Michigan Law Review

Volume 98 | Issue 2

1999

Zoning Speech on the Internet: A Legal and Technical Model

Lawrence Lessig
Harvard Law School

Paul Resnick
University of Michigan School of Information

ZONING SPEECH ON THE INTERNET: A LEGAL AND TECHNICAL MODEL

*Lawrence Lessig**
*Paul Resnick***

Speech, it is said,[1] divides into three sorts — (1) speech that everyone has a right to (political speech, speech about public affairs); (2) speech that no one has a right to (obscene speech, child porn); and (3) speech that some have a right to but others do not (in the United States, *Ginsberg*[2] speech, or speech that is "harmful to minors," to which adults have a right but kids do not). Speech-protective regimes, on this view, are those where category (1) speech predominates; speech-repressive regimes are those where categories (2) and (3) prevail.

This divide has meaning for speech and regulation within a single jurisdiction, but it makes less sense across jurisdictions. For when viewed across jurisdictions, most controversial speech falls into category (3) — speech that is permitted to some in some places, but not to others in other places. What constitutes "political speech" in the United States (Nazi speech) is banned in Germany; what constitutes "obscene" speech in Tennessee is permitted in Holland; what constitutes porn in Japan is child porn in the United States; what is "harmful to minors" in Bavaria is Disney in New York. Every jurisdiction controls access to some speech[3] — what we call "mandatory access con-

* Jack N. and Lillian R. Berkman Professor for Entrepreneurial Legal Studies, Harvard Law School. B.A./B.S. 1983, Pennsylvania; M.A. 1986, Cambridge; J.D. 1989, Yale. — Ed. Thanks to Lorrie Cranor for initially suggesting the symmetry between tagging speech and tagging people. Thanks also to Robert Cooter, Mark Lemley, Richard Posner, and the GALA Workshop at Boalt Law School for helpful comments on an earlier draft. Karen King and Alexander Macgillivray provided valuable research assistance. An earlier version of this paper has been published as "The Architectures of Mandated Access Controls," *in* COMPETITION, REGULATION AND CONVERGENCE: SELECTED PAPERS FROM THE 1998 TELECOMMUNICATIONS POLICY RESEARCH CONFERENCE (Sharon Eisner Gillett & Ingo Vogelsang eds., 1999).

** Associate Professor, University of Michigan School of Information. S.B. 1985, University of Michigan; S.M. (Electrical Engineering and Computer Science) 1988, Ph.D. (Electrical Engineering and Computer Science) 1992, M.I.T. — Ed.

[1] *See* Lawrence Lessig, *What Things Regulate Speech: CDA 2.0 vs. Filtering,* 38 JURIMETRICS J. 629, 638-39 (1998).

[2] *See* Ginsberg v. New York, 390 U.S. 629 (1968).

[3] We reserve the term "censorship" for blanket restrictions on the distribution of speech that apply regardless of the recipient or the context. Access control is a broader concept that includes not only censorship but also restrictions on speech that may depend on the recipient or context.

trols" — but what that speech is differs from jurisdiction to jurisdiction.

This diversity creates a problem (for governments at least) when we consider speech within cyberspace. Within cyberspace, mandated access controls are extremely difficult. If access control requires knowing (a) the identities of the speaker and receiver, (b) the jurisdictions of the speaker and receiver, and (c) the content of the speech at issue, then as cyberspace was initially designed, none of these data are easily determined. As a result, real space laws do not readily translate into the context of cyberspace.

One possible response to the change caused by the initial architecture of the Internet ("Net") would have been for governments simply to give up on access controls. Experience suggests that this is unlikely. As the popularity of the Net has grown, governments have shown an increasing interest in reestablishing mandated access controls over certain kinds of speech now published on the Internet. In the United States, this speech is sex-[4] or spam[5]-related; in Germany, it is both sex- and Nazi-related;[6] in parts of Asia, it is anything critical of Asian governments.[7] Across the world, governments seek to reregulate access to speech in cyberspace, so as to reestablish local control.

We take as given this passion for reregulation. It features prominently in the current political reality of cyberspace. This reality should push us to consider the options that regulators face — not because regulators need encouragement, but because we should understand the consequences of any particular regulatory strategy. Some strategies pose greater costs than others; some strike at more fundamental features of the Net than do others. We aim to understand the tradeoffs that this reregulation presents.

This inquiry is particularly salient in the United States just now. In what may have become a biannual event, the United States Congress in 1998 passed its second attempt at regulating "indecent speech" on the Net — the Child Online Protection Act[8] ("COPA"). Its first statute, the Communications Decency Act of 1996[9] ("CDA"), was struck

[4] *See* Communications Decency Act of 1996, § 502, 110 Stat. 133, *invalidated by* Reno v. ACLU, 521 U.S. 844 (1997); Child Online Protection Act, § 1403, 47 U.S.C.A. § 231 (Supp. 1999).

[5] "Spam" signifies unsolicited commercial email. *See infra* Part IV.

[6] *See* Kim L. Rappaport, *In the Wake of Reno v. ACLU: The Continued Struggle in Western Constitutional Democracies with Internet Censorship and Freedom of Speech Online*, 13 AM. U. INT'L L. REV. 765, 766-67, 788-90 (1998).

[7] *See* Geremie R. Barmé & Sang Ye, *The Great Firewall of China*, WIRED, June 1997, at 138, 147; Philip Shenon, *2-Edged Sword: Asian Regimes on the Internet*, N.Y. TIMES, May 29, 1995, at 1.

[8] Child Online Protection Act, 47 U.S.C.A. § 231 (Supp. 1999).

[9] Communications Decency Act of 1996, § 502, 110 Stat. 133, *invalidated by* Reno v. ACLU, 521 U.S. 844 (1997).

down by the Supreme Court in 1997.[10] Now two years later, a federal district court in Philadelphia has enjoined enforcement of COPA.[11] And if the ACLU succeeds in striking this statute, Congress no doubt will be at it again. Among the headaches of Y2K will be another CDA; and among the more significant (if repetitive) cases of 2001 will be *ACLU v. [the next attorney general]*.

Congress may never pass a statute that satisfies the Court,[12] but we think it could. There exists a type of "decency act," which we sketch here, that would pass constitutional muster. That act is not COPA. To see why this "decency act" would be constitutional where COPA was not, and to understand this alternative act, requires a broader view. It requires an analysis that makes clear the different values at stake.

Our aim in this essay is to provide just such a perspective. We offer in Part I a model of mandated access control that will clarify the issues in play. While this model will help resolve the constitutional questions raised by COPA, it will also help see the issues that mandated access controls present more generally. Given that different jurisdictions will want different restrictions, and given that those restrictions would be differentially costly, we provide in Part II a map of the different architectures and assignments of responsibility that might effect these restrictions. We then consider the trade-offs among these alternatives — both generally, and in particular in the American context.

This approach is a type of sensitivity analysis. Regulation, in the view that we take of it here, is a function of both law and the *architectures* of the Internet within which law must function. By "architectures" we mean (a) the Internet's technical protocols (for example, TCP/IP), (b) its standards and standard applications (for example, browsers or a digital certificate standard), and (c) its entrenched structures of governance and social patterns of usage that themselves are not easily changed — or at least not without coordinated action by many parties. These architectures are not fixed. They change, partly in response to both direct and indirect regulation by law. Thus in Part II we ask first how access can be controlled given the existing array of legal and architectural constraints. We then consider how changes in the current array might yield a different mix of costs and benefits.

We evaluate the various outcomes of these different legal and architectural choices along four separate dimensions. For any particular mix, we consider, first, the effectiveness at controlling access; second,

[10] *See* Reno v. ACLU, 521 U.S. 844 (1997).

[11] *See* ACLU v. Reno, 31 F. Supp. 2d 473 (E.D. Pa. 1999).

[12] A cynic might believe that this repetition is no accident. After all, Congress gets rewarded for what it passes, not what sticks. Protecting kids is great politics. Why do it only once, the cynic might ask, when one can do it every two years?

the cost to participants, whether sender, receiver, or intermediary; third, the costs to a system of "free speech" that such access controls impose; and fourth, other second-order effects, including in particular how different architectures might enable other regulation, beyond the specific access control that a given change was designed to enable.

For concreteness, we will focus on sexually explicit speech. We pick this type of speech because, in the American context at least, there exist at least two permissible levels of regulation for such speech. Some sexually explicit speech is prohibited generally (obscene speech, child porn); some sexually explicit speech is prohibited only to minors (speech that is "harmful to minors"); and the balance of sexually explicit speech is permitted to everyone.[13] This range of regulations will therefore illustrate the more general problem of access control across jurisdictions.

We then apply our model to COPA. COPA has a significantly narrower reach than the original CDA. Although Congress was, we believe, responsive to the Supreme Court's opinion in *Reno,* there is a structural feature of COPA that still renders it unconstitutional, at least when compared to a second possible statute that would have achieved Congress's legitimate end.[14] Those attacking COPA are not in a position to suggest this alternative, because they believe that private regulation is better than any law. But while we agree that private regulation may be better than COPA, we will suggest that private regulation may be more costly for free speech interests than the alternative regulation that we sketch here.

Part III focuses on this cost differential. There we consider the unintended consequences of the various regulatory strategies proposed. We argue that any reckoning of the costs of mandated access control must consider these secondary costs (and benefits) as well. In our view, these have been ignored in the debate so far. Yet arguably, they will be the most significant. Long after the "problem" of "indecent speech" is solved, the consequences of our choices to deal with indecent speech — these secondary effects — will continue to influence the culture of the Net. Legal and policy analyses would do well to account in the first place for these secondary effects.

The last section, Part IV, applies the same model to efforts to control "unsolicited commercial email," or "spam." The motivation for spam control differs from the reasons for controlling "indecency." Spam control protects recipients from unwanted information pushed into their mailboxes rather than preventing them from pulling information that they want. Our model and analysis, however, apply equally well to controlling spam, and shed light on the likely effective-

[13] See the cases cited in Lessig, *supra* note 1, at 638 nn.26 & 27.

[14] We describe this model below in Section II.D. Congress was aware of this alternative. *See* 144 CONG. REC. S12795 (daily ed. Oct. 21, 1998) (comments of Sen. Leahy).

ness and side effects of various legislative and architectural changes that have been proposed.

I. A Model of Access Control

A. *Elements*

In our model of mandated access control ("MAC"), we consider three relevant actors — a sender, a recipient, and an intermediary. The sender makes available the relevant speech; the recipient gets access to the relevant speech; and an intermediary is an entity that stands between the two. As these definitions suggest, nothing in our description hangs upon whether the sender actually *sends* material to the recipient, or upon the mode with which the recipient gains access.

These actors, we will assume, know different things about the speech that is to be regulated. We assume the sender knows about the contents of the item being sent. We assume the recipient has information about her own identity and residence. And finally we assume the intermediary has information neither about the content, nor about who the recipient is or where she resides. Obviously, these assumptions are not necessary. A sender might not have knowledge about the speech she makes available; and a recipient may not know where or who she is. But we assume a general case.

Given this mix of knowledge, a government effects mandated access control through four separate steps. It first defines which transactions are illegal, where "transaction" means the exchange of speech of a certain kind between two kinds of individuals. Second, it assigns responsibility to one or more actors to effect that restriction. Third, it creates a regime to detect when assigned responsibilities are violated. And fourth, it sets punishments for these violations. In the balance of this Part, we sketch issues relevant to each of these elements of a regulatory regime, and we conduct, for each element, a sensitivity analysis.

B. *Step One: Defining Blocked Exchanges*

A regulatory regime first defines a set of illegal transactions, or "blocked exchanges." The criteria for deciding whether an exchange is blocked include: (1) the type of speech item exchanged ("*I*"); (2) the recipient ("*R*"); and (3) the rules of the recipient's jurisdiction ("*J*"). We can state this relation as follows:

(a) Blocked Exchange: $B(I, R, J) = \{Y, N\}$

Where I = item type, R = recipient type, and J = jurisdiction type. $B(x,y,z)$ is a function determining whether exchange of the speech item is blocked. If the exchange is blocked, the function yields Y; if the exchange is not blocked, the function yields N.

Stated alternatively, a blocked exchange equates with access to a given item type, by a given individual within a given jurisdiction, that the law deems illegal.

Within this model, there may be "floor" recipients and "floor" jurisdictions. In the specific context of sexually explicit speech within American jurisdictions, children represent a type of floor recipient (anything permitted to children is permitted to adults as well), and a Bible Belt small town may be a floor jurisdiction (anything permitted there would be permissible everywhere). More formally, with J_f denoting a floor jurisdiction:

(b) Floor Recipient. For all I, J: $B(I, \text{child}, J) = N$ implies for all R, $B(I, R, J) = N$

(c) Floor Jurisdiction. For all I, R: $B(I, R, J_f) = N$ implies for all J, $B(I, R, J) = N$

The two floors can be combined. Anything that the law permits to children in a floor jurisdiction it will permit to everyone in every jurisdiction:

(d) Floor recipient and jurisdiction. For all I: $B(I, \text{child}, J_f) = N$ implies for all J and R, $B(I, R, J) = N$

In the general case, either the sender's or the recipient's jurisdiction may determine that an exchange is blocked. United States laws regulating cryptography, for example, restrict a sender's right to send certain encryption-related material to another jurisdiction; French cryptography laws regulate a receiver's right to receive such material. For simplicity, however, we will focus only on exchanges blocked by the recipient jurisdiction. Our analysis would apply with equal force if the exchange were blocked in the sender's jurisdiction because, aside from the effect on enforcement, the factors analyzed here do not depend on whose jurisdiction regulates.

A jurisdiction,[15] on this model of blocked transactions, may specify that a particular transaction must be blocked in at least two different ways:

1. The jurisdiction might publish criteria defining what must be blocked, but require a judgment by the parties about how to apply that criteria. The jurisdiction may or may not then hold parties responsible for correctly making such judgments prior to a determination by the regulating jurisdiction.
2. The jurisdiction might classify specific items as acceptable or blocked for particular recipient types, or, alternatively, create a list of prohibited speech. Determinations of acceptability could occur through a judicial or administrative process, or the jurisdiction could delegate its authority to an independent rating

[15] There is an important ambiguity in the concept of "jurisdiction" that we ignore here. Some rules depend upon where the person acts, rather than where the person is a citizen. If the drinking age in one state is 21, it does not matter that in the jurisdiction where X comes from, the drinking age is 18. But some rules may depend upon where someone comes from. We do not distinguish those cases in this version of the argument.

service.[16] A jurisdiction could even rely on a computer program to provide an initial classification of the speech at issue, and publish that classification as a preclearance, perhaps with a stipulation that the initial classification might be changed in the future after human review.

In the American context, jurisdictions ordinarily follow the procedure of case (1). If a jurisdiction follows case (2), publishing a list of blocked items for a given recipient type, then the list of items must, ordinarily, be judicially specified.[17] If, however, the lists are used on a voluntary basis for preclearance of acceptable items, nonjudicial determinations might be acceptable, a possibility we will analyze in Part II.

C. *Step Two: Assignments of Responsibility*

In the second step the regulator must define how best to allocate responsibility among actors to assure that access is controlled. In addition to the sender and recipient, it will sometimes be useful to distinguish among intermediaries. Internet Access Providers, such as AOL or AT&T WorldNet, serve as intermediaries closest to the senders and recipients. Internet backbone providers, such as WorldCom and Sprint, carry data between access providers. Responsibility for controlling access could be assigned either exclusively to one actor or jointly to any combination. We analyze only exclusive assignments of responsibility for blocking, as opposed to shared responsibility, though we do consider requiring other parties to provide information to the blocking party.

By hypothesis, no party knows enough to determine whether a particular exchange should be blocked.[18] The law must therefore create an incentive for parties to produce sufficient information to determine whether access should be blocked.

[16] An example would be Cyber Patrol's CyberNOT list. *See* Cyber Patrol *Main Page* (visited Sept. 9, 1999) <http://www.cyberpatrol.com/>.

[17] See *Paris Adult Theatre I v. Slaton*, 413 U.S. 49, 55 (1972), which held that an injunction could be used so long as adequate procedures to determine obscenity had been used. This would probably not be permitted absent a judicial finding. *See* Rowan v. United States Post Office Dep't, 397 U.S. 728, 738-39 (1970).

[18] Again, the sender does not know the recipient; the recipient does not know the content of the item; the intermediary does not know either. *See supra* Section I.A. This does not mean that there would not be extreme, and therefore easy, cases. The speaker would certainly know, therefore, whether some kinds of speech were highly likely to be permitted. Banalities about the weather constitute fairly safe speech acts anywhere; sadistic child porn is fairly unsafe in most jurisdictions.

The law ordinarily creates incentives through property or liability regimes. While a property regime in this area seems conceivable,[19] we focus here on a liability regime. The law can create an incentive to produce the information necessary to determine whether an exchange should be blocked by assigning liability to an actor for failing to block properly a transaction,[20] or by setting a default rule about whether to block properly a transaction when there is uncertainty.[21]

We consider two such defaults.[22] Under the first default, the sender incurs liability if she enters a transaction without reliable indicators that the transaction was in fact legal, and that transaction is later determined to be illegal. We call this the "prohibited unless permitted" rule. Because liability turns on the steps taken to comply with the law, it is distinct from a prior restraint.[23]

Under the second rule, the sender incurs liability only if she enters a transaction in the face of indicators that the transaction was in fact illegal, and that transaction is later determined to be illegal. We call this the "permitted unless prohibited" rule, and it is equivalent to a rule punishing a specific intent to violate the law.[24] One modification of this second rule would hold the sender responsible if the sender should have known that the transaction was illegal. This would com-

[19] For an excellent analysis of a property regime for dealing with access control, see *Developments in the Law — The Law of Cyberspace,* 112 HARV. L. REV. 1574, 1634-57 (1999).

[20] *See infra* Part II.

[21] By "uncertainty" we mean simply not having a given type of information — for example, information about the jurisdiction from which a receiver comes.

[22] We do not claim at this point that either default would, for all types of speech, be constitutional under the U.S. Constitution. Nor do we speak about the burdens of proof under a particular statute. We assume throughout that the state bears the burden for all elements of the charge. *Cf.* Smith v. California, 361 U.S. 147 (1959) (finding it unconstitutional to hold a bookseller criminally liable regardless of the bookseller's knowledge of the obscene contents of books sold). Rather than claim what is constitutionally possible, our defaults help clarify the relationship between the proscription and uncertainty. Like Schauer's article, our objective is to further explore this relationship, and the constitutional implications of uncertainty. *See* Frederick Schauer, *Fear, Risk and the First Amendment: Unraveling the "Chilling Effect,"* 58 B.U. L. REV. 685, 725-29 (1978).

[23] It is distinct because there exists no requirement to not send, but rather simply a punishment for sending without indication that the sending is legal. We concede this is a fine line, but with our defaults we aim, as we have explained above, not so much to limn the contours of American constitutionalism, but to understand the relationship between these rules and uncertainty.

[24] The Model Penal Code equates specific intent with "acting knowingly." The relevant section reads:

> A person acts knowingly with respect to a material element of an offense when:
> (i) if the element involves the nature of his conduct or the attendant circumstances, he is aware that his conduct is of that nature or that such circumstances exist; and
> (ii) if the element involves a result of his conduct, he is aware that it is practically certain that his conduct will cause such a result.

MODEL PENAL CODE § 2.02(2)(b) (1962).

port with a negligence standard, and we consider this alternative where relevant in the analysis below.[25]

These default rules will have significant consequences for behavior if there is systematic uncertainty about either the nature of Internet content or the character of Internet users (as of course there is). In cases of uncertainty, the "prohibited unless permitted" rule will be overbroad (it will block more speech than the state has a legitimate interest in blocking), while the "permitted unless prohibited" rule will be ineffective (since there will be insufficient incentive to discover the relevant information about what speech should be blocked).[26] Thus in the face of uncertainty, the default rule will be important, especially if one default is constitutionally compelled.

We will focus on changes in the architecture that might reduce the uncertainty. Stated abstractly, these changes will either tag speech, or tag people. If speech is tagged, then an intermediary or recipient can more easily determine item types and block accordingly; if people are tagged, then an intermediary or sender can more easily identify recipient and jurisdiction types and block accordingly.

D. *Steps Three and Four: Monitoring and Enforcement*

In the final two steps the regulator must first devise schemes for monitoring compliance and, second, implement schemes of enforcement. In both cases where the target of regulation sits, relative to the regulating regime, is an important factor in selecting among regulatory regimes. And in the case of monitoring, the technology used to effect the access control will significantly alter the costs of monitoring. Some technologies, that is, would be open for an automated and random verification; others would not.

The major issues for enforcement all involve the question of whether the regulating jurisdictions can easily, or cheaply, reach the target of enforcement. We assume there are more receivers than senders, so one might believe targeting senders would be cheaper than targeting receivers. This, however, becomes complicated when the sender operates from outside the regulating jurisdiction, making the sender sometimes legally, or at least practically, beyond the reach of the regulating jurisdiction. The cost of enforcement against these aliens may make it cheaper to enforce a rule against receivers than senders.

Whether more receivers or listeners exist, however, there are certainly fewer intermediaries than either. Intermediaries, as we discuss

[25] *See infra* Section II.A.

[26] The RESTATEMENT (SECOND) OF TORTS §282 (1965) defines negligence as "conduct which falls below the standard established by law for the protection of others against unreasonable risk of harm."

below, could be good targets of regulation, even though they possess even less information than either the sender or receiver. Again, the savings in enforcing a rule against them may outweigh the cost of their obtaining the necessary information. Thus from a social cost perspective, making them liable could be efficient.[27]

II. ALLOCATING RESPONSIBILITY

We now consider the consequences, under each of our two default rules, of allocating responsibility among our three actors — first to the sender, then to the recipient, and finally, to the intermediary. Within each allocation, we also consider how changes in existing law and Internet architecture might more efficiently achieve the aim of access control — more control at less free speech cost. This comprises our "sensitivity analysis" within each allocation. Finally, at the end of this Section, we consider a "mixed" strategy for the special case of "indecent speech" and children.

A. *Sender Responsible for Blocking Access*

Our first rule would make the sender responsible for controlling access. To comply with this rule, the sender must determine both the law of the jurisdiction of the recipient and, depending upon that law, certain characteristics of the recipient. Material considered "harmful to minors" present the obvious case, because many states require that providers of such material keep it from kids.[28] But the rule anticipates more general possibilities: rules regulating SEC filings, for example, make the content of that filing depend upon whether the reader is or is not a U.S. citizen.

Under the present Internet architecture, both determinations entail high costs. There is no simple way to identify the jurisdiction within which the recipient resides,[29] and no cheap way to be certain of

[27] Our conclusion, however, will be that the social costs of enabling blocking by intermediaries are in fact quite high. *See infra* text accompanying notes 58-60.

[28] *See, e.g.,* Reno v. ACLU, 521 U.S. 844, 887 & n.2 (1997) (O'Connor, J., concurring in the judgment in part and dissenting in part).

[29] A web server, for example, knows the IP address of the client computer that requests a web page, but usually knows little else about the recipient. An IP address does not readily identify a geographic location, because the administrative practices surrounding IP address allocation have not been based solely on geography. By analogy with the telephone numbering system, IP addresses have been allocated more like 800-numbers than like the numbers in regular area codes. Moreover, there is currently no single up-to-date database indicating the location of the computer using each IP address. (In practice, to facilitate routing, address allocations do roughly follow geography, which means that such a database might not be too unwieldy if it were assembled.) An IP address does not even uniquely identify a recipient computer, since dial-up connections through an Internet service provider typically are assigned a different address each time they dial.

characteristics of the individual. The rule would therefore be quite costly to a speaker — unconstitutionally costly, according to *Reno v. ACLU*, though differently costly under each of our two default rules.

Under the "prohibited unless permitted" rule, the cost falls on "free speech" interests. The burden of determining eligibility will likely create a significant chill on the speaker's speech.[30] The sender would have to take steps outside of the architecture of the Net to determine where a recipient is — by verifying an address, for example, or by using an area code on a telephone number as a proxy for the location. And the sender would need to rely upon proxies from credentials (such as a credit card) to guess whether the individual is of a proper age or not.

The United States Supreme Court has permitted this regime in the context of obscenity, where the sender must determine both the jurisdiction relevant for the recipient and the law of that jurisdiction.[31] It has not directly addressed the same question in the context of speech "harmful to minors" on the Internet, where the sender must determine, in addition to the jurisdictional information, the age of the recipient. In *Reno v. ACLU,* the Court did cite the burden of verification as one reason that the CDA's "indecency" provision was constitutionally suspect.[32] But *Reno* did not consider the "harmful to minors" standard — or, as described by some, the obscene-as-to-minors standard[33] — and the Supreme Court has not clearly indicated that the test would be different.

If, on the other hand, the rule is "permitted unless prohibited," the cost lies in the effectiveness of the regulation. Under this rule, the existing architecture would make any access control ineffective. While in real space, certain facts about an individual are unavoidably self-authenticating (a ten-year-old boy does not look much like a twenty-year-old man), in cyberspace, such facts are not. To determine either the jurisdiction or the age of the recipient requires affirmative steps by the sender. If no obligation to take such steps exists, or if no require-

[30] Though the use of the word has become quite general, we attempt in this essay to follow Schauer's definition of "chill," which refers "only to those examples of deterrence which result from the indirect governmental restriction of protected expression." Schauer, *supra* note 22, at 693.

[31] *See* Hamling v. United States, 418 U.S. 87, 104-06 (1974).

[32] *Reno v. ACLU,* 521 U.S. at 876.

[33] *See, e.g.,* Upper Midwest Booksellers Ass'n v. City of Minneapolis, 780 F.2d 1389, 1394 (8th Cir. 1986); *see also* M.S. News Co. v. Casado, 721 F.2d 1281, 1287-91, 1295 (10th Cir. 1983) (upholding a requirement that obscene-as-to-minors magazines be placed in "blinder racks"). Under *Ginsberg,* "minors may constitutionally be denied access to material that is obscene as to minors," but adults may not. *Reno v. ACLU,* 521 U.S. at 895 (O'Connor, J., concurring in the judgment in part and dissenting in part) (citing Ginsberg v. New York, 390 U.S. 629, 633 (1968)). Material is obscene as to minors if it is patently offensive, appeals to minors' prurient interest, and completely lacks socially redeeming value for minors. *See id.*

ment exists to block unless such steps are taken, then the rule will not effect the intended access control.

The existing architecture of the Internet therefore creates a great burden for the sender if the default is "prohibited unless permitted," and it defeats access control if the default is "permitted unless prohibited."

1. *Sensitivity*

Some of the burden on the sender could be reduced by architectural and legal changes. In this Section we describe four, and consider the potential costs and benefits of each.

The first two changes involve ways to identify more cheaply facts about the recipient. The two facts unknown by the sender are the jurisdiction of the recipient, and characteristics of the recipient (that she is, for example, over eighteen). The changes described here would facilitate the sender knowing both facts at a relatively cheap cost.

The first technique relies on digital certificates.[34] In the standard model of certificates, certificates identify who someone is. They are digital objects cryptographically signed by a certificate authority, a widely trusted entity that verifies an individual or organizational identity before issuing a certificate. The dominant use of such certificates today is to certify the identity of the holder.[35] This is the model, for example, of the VeriSign Digital ID, which VeriSign, one of the best-known certificate authorities on the Internet, describes as a "driver['s] license[] for the Internet."[36]

But there is no reason that the same technology could not be used to certify facts about the holder — or, more generally, to certify any assertion made by the signer. In our case, a signing certificate authority could then certify that *X* is from Massachusetts, and that *X* is over the age of eighteen, without identifying who *X* is.[37] Senders would then examine these certificates before granting access to regulable speech. Access would then be granted without a cumbersome system of passwords or IDs.

[34] *See generally* A. Michael Froomkin, *The Essential Role of Trusted Third Parties in Electronic Commerce*, 75 OR. L. REV. 49 (1996) (describing digital certification techniques).

[35] *See id.* at 58-62.

[36] *See VeriSign Digital ID Center* (visited Sept. 16, 1999) <http://www.verisign.com/client/index.html>.

[37] David Chaum was an early proponent of such characteristics certificates rather than identity certificates. *See* David Chaum, *Security Without Identification: Transaction Systems to Make Big Brother Obsolete*, 28 COMM. ACM 1030 (1985).

We can call this a "credentialling" solution.[38] It requires that the sender make certain judgments about the speech at stake; but it allows the sender to rely upon representations about the jurisdiction and the recipient that are necessary to determine whether an exchange is or is not blocked.

Under a "prohibited unless permitted" regime, access would be blocked except to those who could show that they carry the proper credentials. In the case of "harmful to minors" speech, the credential would be an adult ID indicating that the recipient is over eighteen. Recipients interested in receiving restricted materials will have an incentive to show such credentials. All else being equal, certificates would lower the cost of such a showing, and therefore reduce the burden, and hence chill, of the access control regime. Moreover, the burden on individuals under such a regime would be lower than under a regime where they must show a credit card or other form of identification. The cost of a certificate should be less than the cost of a card, and the possibilities for anonymity should be greater.

While no legal mandate on recipients would be needed to encourage showing age or jurisdiction certificates under a "prohibited unless permitted" regime, sanctions would be needed to reduce fraudulent use of certificates. If, for example, it were easy to obtain an anonymous adult-ID certificate, one might imagine a black market emerging, with children acquiring certificates from adult intermediaries. This potentiality creates a practical limitation on any regime where a credential grants access, since it creates an incentive to construct a false credential. One way to limit the transferability of anonymous certificates would be to include an IP address in the certificate, so that it could only be used with a single computer, or for the duration of a single dial-up connection if an access provider assigns different addresses for each dial-up session. (Each computer on the Internet has a numeric identifier, called an IP address.) Another technique for limiting transfers would be to make the certificates traceable, so that if abuse is detected, the identity of the original acquirer could be revealed, and that person could be punished.

Alternatively, widespread use of digital certificates could also improve the effectiveness of a "permitted unless prohibited" regime, by providing senders with enough information to block correctly exchanges that would otherwise have been permitted by default.[39]

To minimize the burden of this rule, the rule could require that the recipient provide the certificate only if the server asks, and the server

[38] Note that even though the technology for this solution is already in place, we refer to it as a possible architectural change because a widespread change in social practices would be necessary for the technology to be used in this way.

[39] One version of this regime would have recipients provide child-IDs, which we discuss below when considering COPA. *See infra* Section II.D.

would be required to ask only if the material is illegal in at least one jurisdiction. This regime would still somewhat burden those recipients living in jurisdictions where the speech was wholly legal; its viability would rest then upon the significance of that burden.[40] Alternatively, the rule could require that intermediaries provide or assure that users have valid certificates. In this case, the appropriate intermediaries would be the Internet Access Providers who serve recipients. If the state requires such intermediaries to assure the supply of certificates, then the costs of monitoring and compliance might be lower than if the same role were performed by the state. The intermediary has an advantage not in executing the primary conduct — certainly receivers stand in a better position to certify than intermediaries — but in assuring that the primary conduct is properly regulated.

A second architectural change to help the sender identify the recipient's jurisdiction would be an IP map — a table that would give a rough approximation of the location of the recipient's computer.[41] No doubt the map could not be perfect, and senders or recipients could use proxies to escape the consequences of the map. But in the main, the map might sufficiently segregate restrictive jurisdictions from nonrestrictive ones.

An IP map would provide benefits over a certificate system. Under the "prohibited unless permitted" regime, an IP map may burden speech even less than the certificate regime, since the cost to the recipient of this form of identification is zero, and the processing costs to the server would be lower than processing a certificate. The "permitted unless prohibited" regime becomes more effective as well, since now the sender has an assured way of knowing the jurisdiction into which the material is being sent, though not information about the recipient's age or other characteristics.

But there are important social costs associated with this IP-to-geography mapping that flow from its generality. Since jurisdiction identification would be determinable with any IP transaction, the regime would effect jurisdiction identification independent of the kind of speech being accessed. This raises obvious privacy concerns, which might be mitigated by structures that would limit the use of the map-

[40] Another possibility would be for the server to send a request of the form "if you are in jurisdiction *X* or *Y* and you are under 18, please provide a child ID," which would further reduce the burden of the system.

[41] Currently, the InterNIC maintains a database of the assignations of IP addresses to organizations. This database is public, and a copy of it may be queried from any computer on the Internet. Unfortunately, some entries in the database are incomplete or out of date, and they do not necessarily identify the location of computers using the IP addresses. It has been suggested, however, that such a database be used as a starting point for developing an IP to jurisdiction mapping. *See* Philip McCrea et al., *Blocking Content on the Internet: A Technical Perspective*, app. 5 (visited Sept. 16, 1999) <http://www.noie.gov.au/>.

ping for specific purposes. But for obvious reasons, it would be difficult to limit the use of this information.

The final two architectural changes would aid senders in classifying their speech according to the categories of various jurisdictions. The first is an automated preclearance technology. While we presume that the sender knows about its speech, the sender may not understand the classification scheme of every legal jurisdiction. Preclearance of the sender's materials can reduce or eliminate the uncertainty. Judicial preclearance would entail high costs. Determinations could be performed more cheaply, however, by third parties or even computer programs, but there would inevitably be some errors when compared with the gold standard of judicial determinations. Thus, in the United States at least, a judicial determination is required to block exchanges proactively.[42]

A voluntary preclearance regime, however, might be acceptable, even with nonjudicial determinations. Suppose that the government promised not to prosecute a sender for exchanges that had been precleared as acceptable. Uncertainty would remain about other items. It would seem initially that such a voluntary regime would be speech-enhancing. On the margin, if a speaker could be certain that her speech were permissible, she would be more likely to utter it than if she faced the risk that it would be illegal. But some who have considered the matter believe that if this voluntary regime became effectively mandatory, and if speech that did not appear on a preclearance list thus became effectively restricted, such a list would become constitutionally suspect.[43] The Constitution notwithstanding, we believe that the voluntary regime's overall effect is unclear: preclearance could lead to less chilling of speech (if it is clearer what is prohibited and what is not) but to more control of speech (if it results in greater prosecution).

A second way to reduce uncertainty about how to classify items according to particular jurisdictions' categories would be a thesaurus that relates the categories of different jurisdictions. Thus, if the sender is able to classify an item according to one jurisdiction's categories, it could infer the classification in some other jurisdictions. For example, it may be that anything classified as child pornography in jurisdiction *A* would be classified as obscene in jurisdiction *B,* though the converse inference might not hold. The thesaurus functions as a more complex

[42] *See, e.g.,* Paris Adult Theater I v. Slaton, 413 U.S. 49, 55 (1972).

[43] *See* Schauer, *supra* note 22, at 725-29. The closest case is perhaps *Bantam Books v. Sullivan,* 372 U.S. 58 (1963), where the Court invalidated a "blacklist" Commission. The preclearance idea is not quite a blacklist — the result of the submission would be a promise not to prosecute, not a determination that the material was "obscene." Again, however, we concede that the line is a difficult one to sustain.

version of the base jurisdiction model that we described in equation (c).[44]

B. *Recipient Responsible for Not Taking Access*

Our second rule would make the recipient responsible for illegal transactions — targeting the buyer, that is, rather than the seller. Under this rule, then, it is the recipient who incurs liability if an improper transaction occurs.

This rule has some advantages over the sender-responsible rule — the recipient, for example, may be in a better position to know about the law of its jurisdiction, and about its own recipient type. But obvious disadvantages exist as well. The recipient stands in a worse position, relative to the sender, to know about the kind of information that the sender is making available. While a sender may find it burdensome to classify its speech according to any given jurisdiction's categories, at least the sender begins with knowledge about the content of the speech at issue.[45] The receiver does not. This lack of knowledge means that a recipient cannot determine the legality of an exchange until after the exchange has occurred. Thus, under a "prohibited unless permitted" rule, the receiver risks liability[46] in the very act of determining whether a particular exchange complies with the law. And under the "permitted unless prohibited" rule, restrictions would likely be completely ineffective because the recipient would have a significantly reduced incentive to accurately assess the legality of the exchange.

A second problem with placing liability on the receiver results from the costs of classification. Because receivers outnumber senders, this rule shifts the cost of classification to the many, rather than to the few. This cost shift will result in either too much or too little blocking. For those who have a strong interest in blocking certain speech, the costs of classification will push the classifier to an overly conservative strategy. For those who have little interest in blocking certain speech, the costs would likely push the classifier not to classify at all.

Finally, putting the responsibility on the receiver may increase the costs of enforcement. Receivers are ordinarily individuals, and therefore more difficult to target. Whether this would increase the cost of

[44] *See supra* Section I.B.

[45] It would be different, of course, if the sender were considered as a bookstore, without knowledge, or any simple way to get knowledge, about the content of its books. *See* Cubby, Inc. v. Compuserve, Inc., 776 F. Supp. 135, 139-40 (S.D.N.Y. 1991). We would consider such a "sender" to be an intermediary in our analysis.

[46] This depends upon the level of knowledge required for someone to be guilty under such a provision. If the statute were criminal, the knowledge requirement would be quite strong, so inadvertent liability would not be possible. But for a lesser prohibition, the knowledge requirement may be less.

enforcement generally, of course, depends upon whether the alternative targets — senders or intermediaries — are more easily regulated. If they operate primarily from outside the regulating jurisdiction, then regulating recipients may be less costly than regulating senders or intermediaries.

1. *Sensitivity*

A recipient-responsible rule could be made less costly if there were cheaper ways to identify the speech before the transaction. Labels or content rating is an obvious solution here. Two sorts of labeling are possible. One we have already described — prescreening[47] — and here the same techniques for reducing the costs of preclearance would apply, including the use of automatic text classification and delegation of the preclearance powers to an independent third-party rater. As we mentioned before, however, there remains a concern about the constitutionality of even a voluntary preclearance regime.[48] In the American context, despite the reduction in uncertainty, this might constitute a prohibited regulatory change.

The other labeling solution is to rely on senders to label their own materials. The labels might directly indicate whether the item is permitted or prohibited to recipients of various ages in particular jurisdictions, or it could describe the item in detail (on dimensions such as sex-related) sufficient to infer whether it should be blocked.[49] This solution simply inverts the certificate solution — here the sender offers a "certificate" upon which the recipient relies, while in the case above, it was the recipient providing the certificate upon which the sender would rely. The analysis is also analogous.

Under a "prohibited unless permitted" regime, the labels would convey information that the speech is permitted (for example, no sex or hate speech). Recipients would receive immunity if they in good faith relied upon a sender's labels to determine that access is permitted. Senders would have a natural incentive to provide labels, since they would allow more recipients to receive the speech, although penalties for inaccurate labels might be needed to prevent widespread mislabeling. There would of course be a transition period, during which only a small percentage of materials would carry self-rating labels, rendering most of the Net blocked under a strict "prohibited unless permitted" rule. To minimize the transition period, authorities

[47] *See supra* Section II.A.1.

[48] *See id.*

[49] The labels could be expressed in PICS format, *see generally Platform for Internet Content Selection* (last modified Aug. 4, 1999) <http://www.w3.org/PICS/>, or the new RDF format, *see generally Resource Description Framework* (last modified Aug. 9, 1999) <http://www.w3.org/RDF>, and distributed along with the items.

might publicize well in advance the imposition of such a regime, in the hopes that most senders would label before the filtering took effect. It is not clear how effective such advance publicity would be, and there would probably be a great public outcry during the transition period, perhaps enough to cause a reversal of the regulation.

Under a "permitted unless prohibited" regime, the labels would indicate that access to an item was prohibited (to some groups in some jurisdictions). The obvious problem here is that the sender would have little incentive to label, since that could only reduce legal access.[50] To bolster the effectiveness of this regime, a government might require senders to provide labels. This may raise a constitutional question in the United States if labels were considered compelled speech.[51] Some have argued that they would not,[52] but we believe this is a close question. To reduce the cost to senders of labeling, a government might subsidize third-party ratings or itself produce suggested ratings. In the United States, its ratings could not be treated as definitive[53] in such a system, but they may provide an aid to senders in self-labeling.

The burden of labels might be minimized by simply requiring labels only where speech is potentially regulable (comparable to requiring that people up to the age of twenty-six carry IDs to purchase cigarettes, even though the prohibition reaches only those eighteen and under). Even here, however, the requirement raises difficult questions, since it would require speech by the sender in the form of a label even when the underlying speech is clearly legal in the receiving jurisdiction. Thus the most restrictive jurisdiction would in effect determine whether the speaker must label.

[50] If many people voluntarily adopted a "prohibited unless permitted" filter, then the market demand for labels might sufficiently encourage sender self-labeling, even if the state mandated only the less strict "permitted unless prohibited" regime. For example, consumers might turn on the facilities in Microsoft's Internet Explorer (version 3 and higher) or Netscape Navigator (version 4.5) to voluntarily block access based on senders' PICS-formatted self-labels.

[51] The "compelled speech" doctrine forbids the government from forcing individuals to assert the views of the government. See LAURENCE H. TRIBE, AMERICAN CONSTITUTIONAL LAW 804-06 (2d ed. 1988) and the cases cited therein.

[52] *See, e.g.*, R. Polk Wagner, *Filters and the First Amendment*, 83 MINN. L. REV. 755, 777-98 (1999).

[53] The government's own ratings are not always determinative of whether speech was delivered or not, absent a judicial finding. *See* Rowan v. United States Post Office Dep't 397 U.S. 728, 738-39 (1970). To force individuals to label their content would, we believe, often require them to make judgments about the character of the material they were labeling. But the labels in this context are not objective, nor independent of a viewpoint about the nature of the material. To have to assert the character of the material, then, can be to require individuals to make what is in essence a political statement. If a newspaper cannot be forced to publish a story it does not otherwise want to print, see *Miami Herald Publishing Co. v. Tornillo*, 418 U.S. 241 (1974), then neither should a web site be forced to publish a story about itself (i.e., that it carries materials of type *X*) that it does not otherwise want to print.

As with recipient certificates, the responsibility for assuring a supply of sender labels might be assigned to intermediaries, in this case to the sender's Internet Access Provider. One important asymmetry exists, however. While age and jurisdiction are objective properties that one might reasonably expect an access provider to verify, correct assignment of rating labels to items will involve subjective judgements. One intermediate form of responsibility might be to require an access provider to assure the availability of some sender self-label, but to make only the sender and not the access provider responsible for any inaccuracies in the label.

There is a practical enforcement problem with mandating that senders provide labels. Just as it may be difficult to enforce blocking requirements across jurisdictional boundaries, it may be difficult for authorities in one jurisdiction to enforce a labeling requirement in another.

C. *Intermediary Responsible for Blocking*

We have assumed that the intermediary possesses information about neither the recipient nor the item the sender would send. It might therefore seem odd to consider the intermediary as a possibly responsible actor.

But intermediaries provide a cheap target of regulation. Fewer of them exist than either receivers or senders, and they are typically more stable, or harder to move. Just as the government can more easily regulate telephone companies than it can telephone users, it would be easier for the government to set requirements on intermediaries, which intermediaries could then enforce upon their customers. More importantly, because intermediaries have an interest in reducing the cost of compliance, regulating intermediaries will more likely catalyze innovation in compliance methods.

In addition to a lack of information, intermediaries may have limited capabilities for implementing blocks. Blocking can either be implemented at the application layer (for example, web page requests) or at the network layer (for example, individual packets). Whereas application layer blocking allows tailored blocking of URLs,[54] network layer blocks are of necessity much cruder: only the sender's and receiver's IP addresses and the port number (a rough indicator of whether the connection is being used for a web transfer, email, or something else) are available. Thus, a network layer block can either block all web requests to a particular IP address, or none of them.[55]

[54] A URL identifies both a computer to connect to and a path or file name to request. Thus, URL blocking enables some files from a Web server to be blocked, while others are not blocked.

[55] For a more complete description of application layer and network layer blocking, see McCrea et al., *supra* note 41, at 25-31.

We consider two types of intermediaries. One type provides Internet access, such as Internet access or service providers, or even employers and schools (for simplicity, we will refer generically to any of these as an IAP). It is reasonable to assume that an end-user and his or her IAP lie in the same jurisdiction.[56] Many, but not all, IAPs run proxy servers (and other application layer gateways) which intercept some kinds of Internet traffic. Most commonly, a web proxy at an IAP will keep copies in a cache of frequently accessed web pages; when a customer requests a cached page, the proxy sends it to the customer, without fetching it again from the sender's web server. Proxy servers permit application layer blocking: requests for certain URLs can be blocked. Moreover, an IAP may configure a firewall that forces all requests to use the proxy server. This is done most frequently to enhance corporate security, by restricting the Internet traffic entering and leaving a corporation to only that which passes through proxies. In those cases where an IAP does not employ proxy servers, however, only cruder network layer blocking is possible.

The second type of intermediary is a backbone provider, which carries data across jurisdictional boundaries. In practice, the IAP may also run backbone services, but the services are conceptually distinct because they have different technical filtering capabilities. Consider the cross-jurisdiction transit point, the place in the backbone provider's network where data crosses a jurisdictional boundary. Such transit points do not normally employ proxy servers or other application layer gateways. Thus, only the cruder network layer blocking is possible at cross-jurisdiction transit points, given the current Internet architecture.

One final difficulty with blocking by intermediaries is that recipients may find ways to bypass the blocks, especially if the senders cooperate. For example, the same prohibited document may be available from several different URLs, so that a recipient can access one even if the others are blocked. A technique known as tunneling, where the contents of one packet are wrapped inside another packet, may bypass a network layer block.[57]

1. *Sensitivity*

These architectural features yield the conclusion that intermediaries cannot effectively control access, and given the fundamentality of these features, it might seem unadvisable to make changes that would increase their ability to control. Because intermediaries are also prac-

[56] It would be possible, though expensive, to make an international phone call to access an IAP in another jurisdiction.

[57] McCrea et al. detail these and other ways that senders and recipients might bypass intermediaries' blocks. *See* McCrea et al., *supra* note 41, at 35, 37.

tically easier for a jurisdiction to regulate, however, we will consider what changes might make this control possible.

A combination of the architectural changes discussed in previous Sections could provide intermediaries with enough information to decide which exchanges to block. That is, information about item types could come either from senders' labels or from preclearance lists provided by jurisdictions. Information about recipient type could come from certificates, and information about recipient jurisdiction could come either from certificates or from a database lookup on the IP address.

One potential change in the architecture to facilitate the implementation of blocking would be to require an application layer gateway at IAPs or cross-jurisdiction transit points, and to require that all customer traffic use these gateways (perhaps enforced via a firewall). This would have high costs for Internet flexibility and operation. First, it would be computationally expensive to assemble all packets into messages at cross-jurisdictional transit points, especially for traffic where there is no counteracting performance gain from caching. Second, messages may be encrypted for privacy or security purposes (for example, in SSL connections) so that even at the application layer only crude blocks based on sender and receiver address are possible. Third, innovations that introduce new applications would be stifled, since the application layer gateways would not initially know about the new applications and hence would block them.[58] The Internet's current architecture has enabled experimentation and rapid deployment of new applications (examples of applications that blossomed in part as a result of this flexibility include the world wide web, push services, and ICQ).[59] One final cost might come in the form of reliability. It is relatively easy for a service provider to provide multiple routers, so that a temporarily disabled router would not interrupt the network layer service. It may be more costly to arrange for continued service, however, when an application layer gateway is temporarily disabled.[60]

[58] Many corporate firewalls prevent employees from using experimental applications that the corporate proxy or gateway is not configured to handle. *See* WILLIAM R. CHESWICK & STEVEN M. BELLOVIN, FIREWALLS AND INTERNET SECURITY: REPELLING THE WILY HACKER 76 (1994).

[59] ICQ ("I Seek You") maintains a worldwide registry of users and their status (online, busy, away, etc.), allowing users an easy way to keep track of friends and acquaintances. The ICQ client software interacts with the registry updating a user's information and receiving information about others on that user's "contact list." The ICQ client also acts as a platform for chat and other message exchange between any two registered ICQ users. *See How to Use ICQ* (visited Sept. 9, 1999) <http://www.icq.com/icqtour/>; *What Is ICQ?* (visited Sept. 9, 1999) <http://www.icq.com/products/whatisicq.html>.

[60] *See* McCrea et al., *supra* note 41, at 31.

D. *Incentives for Tagging: "Kids" and COPA*

Several of the regimes that we have considered depend on one party providing information that another party uses to make its decision to block. In these situations, the incentives of the information provider depend on the default rule recognized by the regulator. A blocking rule consistent with the "prohibited unless permitted" requirement creates an incentive (depending upon the regime) for recipients to provide certificates or for providers to label content. To make the more permissive "permitted unless prohibited" regime work, the regulator must threaten a sufficient punishment to induce the needed tagging of information or recipients.

Regulations designed to protect kids, however, present a special case.[61] If one assumes that the parent is the relevant "recipient," then unlike the general case, the recipient has an incentive to facilitate blocking — if, indeed, blocking access is what that parent wants.[62] Sender-based regulations could therefore follow a "permitted unless prohibited" regime without restricting recipients who do not identify themselves.

This difference has constitutional significance when one considers regulations designed to block access to kids. For example, consider one alternative to the regulation prescribed by Congress in CDA and COPA — a regulation that would require senders to block only self-identifying kids, rather than regulation, such as CDA and COPA, that required senders to block all receivers except those identified as adults. The following hypothetical statute will suggest the idea.

1. *Tamper-Resistant Kids-Mode Browsers ("KMB").* A "kids-mode" browser is a browser that signals to servers that the user is a minor.[63]
2. *Server Responsibility.* When a server detects a kids-mode client, it shall (1) block that client from any material properly

[61] They are a special case as well in that, relative to other mandated access control, the content here is easier to identify. The model becomes far more complex if content such as "defamatory" or "seditious" speech were considered. Likewise, the problem becomes far more difficult if the recipient always has an incentive to evade the regulation. Regulating "kids" is a special case because, at least sometimes, the parent has an interest in enforcing the regulation.

[62] The Court in *Reno v. ACLU* made it clear that the relevant question is whether parents are enabled in protecting kids, not whether the state is. 521 U.S. 844, 865 & n.31 (1997). If a parent decides to give kids access, that decision cannot, for the range of speech being discussed here, be overridden by the state.

[63] Manufacturers of browsers and operating systems would presumably make it difficult to modify the kids-mode configuration without a password. In practice, this would require both that the kids-mode browser configuration not be easily changed, and that the operating system prevent installation of a fresh browser where the child could choose whether to set it in kids-mode. Eventually, we might expect that the kids-mode setting would migrate entirely into the operating system, with all browsers' behavior determined by the operating system setting.

> deemed "harmful to minors," and (2) refrain from collecting any identification data about the user, except data necessary to process user requests (such as IP addresses). Any data collected shall be purged from the system within *X* days.

A browser "signals" to a server that its user is a minor in just the way the browser now signals its browser type (for example, Microsoft Internet Explorer or Netscape Communicator). Under the present architecture of the Net, the server "knows" what kind of browser you use, the IP address you are browsing from, whether the client will accept "cookies,"[64] and the site you were viewing before you switched to that site. The statute would simply require that if a browser signaled that a user was a kid, then the server would not transmit material "harmful to minors."

The statute imposes burdens, but burdens that are far less significant — practically and constitutionally — than the burdens of CDA or COPA. The primary practical burden rests upon senders, who must now discriminate on the basis of whether the client is a kid. Relative to the existing constitutional baseline,[65] it would be a trivial change for servers to check for the existence of a kids-mode signal. The statute does not directly burden software manufacturers, since it simply defines what a "kids-mode browser" is. By requiring servers to respect a kids-tag, however, the statute creates an incentive for manufacturers to provide such browsers to parents who would want this option.

Finally, one might believe the statute burdens parents practically because they would have to activate the kids-mode browsing. But the burden here is not legally significant; it merely consists in the difficulty of checking a preference box and keeping a password secret. That, we believe, is far less significant than the alternatives, say, of purchasing and installing blocking software.

Compared with the burdens of COPA, we believe this regulation would be constitutionally preferred. To see this, one must consider the relative burden and effectiveness of the three statutes considered here — COPA, CDA, and our proposed statute.[66]

COPA was modeled on the CDA, but regulates more narrowly than the CDA. Like the CDA, it puts the burden on the speaker to

[64] "Cookies are a general mechanism which server side [sic] connections (such as CGI scripts) can use to both store and retrieve information on the client side of the connection. The addition of a simple, persistent, client-side state significantly extends the capabilities of Web-based client/server applications." *Persistent Client State HTTP Cookie, Introduction* (visited Oct. 6, 1999) <http://home.netscape.com/newsref/std/cookie_spec.html>. *See generally Cookie Central* (visited Oct. 6, 1999) <http://www.cookiecentral.com>.

[65] *Ginsberg* implies that suppliers can be burdened to separate "harmful to minor" speech from other speech. Ginsberg v. New York, 390 U.S. 629, 643-45 (1968).

[66] This analysis follows the test created by the Supreme Court regarding the constitutionality of speech-restrictive regulations involving the protection of children. *See infra* note 80 and accompanying text.

avoid speaking improperly to kids. But unlike the CDA, COPA puts that burden on a narrow class of speakers, in a narrower zone of the Internet. Under COPA, a commercial provider who "knowingly and with knowledge of the character of the material . . . [uses the] World Wide Web [to make] available to any minor . . . material that is harmful to minors" has committed a crime.[67] The statute is thus narrower in three ways: (1) in the breadth of speech regulated ("harmful to minors" rather than "indecent"), (2) in the scope of speakers covered (it does not reach noncommercial providers), and (3) in the range of the Internet affected (it does not reach newsgroups or chat rooms).

Similarly, the defenses provided to a speaker by the statute impose lesser burdens than those imposed by CDA. COPA's defenses are broader than the defenses under CDA. Under section 231(c)(1) of COPA, a provider has a defense if he

> in good faith . . . restricted access by minors . . .
>
> (A) by requiring use of a credit card, debit account, adult access code, or adult personal identification number;
>
> (B) by accepting a digital certificate that verifies age; or
>
> (C) by any other reasonable measures that are feasible under available technology.[68]

Section 231(c)(2) adds immunity from prosecution to this substantive defense of section 231(c)(1). It provides that no action can be brought against a provider who has in good faith attempted to implement one of the defenses from section 231(c)(1).[69]

These defenses are thicker than those in CDA. First, the statute envisions a form of identification not expressly recognized in the CDA — the digital certificate, which as we have described could more cheaply and with greater anonymity certify that someone is an adult.[70] Second, the catchall category of technologies ("by any other reasonable measures that are feasible under available technology"[71]) is broader than the parallel in the CDA. The CDA required that these other technologies be "reasonable, effective, and appropriate."[72] The Supreme Court read this standard not as an ordinary tort standard, but as an absolute effectiveness requirement.[73] COPA's test, by contrast, creates a traditional tort standard: a provider will have a defense if he takes those steps reasonable in the circumstances, given the existing state of technology, whether or not those steps are "effective."

[67] Child Online Protection Act, 47 U.S.C.A. § 231(a)(1) (Supp. 1999).

[68] 47 U.S.C.A. § 231(c)(1).

[69] *See* 47 U.S.C.A § 231(c)(2).

[70] *See supra* Section I.A.

[71] 47 U.S.C.A § 231(c)(1).

[72] 47 U.S.C.A § 223(e)(5)(A).

[73] *See* Reno v. ACLU, 521 U.S. 844, 881-82 (1997).

These differences evince what Congress purported to accomplish — a response to the concerns of the Supreme Court in *Reno v. ACLU*.[74] It followed the outline sketched by Justice O'Connor's concurrence, an outline of what she thought a constitutional regulation would be.[75] If the Court is eager to reward legislative obedience, it might well feel itself compelled to uphold Congress's latest effort.

But so far, lower courts have not been eager to reward Congress. While acknowledging that COPA is less restrictive than CDA, they have still concluded that COPA is too burdensome.[76]

In our view, this analysis is incomplete. We agree with Professor Volokh that the question posed by the Supreme Court is not whether the regulation is in some absolute sense "too burdensome."[77] That form of analysis so far has been restricted to abortion regulation[78] and perhaps also to the dormant commerce clause.[79] Rather, the question the Supreme Court has asked in this context is whether the regulation is more burdensome than needed.[80] If that question could be answered by asking whether COPA mandated the least burdensome *adult-ID* regime possible, then we believe this statute does impose the smallest adult-ID regime burden possible.

We believe this because, unlike CDA, COPA includes a catchall provision that permits "any other reasonable measures that are feasible under available technology."[81] This clearly invokes traditional negligence standards. Whereas CDA required that the technology *be* effective, COPA requires only that it be reasonably effective, given the existing technology. In effect, by definition then COPA creates the least burdensome *adult-ID* regime.

Our proposed regulation, however, creates an even less burdensome regime. The adult-ID regime is not the only ID regime possible. As we outlined at the beginning of this Section, an alternative would be a *kids-mode-browser* regime. By requiring that servers segregate based on whether a KMB was signaled, Congress would thereby enact a sender-based regulation. This sender-based regulation would be far

[74] *See* 144 CONG. REC. 139, H9902-11 (1998).

[75] *See Reno v. ACLU*, 521 U.S. at 887-88 (O'Connor, J., concurring in the judgment in part and dissenting in part).

[76] *See, e.g.*, ACLU v. Reno, 31 F. Supp. 2d 473 (E.D. Pa. 1999) (preliminarily enjoining COPA).

[77] The "burdensome" test is structurally similar to the test in abortion cases. *See* Planned Parenthood v. Casey, 505 U.S. 833 (1992).

[78] *See Casey*, 505 U.S. at 833.

[79] *See* Pike v. Bruce Church, Inc., 397 U.S. 137 (1970).

[80] *See* Sable Communications, Inc. v. FCC, 492 U.S. 115 (1989) (plurality permitting regulation of "dial-a-porn" to protect kids); Ginsberg v. New York, 390 U.S. 629 (1968) (upholding New York statute that required keeping of material harmful to minors from minors).

[81] Child Online Protection Act, 47 U.S.C.A. § 231(c)(1)(C) (Supp. 1999).

less burdensome than CDA or COPA, however — or indeed any adult-ID regime. And, we believe, it would satisfy the second part of the Court's test — it would be just as effective.[82]

The advantages of the KMB regime are many.

- First, the burden of signaling that the user is a kid would be far less costly than the burden of signaling that the user is an adult because there would exist no need to verify the signal. An adult ID needs to be verified because, by granting access that otherwise would not be permitted, there would be an incentive to cheat. A KMB, on the other hand, would only block access; there would be no incentive to lie.
- Second, the absolute number of people burdened by the regulation would likely be lower. Rather than requiring an expensive ID for every adult wishing full access to the web, only parents who want their kids to be blocked from access on the web would have to enable the kids mode. The burden here would thus fall on a much smaller proportion of the population, and, as described above, that burden would be less than the burden of adult IDs.
- Third, the burden on the parents of obtaining the software to enable this blocking is less than the burden of purchasing blocking or filtering software. Browsers are (for the moment) free; the district court found that the cost of blocking software was approximately $40.00.[83]
- Fourth, the burden of this regulation falls on browser manufacturers and web sites with "harmful to minors" material. The burden on the browser manufacturers is relatively slight; the burden on the web sites is the same burden that real space sites bear when they distribute such material. We don't mean to minimize these costs, but they are less extensive than COPA, and they do not impose any burden on recipients without kids.[84]
- Fifth, rather than an elaborate identification system, maintained either by companies such as AdultCheck or by content providers, this regime would provide only the single unverified assertion that the user is a kid. No other personal data would need to be provided; no compromise of financial information would be risked.
- Sixth, and relatedly, the cost of providing this identification data is far cheaper with the KMB than with the adult-ID system. The

[82] *See supra* note 81 and accompanying text.

[83] *See* ACLU v. Reno, 31 F. Supp. 2d 473, 492 (E.D. Pa. 1999).

[84] *Reno v. ACLU* indicates quite clearly, we believe, that the state's interest is limited to facilitating the choice by parents. *See supra* note 62. The government in *Reno* had argued that the state had an interest, beyond the interest of parents, to protect kids from speech "harmful to minors" even if the parents did not so wish. *See* Transcript of Oral Argument, *Reno v. ACLU,* 521 U.S. 844 (1997), *available in* 1997 WL 136253, at *19-24 (Mar. 19, 1997). But the Court did not embrace this broader restriction. *See Reno v. ACLU,* 521 U.S. at 865 & n.31.

code required to enable the two or three dominant browsers to identify users as kids is relatively trivial; the code to protect anonymity with the adult-ID is quite severe because it requires careful implementation of cryptography and security features.

- Seventh, this technique would easily apply to other kid-protective regulations. In the very same Act enacting COPA,[85] Congress enacted the Children's Online Privacy Protection Act.[86] That Act regulated the data that can be collected from a child online. The weakness in that statute is that there is no easy way to identify a child. But the change in browsers suggested here would be a way to identify a child.
- Finally, this technique would provide an easy way for schools to regulate access to the Net. A common profile for all users in a school could be set by a network administrator. That common profile would then control the types of sites to which the user could gain access.

These reasons together suggest why this alternative to COPA would be less burdensome. That is one half of the Court's test. The other half requires that the alternative be "at least as effective."[87] Here again, we believe that it would.

- First, if the ultimate test is whether the statute enables parents to control their children, we believe this alternative would be as effective as COPA. Of course, parents would be required to set the profile for use by kids, but this profile need not be difficult to set. Indeed, it would be easier to establish this profile than to establish a profile to collect email through a browser, which users already do routinely.
- Second, the kids profile would be easier to implement in places where kids are most likely to use the net. Schools could establish a common profile for all users within the school and disable the ability to build alternative profiles. These locations would then be protected locally, while under COPA, they are protected only if the kid does not get access to an adult ID.
- Third, while it is always possible for a child to take steps to evade the profile, there is no reason to believe it would be any easier to evade a profile than to evade an adult check requirement. The simplest way for a child to evade COPA is to steal a credit card number, and this is certainly as easy as cracking a security provi-

[85] *See* Omnibus Consolidated and Emergency Supplemental Appropriations Act, 1999, Pub. L. No. 105-277, 112 Stat. 2681.

[86] Children's Online Privacy Protection Act of 1998, Pub. L. No. 105-277, 112 Stat. 2681-728 (to be codified at 15 U.S.C. §§ 6501-6506).

[87] Reno v. ACLU, 521 U.S. 844, 874 (1997).

> sion built into a browser.[88] Indeed, it may be easier. Whereas a KMB would be protected by a password, the blocks of COPA are protected by credit card numbers; and while there are plenty of places that a credit card number can be found (credit card statements, receipts, etc.), only the careless would leave a password lying around.

These considerations suggest that this alternative is at least as effective as COPA. This is not to say, however, that either would truly be effective. Given the flood of sites from jurisdictions beyond the United States, any effort to regulate United States web sites would seem plainly ineffective. But between the two, we believe the kids-mode browser dominates COPA and would therefore render COPA unconstitutional.

We consider one final twist to the KMB. Compare the KMB regime that we have described to a quite different, though functionally similar, regime. Suppose the responsibilities for tagging and filtering were reversed. The sender would be responsible for tagging contents according to whether they are prohibited to children, and the kids-mode browser would block such items. Although this system would block exactly the same items as the KMB regime would, a difference exists. The difference lies in the information thereby revealed that could be used for purposes other than the legally required blocking. In the KMB regime that we have described, all senders (even of child-acceptable materials) would know which receivers were children. This could be useful, as mentioned above, in enabling other child protection regulation; it would also create a risk of other abuse.[89] In the alternative we consider here, all receivers would know which materials were labeled as inappropriate for children, even those receivers who did not signal themselves as children. Thus, these two regimes that have the same immediate consequences might have quite different secondary effects. One produces information for the servers; the other produces information for the surfer.

III. SECONDARY CONSEQUENCES

As we said at the start, our aim in this Article extends beyond COPA to any attempt to regulate access that does not fully consider the costs. In this Section, we extend our analysis of these costs to a

[88] One possible way to evade the limitation would be for a kid to download another browser and set it up to be free of the kids-ID restriction. But this possibility could be addressed. Again, the browser manufacturers could easily segregate download locations, based on whether the browser making the request were kids-ID enabled. If it were kids-ID enabled, then the company would download a kids-ID set browser only. Alternatively, the kids-ID could be enabled in the operating system ("OS"), making substitution of an adult OS for a kids OS significantly more difficult.

[89] *See infra* Part III.

category of costs that we believe have not fully been reckoned in the debate so far. These costs, in the end, may well be more significant than the costs of the "problem" that access controls seek to remedy.

In our analysis so far, we have considered three techniques for regulating access — tagging the sender, tagging the recipient, and regulating the intermediary to help effect either of the two taggings. All three strategies, we suggest, have effects that reach beyond their primary objective. All three envision a general infrastructure that can be used for purposes beyond those initially intended. This potential, we believe, should also be counted when reckoning the cost of a given regulatory strategy.

A. *IDs and Regulability*

To effect sender or intermediary control, we envisioned the development of identity certificates designed to facilitate the credentialling of certain facts about a recipient — how old that person is, where she is from, etc. We also proposed the development of a database that maps IP addresses to jurisdictions.

But it should be clear that if these architectures were enabled for this speech-regulating purpose, they would both have uses that extend well beyond this purpose alone. That is, these architectures might facilitate other jurisdiction-based regulation or access control imposed on senders, beyond the narrow purposes that motivated the initial change. We might make the Net safe for kids, but in consequence make it a fundamentally *regulable* space.

How? Certificates or IP databases would facilitate a more general structure of jurisdiction-based control, including taxation and privacy regulations. The reason is straightforward. Local jurisdictions have the legal authority to regulate their own, both while the citizens are at home and while they are away.[90] A certificate-rich Internet could facilitate the identification of who could be regulated by whom, or what standards could be imposed upon whom. And this, in turn, could facilitate a more general regulation of behavior in cyberspace.

We might imagine a scheme that looks something like this. States would enter a compact whereby they, as home jurisdictions, agree to require senders or intermediaries, within their own jurisdictions, to respect the rules of other jurisdictions, in exchange for senders or intermediaries in other jurisdictions reciprocating. These rules would specify the restrictions imposed on citizens from a given jurisdiction,

[90] *See* RESTATEMENT (THIRD) OF FOREIGN RELATIONS § 402(2) & cmt. e.(1987) For example, the Child Sexual Abuse Prevention Act, 18 U.S.C. § 2423(b) (Supp. I 1995), criminalizes traveling abroad to engage in illegal sexual acts with children. For an interesting discussion of this provision, see Margaret A. Healy, *Prosecuting Child Sex Tourists at Home: Do Laws in Sweden, Australia, and the United States Safeguard the Rights of Children as Mandated by International Law?*, 18 FORDHAM INT'L L.J. 1852, 1902-12 (1995).

and the range of citizens for whom the restriction applies. For example, a jurisdiction might specify that its citizens may not engage in Internet gambling; the jurisdiction within which a gambling server sits, then, would require the server to check for a person's citizenship, and condition access based on whether they held the proper credential. And presumably the jurisdiction would do this only if there were restrictions that it wanted imposed in other places and that it needed other jurisdictions to respect.

Thus, if a jurisdiction database or a credential-rich Internet were in place, we might expect voluntary uses of that infrastructure to proliferate, although with differing degrees of desirability. Some voluntarily imposed restrictions might seem reasonable. For example, recording companies might refuse access to their web sites from countries where pirated copies of intellectual property were rampant. Other voluntary uses might not have such sanguine effects. For example, some Serbs and Croats might refuse to allow each other access to their web pages. In both cases, the certificate infrastructure would enable a form of discrimination.

Not every ID architecture, however, would effect this increase in private regulability. Obviously, the more data a certificate architecture transmits, the more regulability increases. Likewise, an ID architecture with a narrow focus on children, like the KMB, would facilitate very little regulation beyond regulation protecting kids. This, by the way, constitutes a second, and we believe, compelling reason to prefer it over the adult-ID regulation of COPA; for if the government has two means available for protecting kids, we believe it should select the means that produce the least significant secondary effects, unless there is some analysis showing that the secondary effects also advance some legitimate government claim.

B. *Labels and Improper Control*

The other general solution we have identified for effecting mandated access control relies upon labels, designed to facilitate filtering by recipients or intermediaries. The labels might be provided by senders or by governments in the form of preclearance lists. But as should be obvious, an inexpensive and widely used labeling infrastructure would have its own secondary impacts, including the possibilities of more widespread speech regulation and of voluntary individual or collective uses of labels for blocking beyond the state's legitimate interest.

First, if available speech labels describe categories beyond those that a jurisdiction would normally regulate, the mere availability may tempt regulation within these new categories. Thus, the widespread use of a general labeling infrastructure may start governments on a slippery slope toward regulating all sorts of speech, even if the initial impetus for labeling is limited to only a few kinds of speech.[91]

Second, labels might be used for voluntary access controls as well as mandated access controls.[92] That is, recipients or intermediaries might choose to block more exchanges than governments require. Parents in the United States, for example, may choose to block young children's access to hate speech or speech about sex education, even though such speech is legal for children in the United States. Alternately, a search engine may provide a filtered search service that, when queried for "toys," returns links to pages describing children's toys rather than sex toys, without necessarily reporting that certain sites have been blocked.[93]

The availability of voluntary access controls to parents and teachers is widely viewed as socially beneficial, since it gives control to people who can tailor restrictions to individual and local needs. In a world of perfect transparency and competition, such control imposed by IAPs or search engines may be unproblematic.

But in practice, a number of reasons suggest that these access controls might be less than ideal.

- First, consumers may have a hard time determining which blocks are in their own best interest, as the criteria for selection may not be transparent or readily understandable.[94]
- Second, even if the criteria were transparent, the present architecture would still allow filtering "upstream" (for example, by a search engine) without the consumer knowing (thus a nontransparency not about the rating, but about who is effecting the filter).[95]
- Third, individuals may face a social dilemma about whether to adopt filters. Individuals may themselves prefer to have filtered content (to perfect their own choice), but not want society to have filtered content (to preserve social diversity).[96] If everyone can

[91] Obviously, the most significant concern here would be jurisdictions outside of the United States or outside of places where a strong free speech right exists. The norms that the United States sets for the Net, however, would certainly spill over into those places and our view is that this spillover ought to be reckoned in any regulatory regime.

[92] In fact, voluntary access controls were the main motivation for the creation of PICS.

[93] For a demonstration of Alta Vista's "Family Filter," using ratings from SurfWatch, click on the Family Filter link at <http://www.altavista.com/> (visited Sept. 17, 1999). For a discussion of the implications, see Jonathan Weinberg, *Rating the Net*, 19 HASTINGS *Comm. & Ent.* L.J. 453 (1997).

[94] *See* Rikki McGinty, *Safety Online: Will It Impede Free Speech?*, MEDIA DAILY, Dec. 5, 1997.

[95] *Cf.* Weinberg, *supra* note 93, at n.108.

[96] *See* CASS R. SUNSTEIN, DEMOCRACY AND THE PROBLEM OF FREE SPEECH 21-23 (1995).

easily satisfy their individual preference for filtering, the collective preference for social diversity may be ignored.

- Fourth, the very act of labeling can have destructive consequences for the evolution of ideas, at least if those labels are exclusive in form or in fact. As Niva Elkin-Koren describes, one great virtue of the Internet is its democratization of the process for drawing categories.[97] Rather than labels imposed by a librarian, search engines allow the users to construct different ways of pulling the material together.
- Finally, if IAPs bundle filters with service, then the choice among filters might be less robust than ideal. Put another way, in practice, the competition among filters may not be sufficiently diverse. This could yield very broad filters, which, if common, could create secondary impacts on the variety of speech available on the Internet — since senders may tailor their speech to what will pass the filters.

These secondary effects — a slippery slope of regulation and potentially chilling voluntary uses of labels — have led one of the authors previously to describe PICS, which provides the technical infrastructure for labeling, as "the devil."[98] The other author (one of the developers of PICS) believes that the net impact of a widespread labeling infrastructure would be positive, because of the many positive voluntary uses.[99]

But whether one supports labels for these secondary uses or not, we both acknowledge that the consequence of these labels is the enabling of this secondary use. And if one were sufficiently troubled by this secondary use — as for example the ACLU and other civil rights organizations are[100] — then this secondary consequence might well affect one's judgment about whether a law mandating KMB is preferable to a world with private labels. In other words, if part of the motivation for private labels comes from the need to protect kids, that motivation would be undermined if there were other ways to protect kids.

[97] *See* Niva Elkin-Koren, *Cyberlaw and Social Change: A Democratic Approach to Copyright Law in Cyberspace*, 14 CARDOZO ARTS & ENT. L.J. 215, 238-40 (1996).

[98] Lawrence Lessig, *Tyranny in the Infrastructure*, WIRED, July 1997, at 96.

[99] *See* Paul Resnick, *Filtering Information on the Internet*, SCI. AM., March 1997, at 62; *PICS, Censorship, & Intellectual Freedom FAQ* (Paul Resnick, ed.) (last modified Aug. 4, 1999) <http://www.w3.org/PICS/PICS-FAQ-980126.html>.

[100] *See* Ann Beeson & Chris Hansen, *Fahrenheit 451.2: Is Cyberspace Burning? How Rating and Blocking Proposals May Torch Free Speech on the Internet* (visited Oct. 5, 1999) <http://www.aclu.org/issues/cyber/burning.html>.

IV. EMAIL SPAM CONTROLS

We now turn our attention to controls on the distribution of spam email.[101] On the surface, the problem of spam may seem quite different from those we have considered earlier. With spam the incentives of the recipients and the regulator are aligned (neither wants the spam). Only the sender has a contrary interest.

Our model of mandated access controls, however, applies equally well. Regulations in both cases define some information exchanges as blocked. They also must define which parties carry the responsibility for providing the information necessary to enable that blocking. The effectiveness of any regime will depend on the ability of the responsible parties to comply (especially whether the party responsible for blocking has sufficient information to decide which items to block) and on the enforceability of the regulations (especially the ability to reach senders and intermediaries in other jurisdictions).

The sensitivity analysis for spam controls is also analogous to the analysis of access controls. For example, various architectural changes could make recipient type and jurisdiction information available to senders. Since email transmission protocols do not involve a two-way communication session between sender and receiver,[102] recipients would have to preregister their types and jurisdictions with some server that senders would check with prior to sending mass mailings. If the sender cannot determine a recipient's type and jurisdiction, the default may be either to permit or prohibit sending. A "permitted unless prohibited" regime would be similar to current "opt out" lists for which people can now register. Senders, however, would be required to use these lists, either by sending their spam through a remailer that excludes those on the list, or by checking their recipients against the list.[103] As with other schemes for tagging people, secondary effects exist that we should account for here. For example, this system could facilitate the use of opt-out lists for other purposes, or the use of the registry infrastructure to facilitate creation of other lists.

Alternately, senders may be required to label their messages if they are spam (according to the definitions of various jurisdictions) to enable either mandated or voluntary filtering by intermediaries or re-

[101] Spam is defined roughly as unsolicited commercial email. Various proposed regulations have grappled with how to operationalize this definition. For an excellent analysis, see ALAN SCHWARTZ & SIMSON GARFINKEL, STOPPING SPAM: STAMPING OUT UNWANTED EMAIL AND NEWS POSTINGS (1998).

[102] In email transmission protocols such as SMTP, there is no preliminary end-to-end session set up that would enable the recipient to send a certificate indicating its jurisdiction or type. Email transmission is effectively a one-way communication.

[103] This is the essence of one part of Senator Robert Torricelli's Electronic Mailbox Protection Act of 1997, S. 875, 105th Cong.

cipients.[104] And here also possible secondary effects exist — for example, the use of spam tags for other purposes (such as monitoring what employees receive), or added pressure to label for other criteria (such as sex or politics).

One essential difference between spam controls and access controls is that recipients do not want to receive spam, but some people may want to receive obscene or other materials over which governments may mandate access controls. Thus it is not necessary to mandate that spam recipients filter out certain messages or that they reveal their type and jurisdiction. Recipients will do so voluntarily if it will reduce spam. Hence, sender blocking under a "permitted unless prohibited" regime may be more effective for spam control than for access controls. Where intermediaries carry the blocking responsibilities, however, mandates may be necessary. Otherwise, spam senders may pay intermediaries enough to make them prefer not to filter.

Any regulations of sender behavior suffer from the same jurisdictional problems that plague efforts to mandate access controls. A small country may refuse to enforce other countries' regulations on its senders and thus become a spam haven. One possible solution would place the responsibility on intermediaries not to forward improper email (either spam or untagged spam, depending on the regulation). If most jurisdictions imposed such a regulation, it would put pressure on intermediaries in other jurisdictions to voluntarily impose similar restrictions, else they would not be able to pass messages to restrictive jurisdictions. If, however, not enough jurisdictions imposed such a regulation, there could be an overall loss of email connectivity, which would be inconvenient for legitimate users.

These phenomena have already occurred as a result of voluntary attempts by IAPs to restrict spam. Some large IAPs have refused to accept email originating from servers that are known spam sources (they may also be sources of legitimate email). This results in irate customers of both IAPs who resultantly cannot communicate non-spam messages to each other. The hope is that these irate customers will pressure the spam source to mend its ways, and this pressure has indeed been effective on occasion.[105]

Given the difficulties and potential side effects of mandated spam controls, it is worth exploring architectural changes to enable voluntary access controls. Monetary filtering presents one intriguing possi-

[104] Without sender tagging, automated filtering is still possible, but will never be completely accurate at separating spam from non-spam messages. This is the essence of Senator Frank Murkowski's Unsolicited Commercial Electronic Mail Choice Act of 1997, S. 771, 105th Cong.

[105] *See* SCHWARTZ & GARFINKEL, *supra* note 101, at 85-93.

bility. A message sender would attach some ecash to each message.[106] If the recipient is unhappy about having received a message, she can keep the money. Each recipient would have a filter automatically reject messages with insufficient money attached, with the sender being notified of the minimum amount required to get the message through. If widely adopted, this scheme would set up market incentives to voluntarily curtail sending of unwanted messages.

But here again, the secondary consequences are crucial. The spam-stamp could constrain noncommercial speech as much as commercial speech. It would restrict political messages as well as nonpolitical messages. Thus how one changes the architecture to solve the problem of spam will have effects far beyond the problem of spam. Our point has been to highlight these effects, and to argue that they must be accounted in any decision to regulate.

CONCLUSION

This Article has proposed an abstract model of mandated access controls, and it has applied that model to two concrete cases. The model includes three types of actors: senders, intermediaries, and recipients. Control decisions are based on three types of information: the item, the recipient's jurisdiction, and the recipient's type.

With the architecture of today's Internet, senders are ignorant of the recipient's jurisdiction and type, recipients are ignorant of an item's type, and intermediaries are ignorant of both. It is easy to see, then, why, with today's Internet architecture, governments are having a hard time mandating access controls. Any party on whom responsibility might be placed has insufficient information to carry out that responsibility.

While the Internet's architecture is relatively entrenched, it is not absolutely immutable. Our abstract model suggests the types of changes that could enhance regulability. Senders could be given more information about recipient jurisdiction and type, either through recipients providing certificates, or through a database mapping IP addresses to jurisdictions. Recipients could be given more information about item types, either through senders providing labels or through government preclearance lists of permitted or prohibited items.

Table 1 summarizes this sensitivity analysis. Since the two interventions are analogous, the analyses of their costs and effectiveness are analogous as well. In either case, there will be a natural incentive to provide information if the default action of the responsible party is to block access unless the information is provided (a "prohibited un-

[106] Ecash is simply a digital object that could be attached, as a file is, to an email message. For a discussion of ecash, see David Chaum, *Achieving Electronic Privacy*, SCI. AM., Aug. 1992, at 96, 96-97.

less permitted" regime). Otherwise, there will be no natural incentive, and the government will have to require the provision of that information.

TABLE 1: SENSITIVITY TABLE

	Sender	**Intermediary**	**Recipient**
Missing information	• jurisdiction • recipient type	• jurisdiction • recipient type • content of item	• content of item
Possible architectural and legal changes	• IP to geography mapping, jurisdiction certificates • recipient-type certificates • preclearance, thesauri	As for sender and recipient, plus: • responsibility to assure sender/recipient compliance • use of proxies and application gateways	• preclearance • sender's self-rating • third-party rating
Consequences	Enables general regulability of behavior on the Net based on recipient type and jurisdiction	Enables private parties (IAPs and ISPs) to regulate behavior on the Net	Enables greater control of speech content on the Net beyond that initially required by governments
Notes	Enforcement problems significant, if sender outside the jurisdiction	Enforcement is easier: since IAPs are not mobile, there are few players, and they have commercial assets	Enforcement problem: number of recipients leads to selective enforcement, though a greater portion of the regulable public is within a given jurisdiction

The secondary effects of these two infrastructures are also analogous, but quite different. The by-product of a certificate regime is a general ability to regulate based on jurisdiction and recipient characteristics, even for issues beyond content control, such as taxation and privacy. Such a regime also enables senders voluntarily to exclude recipients based on jurisdiction or type, a facility which might be used for negative as well as positive purposes. The by-product of a widely used labeling infrastructure is a general ability to regulate based on item characteristics, even characteristics that governments have no legitimate reason to regulate. Such a regime also enables intermediaries and recipients to exclude voluntarily some item types, a facility that may empower parents and teachers but may also be overused if it is poorly understood or difficult to configure.

If intermediaries have responsibility for blocking, they will need both types of information. In addition, architectural changes will be necessary to enable application layer blocking of individual items rather than cruder network layer blocking of all traffic from or to an IP address. A requirement of application layer blocking, however, introduces significant costs in terms of openness to innovation and vulnerability to hardware and software failures. Intermediaries, then, arguably constitute the most costly party upon which to impose responsibility. On the other hand, they are the most easily regulated, since there are fewer of them, they are more stable, they have assets, and their governing jurisdictions are clear.

While our sensitivity analysis does suggest consequences that might not have been readily seen, our ultimate conclusion is one others have reached as well. It will be difficult for governments to mandate access controls for the Internet. Given today's architecture, any such mandates would of necessity be draconian or ineffective. Changes to the technical infrastructure or social practices could enhance regulability, although such changes would both entail direct costs and create secondary by-products whose value is debatable. Given the significant costs of any such architectural change, governments must answer the fundamental question regarding the importance of such changes — perhaps a lessening of governments' traditional power to control the distribution of harmful information would be preferable.