a new warrant[222] or under the plain view doctrine.[223]

Goldfoot further asserts that, even with sophisticated tools, the art of computer forensics is not amenable to mechanization, "because forensics is detective work . . . [which] involves applying background knowledge, intuition, and professional judgment."[224] Happily, the semantic-zones approach withdraws none of these faculties from forensic examiners. A semantic zone includes not only areas that are directly responsive to the zone description, but also areas that hold ancillary data like metadata and configuration data.[225] The only restriction is that inferential leaps that take the examiner outside the zone of established probable cause must be vetted by a magistrate.[226]

4. *Semantic Zones Versus Search Protocols*. Semantic zones avoid the constitutional and normative hazards that, as Kerr pointed out, afflict search protocols.[227] Semantic zones, unlike search protocols, mandate no particular method of search, but merely

219. *See supra* notes 199–206 and accompanying text. In particular, the repeated canard that "much evidence could escape discovery simply because of [the defendant's] labeling of the files documenting [his] criminal activity," United States v. Adjani, 452 F.3d 1140, 1150 (9th Cir. 2006), simply does not hold water. None of the techniques enumerated above rely on the files' labeling. This is not like "saying police may not seize a plastic bag containing a powdery white substance if it is labeled 'flour' or 'talcum powder,'" United States v. Hill (*Hill II*), 459 F.3d 966, 978 (9th Cir. 2006), but more like requiring that the police test it with a machine to investigate whether it is cocaine, rather than take a bump.

220. *See supra* note 207.

221. For a concrete discussion in the context of a child pornography investigation, see *infra* notes 246–48 and accompanying text.

222. *See infra* Part IV.A.3.

223. *See infra* notes 263–67 and accompanying text.

224. Goldfoot, *supra* note 2, at 138.

225. *See supra* note 211.

226. *See* Johnson v. United States, 333 U.S. 10, 14 (1948) ("[T]he usual inferences which reasonable men draw from evidence [must, under the Fourth Amendment,] be drawn by a neutral and detached magistrate instead of being judged by the officer engaged in the often competitive enterprise of ferreting out crime.").

227. *See supra* notes 109–19 and accompanying text.

restrict the search to particular areas of the hard drive. It is only as a corollary of how those areas are identified (by human meaning) and how "searching" data is defined (exposure to human observation) that the forensic process (using automated techniques to locate the semantic zone) is constrained. Thus, there is no ex ante attempt to "guess what would be reasonable."[228] A semantic zone warrant restricts the search using only information available ex ante—the probable cause for the search in the first place.[229] Kerr's constitutional arguments are also largely inapplicable for the same reason—the Constitution *requires* ex ante restrictions on the "particular place" to be searched, and semantic zones fill that doctrinal and normative gap for searches of ESI.[230]

Much like semantic zones, search protocols aim to identify the class of data relevant to the investigation; unlike semantic zones, search protocols go on to dictate *how* agents may locate that class of data. This difference does have important ramifications. In *United States v. Ganias*,[231] for example, Ganias was not initially a suspect, but rather a third party whose computer may have contained tax data incriminating the targets of the investigation. Accordingly, the search protocol limited the search to data "intimately related to the subject matter of the investigation"[232]—essentially a very meaningful[233] semantic zone. The warrant also specified how the agents should find

228. Kerr, *supra* note 3, at 1277–78, 1287 (summarizing his basic normative argument against ex ante limits).

229. *Cf.* Kerr, *supra* note 3, at 1277 (decrying search protocols as "error-prone ex ante judicial review" whose utility diminishes as rules of reasonableness are developed ex post).

230. Kerr's discussion of *Lo-Ji Sales, Inc. v. New York* may require further distinguishing. *See* Kerr, *supra* note 3, at 1261–64. In *Lo-Ji Sales*, the local magistrate accompanied the police officers in the execution of the search of an adult bookstore, purporting to determine on-site which materials to seize for probable cause for obscenity. Lo-Ji Sales, Inc. v. New York, 442 U.S. 319, 322–23 (1979). The Supreme Court invalidated the warrant and seizure for two reasons: because the open-ended warrant was insufficiently particular and because the magistrate had abandoned his "neutral and detached posture" and "allowed himself to become a member, if not the leader, of the search party which was essentially a police operation." *Id.* at 325–27. By contrast, semantic zones *cut back* on the generality of computer warrants. Even though the magistrate is involved in a sort of supervision, that involvement is mediated through the traditional process of warrant applications. Indeed, multiple courts have expressly relied on the officer's having applied for a second warrant, upon discovering probable cause to search what was effectively a new semantic zone, in upholding the search. *See supra* note 144.

231. United States v. Ganias, No. 3:08CR00224(AWT), 2011 WL 2532396 (D. Conn. June 24, 2011), *appeal docketed*, No. 12-240 (2d Cir. Jan. 20, 2012).

232. *Ganias*, 2011 WL 2532396, at *7.

233. *Id.* at *7. For an explanation of "meaningfulness" in this context, see *supra* notes 190–93, 207 and accompanying text.

that data: manual techniques such as "cursorily reading the first few 'pages' of . . . files in order to determine their precise contents," as well as automated techniques such as "key word searches."[234]

On the one hand, such manual techniques allow very meaningful semantic zones to be specified, which can further protection of the suspect's privacy; automated techniques would likely not have been able to differentiate between Ganias's tax data and that of the entities under investigation. On the other hand, search protocols often involve manual techniques to locate relevant data—effectively, peeking at it—that provide little guidance to law enforcement and can severely undercut any privacy-protection goals. In addition, the search warrant here could be read to forbid advanced automated techniques not known to the magistrate—precisely the concern of scholars like Professor Kerr.[235]

IV. TEST CASES

At this point, some examples may help illustrate how semantic zones will work. The primary question ex ante will be what semantic zone ought to be authorized given the information available. The nature of the crime, changes in technology, the way the information is stored, and information that the police can gather without actually searching the data will all be factors in that determination. Ex post suppression litigation will likely focus on whether the evidence was actually within the authorized semantic zone, and if not, whether it falls within the plain view doctrine. This Part tackles each question in turn, using the facts of various real and hypothetical cases.

A. The Appropriate Semantic Zone

In *United States v. Storm*,[236] Storm's then-girlfriend informed the police that she had found child pornography in the defendant's recently viewed files.[237] The magistrate issued a warrant to search all

234. *Id.*

235. *See* Kerr, *supra* note 3, at 1287 ("Ex ante restrictions effectively delegate the Fourth Amendment to magistrate judges, transforming Fourth Amendment litigation away from an inquiry into reasonableness and towards an inquiry into compliance with the magistrate's commands.").

236. United States v. Storm, No. 3:11–cr–00373–SI, 2012 WL 3643845 (D. Or. Aug. 23, 2012).

237. *Id.* at *1.

of the ESM, which the court upheld.[238] Recall that the warrant must only authorize search of the particular area for which there is probable cause.[239] Under the semantic-zones approach, the question is: In what place—what semantic zone—was there probable cause to search?[240]

On the one hand, the entire hard drive would certainly have been too broad—there was no probable cause to search, for example, Storm's calendar or tax returns, because they could not reasonably have contained child pornography. Authorizing such a search would be akin to authorizing the search of a chest of drawers for an adult elephant.[241] On the other hand, Storm argued that the search should be limited to the files his girlfriend saw, his recently viewed files.[242] That would certainly have been too narrow; it was quite probable that he possessed more than what his girlfriend discovered. There was probable cause to suspect that, in his *image and video data*, he possessed child pornography—that is the appropriate semantic zone. That would authorize law enforcement to scan his hard drive for image and video headers and view any responsive files (including misleadingly named files),[243] fragments of files, recoverable deleted files, or files inside "compound files such as ZIP, email, and OLE files,"[244] but not to review his other documents for miscellaneous criminality.[245] *Storm* is a relatively simple case. Other fact patterns can give rise to harder questions, explored below.

1. *Changing Technology*. What about narrowing the semantic zone in *Storm* further, by limiting the search to *pornographic image and video data*? This is a workable semantic zone—modern forensic tools already have the capability to return only images likely to be

238. *Id.* at *11.

239. *See supra* notes 24–27.

240. *See supra* notes 193–94 and accompanying text.

241. *Cf.* Platteville Area Apartment Ass'n v. City of Platteville, 179 F.3d 574, 579 (7th Cir. 1999).

242. *Storm*, 2012 WL 3643845, at *11.

243. *See* Kerr, *supra* note 2, at 545 ("Software can locate image files . . . by searching for file headers characteristic of known types of picture files. . . . The file header remains unchanged regardless of the extension placed on the file, . . . [and] file header characteristics can be located in slack space or in partially deleted files . . . ."); *supra* note 200 and accompanying text.

244. *Forensic Toolkit User Guide*, *supra* note 202, at 60.

245. *See* United States v. Rosa, 626 F.3d 56, 61 (2d Cir. 2010) (suggesting that without such limits, officers "might review expense reports, income-related files and correspondence, and federal filing information in search of evidence of tax evasion").

pornographic.[246] An even narrower semantic zone could be conceived: *child-pornographic image and video data*. The potential for a semantically narrow description will increase with the ability of computers to understand human meaning, but magistrate judges should be aware of technological limitations when requiring semantically narrow zones; a too-meaningful semantic zone, like *guilt-related data*,[247] will not be workable. Here, although there is not yet technology that can specifically identify child-pornographic images while excluding legal pornography,[248] limiting the search to pornographic image data is technologically feasible and would avoid exposure of embarrassing but legally insignificant photos.

2. *Structured Data*. The potential for a semantically narrow description will also vary with the degree of structure the data is given. *United States v. Comprehensive Drug Testing, Inc.*[249] provides illustrative facts. The government had probable cause to believe that ten baseball players were using illegal steroids, but the records they sought were kept in a Microsoft Excel-format spreadsheet containing the drug-testing records of hundreds of other players "and a great many other people."[250] The warrant only authorized search of the records of the ten players.[251] Excel spreadsheets are basically unstructured—even though columns and rows may be labeled, there is no programmatic association between those labels and the associated data. That is, there is no way to "query" an Excel spreadsheet for information only about one person. Once the spreadsheet is open, all the data is visible. Thus, the semantic zone *drug-test results for specific players* was unworkable.[252]

---

246. *See Forensic Toolkit User Guide*, *supra* note 202, at 182 (explaining the Explicit Image Detection feature).

247. For a discussion of the variable meaningfulness of semantic descriptions, see *supra* notes 190–93 and accompanying text. For a discussion of technological solutionism and the generality of the semantic-zones approach, see *supra* note 207.

248. Of course, the technology does exist to identify *specific*, known images of child pornography. *See supra* note 202.

249. United States v. Comprehensive Drug Testing, Inc., 621 F.3d 1162 (9th Cir. 2010) (en banc) (per curiam).

250. *Id.* at 1166.

251. *Id.*

252. There are a few contingent facts that have been omitted for simplicity. As Judge Bea noted in his concurrence, large spreadsheets are usually not displayed in their entirety, and in this case, it was possible to "seize" the data in the authorized semantic zone without exposing the incriminating data as to other players. *Id.* at 1180–81. The search warrant also included a number of other procedural safeguards that the government brazenly ignored, such as initial

By contrast, modern relational-database systems store data in a structured way, using a user-defined *schema* that attaches labels to and establishes interrelationships within the data.[253] Such databases are not accessed by exposing the entire file, but rather through queries that selectively return information based on particular parameters.[254] Essentially, had the *Comprehensive Drug Testing* data been stored in a relational database, it would have been perfectly natural to expose data player by player,[255] and accordingly far more workable to authorize the narrow semantic zone *drug-test results for specific players*.

3. *Broadening the Search*. One common complaint from those who oppose restrictive warrants for digital evidence is that privacy-protective restrictions will allow clever criminals to evade detection. Goldfoot insists that automated techniques will not catch, for example, "pictures of documents."[256] But this is patently false; modern forensics software includes OCR technology that can recognize and extract text in images.[257] The harder question is what agents ought to do when, for example, in connection with tax fraud, a warrant authorizes searching *text-document and spreadsheet data*, but the clever fraudster has hidden his incriminating documents in images, which are off-limits.[258]

Suppose the agents develop a suspicion, based perhaps on the abundance of image files and dearth of incriminating data, that this is the case. According to the warrant, the agents may not search the image data—but that regulates exposure only to *human* observation. They may run a combination of OCR analysis and keyword search on the images, and any resulting match is new evidence with which they can request a second warrant to search *image-related data*. The general principle is that when the authorized semantic zone is unavailing but the government remains suspicious, automated

---

segregation of the data by personnel not involved in the case, to avoid exposure of data for which there was no probable cause. *Id.* at 1168–69. Thus, here, the results were suppressed.

253. GARCIA-MOLINA ET AL., *supra* note 209, at 2, 14.

254. *Id.*

255. *See* ROBERT SHELDON & GEOFF MOES, BEGINNING MYSQL 250–52 (2005) (describing how to select particular rows to view from a MySQL database).

256. Goldfoot, *supra* note 2, at 138.

257. *Forensic Toolkit User Guide*, *supra* note 202, at 72.

258. *See, e.g.*, United States v. Evanson, No. 2:05-CR-805-TC, 2007 WL 4299191, at *5 (D. Utah Dec. 5, 2007) (describing evidence stored in image files, albeit only because the suspect was "transforming his operation into a paperless type office").

analyses can be run on the rest of the data to justify broadening the search.

4. *Physical Crime, Digital Evidence*. Sometimes the connection between the suspected crime and the expected evidence will be more oblique than with tax fraud or child pornography. Drug trafficking is a good example—the evidence tends to be physical, such as "cash . . . , [m]arijuana, [c]ocaine, [m]ethamphetamine, or other illegal controlled substances, along with associated paraphernalia."[259] Some types of evidence that "tend to show conspiracy to sell drugs," however, such as "pay-owe sheets, address books, [and] rolodexes," are increasingly stored electronically.[260] As indicated in the Introduction, this type of evidence is likely to turn up in spreadsheets, text documents, and data files for electronic address books. But, having seized all of a suspect's data, agents may assert a need to search more attenuated semantic zones such as *image data*, averring that drug traffickers sometimes take "trophy photos," or "pictures of a person holding the controlled substance in front of a stack of money."[261] Obviously, looking through all of a suspect's photos entails a deep intrusion upon his privacy, and, in such cases, magistrates should evaluate the strength of the evidence presented against the breadth of the proposed semantic zone in making the probable-cause determination. At the end of the day, the magistrate has broad discretion as to how intrusive the search may be—and, as a "neutral and detached" third party, that is exactly with whom the discretion should lie.[262]

B. Suppression Litigation

The question whether a particular piece of evidence was found within the authorized semantic zone will be most contentious when the semantic zone is narrow, and either the technology is not able to draw such fine lines or the data is insufficiently structured. It is again useful to consider *Comprehensive Drug Testing*.[263] Suppose the

259. *See* United States v. Burgess, 576 F.3d 1078, 1083 (10th Cir. 2009) (quoting a search warrant).

260. *Id.* (quoting a search warrant).

261. *Id.* at 1084.

262. *See supra* note 226.

263. For the facts of the case, see *supra* notes 249–52 and accompanying text. For the purposes of this discussion, again, ignore the contingent facts discussed above, see *supra* note 252.

government opened the lawfully seized spreadsheet and the incriminating nature of hundreds of rows for hundreds of players was immediately apparent.[264] This is essentially an application of the plain view doctrine, and the key question will be whether the incriminating evidence lawfully came into the government's view.[265]

Much as in the traditional scope of search analysis, the test in this case should be reasonableness—whether the government could reasonably have viewed the ten rows without exposing the rest.[266] In an Excel spreadsheet, this is typically not the case. By contrast, had the data been stored in a relational database, it would have been trivial to craft a query that returned only the results for the players for which the government had probable cause, and if the government had chosen instead to dump the data en masse, that would have been unreasonable and the evidence should have been suppressed.[267]

CONCLUSION

As Justice Scalia wrote in *Arizona v. Hicks*, "[T]here is nothing new in the realization that the Constitution sometimes insulates the criminality of a few in order to protect the privacy of us all"; the appropriate question is "where the proper balance should be struck."[268] The arguments presented here could be adapted to call for an end to broad overseizure as well, substituting on-site, automatic assessment and seizure of responsive areas of ESM.[269] This Note, on the other hand, accepts that forensic analysts cannot effectively apply their "art" without the expediency of broad overseizure.[270] But the power to seize everything and search it later is both awesome and terrible, and if we have decided, as a society, to allow it in the interest of preventing crime, there must be an equally significant check to

264. As Judge Bea's opinion highlights, the incriminating nature was *not* immediately apparent—the agents had to expose more data first. United States v. Comprehensive Drug Testing, Inc., 621 F.3d 1162, 1180–81 (9th Cir. 2010) (en banc) (Bea, J., concurring in part and dissenting in part).

265. *See supra* notes 30–33 and accompanying text.

266. *See supra* notes 28–29 and accompanying text.

267. *See supra* notes 253–55 and accompanying text.

268. Arizona v. Hicks, 480 U.S. 321, 329 (1986).

269. At least one author has suggested something along these lines. *See* Jekot, *supra* note 66, at 46–47 ("[A] new 'best practices' search warrant should authorize the on-site search for a particular class or classes of data, and seizures of only the data that is relevant to the crime being investigated.").

270. *See* Kerr, *supra* note 2, at 547 ("[I]t is . . . difficult to plan a computer search ex ante; the search procedures are . . . more of an art than a science.").

ensure against its abuse.[271] Semantic zones are the right check. They are constitutionally grounded, effective, and responsive to the basic apprehension about broad overseizure: that once "back [in] the lab," the government will "have a good look around"[272] without being particularly—well, particular.

271. *See* McDonald v. United States, 335 U.S. 451, 456 (1948) ("Power is a heady thing; and history shows that the police acting on their own cannot be trusted.").

272. United States v. Comprehensive Drug Testing, Inc., 621 F.3d 1162, 1171 (9th Cir. 2010) (en banc) (per curiam).