The Constitutionality of Mandated Access Control: A Model

Lawrence Lessig
Paul Resnick
Circulating Draft 4[*]

forthcoming, MICHIGAN LAW REVIEW (fall, 1999)

---

[*] Jack N. and Lillian R. Berkman Professor for Entrepreneurial Legal Studies, Harvard Law School; Associate Professor, University of Michigan School of Information. Thanks to Lorrie Cranor for initially suggesting the symmetry between tagging speech and tagging people. Thanks also to Robert Cooter, Mark Lemley, 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 in the Proceedings of the Telecommunications Policy Research Conference (1998).

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 to 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 is meaningful for speech and regulation within a single jurisdiction. 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 is "political speech" in the United States (Nazi speech) is banned in Germany; what is "obscene" speech in Tennessee is permitted in Holland; what is porn in Japan is child porn in the United States; what is "harmful to minors" in Bavaria is Disney in New York. Every jurisdiction has some speech to which access is controlled,[3] 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. For within cyberspace, mandated access controls are extremely difficult. If access control requires knowing (a) the identity of the speaker and receiver, (b) the jurisdiction 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.[4] As a result, real space laws cannot readily be translated into the context of cyberspace. Or put

---

[1] *See* Lawrence Lessig, *What Things Regulate Speech: CDA 2.0 vs. Filtering*, JURIMETRICS Summer 1998, at 10.

[2] Ginsberg v. New York, 390 U.S. 629, 88 S. Ct. 1274 (1968).

[3] We reserve the term 'censorship' for blanket restrictions on the distribution of speech that apply regardless of the recipient or context. Access control is a broader concept that includes censorship but also restrictions on speech that may depend on the recipient or context.

[4] The content is not easily identified primarily because content on the net is broken into packets, and not all packets will necessarily pass through the same channels. Even if they did, the content could be encrypted, which would further complicate identification.

another way, the initial architecture of cyberspace in effect places all speech within category (1).

One possible response to this change caused by initial architecture of the 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 re-establishing mandated access controls over certain kinds of speech, now published on the internet. In the United States, this speech is sex related;[5] 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 are moving to re-regulate access to speech in cyberspace, so as to reestablish local control.

We take as given this passion for re-regulation; we consider it a feature of the current political reality surrounding cyberspace. This reality should push us to consider the options that regulators face — not because regulators need to be encouraged, but because we should understand the consequences of any particular regulatory strategy. Some strategies are more costly than others; some strike at features of the net more fundamental than others.

This inquiry is particularly salient in the United States just now. In what may have become a bi-annual event, the United States Congress in 1998 passed its second attempt at regulating "indecent speech" on the net — the Child Online Protection Act (COPA). Its first statute — the Communications Decency Act of 1996 (CDA) — was struck down by the Supreme Court in 1997.[8] Now two years later, a federal district court in Philadelphia has enjoined

---

[5] *See* Telecommunications Act of 1996, Pub. L. No. 104-104, Title V, 110 Stat. 56, 133-43 (1996) (Communications Decency Act); Child Online Protection Act, Pub. L. No. 105-277, 112 Stat. 2681 (1998) (to be codified at 47 U.S.C. § 231).

[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 (1998).

[7] *See* Geremie R. Barmé & Sang Ye, *The Great Firewall of China*, Wired, June 1997, at 138 and Philip Shenon, *2-Edged Sword: Asian Regimes On the Internet*, N.Y. Times, May 29, 1995, §1 at 1.

[8] Reno v. American Civil Liberties Union, 521 U.S. 844, 117 S.Ct. 2329 (1997).

enforcement of COPA.[9] And if the ACLU succeeds in striking the statute again, 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. Reno III*.

It may be that Congress will never pass a statute that satisfies the Court,[10] but we think it could. There is a "decency act" that would pass constitutional muster. It is just not COPA. But to see why, and to understand the alternative, requires a broader view. It requires an analysis that makes clear the different values at stake.

Our aim in this essay is just such a perspective. We offer a model of mandated access control that will clarify the issues in play. And 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 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.

The 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 (e.g., TCP/IP), (b) its standards and standard applications (e.g., 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, and are in part a function of both direct and indirect regulation by law. Thus in this essay we ask first how access can be controlled given the existing array of legal and architectural constraints. And we then consider how changes in the current array might yield a different mix of costs and benefits.

---

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

[10] A cynic might believe that this repetition is no accident. After all, Congress gets rewarded for what it passes, now what sticks. Protecting kids is great politics. Why do it only once, the cynic might ask, when one can do it every two years?

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, the cost to participants, whether sender, or receiver, or intermediary; third, the costs to a system of "free speech"; 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 are 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. This range of regulations will therefore be illustrative of the more general problem of access control across jurisdictions. It is therefore meet for the complex analysis that we believe access-control requires.

We then apply our model to COPA. COPA, we believe, is significantly narrower in reach than the original CDA. Congress was, we believe, responsive to the Supreme Court's opinion in *Reno*. But there is a structural feature of COPA that does render it unconstitutional, at least when compared to a second possible statute that would have achieved Congress's legitimate end.[11] Those attacking the statute are not in a position to suggest such an alternative. Their position is that private regulation is better than any law. But while we agree that private regulation is 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.

The reason is the focus of our last section. For there we consider the unintended consequences of the various regulatory strategies proposed. Our argument is 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

---

[11] Congress was aware of this alternative. See 144 Cong. Rec. S 12741, *S12795 (comments of Senator Leahy).

secondary effects — will continue to burden the culture of the net.

A MODEL OF ACCESS CONTROL

Elements

In our model of mandated access control (MAC), we consider three relevant actors — a sender (S), a recipient (R), and an intermediary (I). The sender is the party who makes available the relevant speech; the recipient is the party who gets access to the relevant speech; and an intermediary is an entity that stands between the sender and the recipient. As these definitions suggest, nothing in our description hangs upon whether the sender actually *sends* material to the recipient. The model is agnostic about 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 that the sender knows about the contents of the item that is being sent. We assume the recipient has information about who she is, and where she resides. And finally we assume that 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 is making 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 being violated. And fourth, it sets punishments when these responsibilities are violated. In the balance of this part, we sketch issues relevant to each of these elements of a regulatory regime. In the next part we conduct the sensitivity analysis.

(1) 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) $B(I, R, J) = \{Y, N\}$

Where I = item type, R = recipient type, and J = jurisdiction type, and B() 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 is 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 are a floor recipient type (anything that is permitted to children is permitted to adults as well), and a Bible Belt small town may be a floor jurisdiction (anything that is permitted there would be permissible everywhere). More formally, with J_f denoting a floor jurisdiction:

(b: floor type) For all I, J:
B(I, 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 is permitted to children in a floor jurisdiction is permitted to everyone in every jurisdiction:

(d: floor recipient and jurisdiction) For any I:
B(I, 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 crypto laws regulate a receiver's right to receive such material.[12] For simplicity, however, we will focus on blocked exchanges in the recipient jurisdiction alone. This focus will be significant in the context of enforcement, since governments can more easily control their own populations than populations in other jurisdictions.

A jurisdiction,[13] on this model of blocked transactions, may specify that a particular transaction is to be blocked in at least two different ways.

---

[12] *See* STEWART A. BAKER & PAUL R. HURST, THE LIMITS OF TRUST 130 (1998).

[13] There is an important ambiguity in the concept of "jurisdiction" that we

1. The jurisdiction might publish criteria defining what is to 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 may classify specific items as acceptable or blocked for particular recipient types. Such classifications would function as a form of pre-clearance for allowable speech, with the government promising not to prosecute parties for decisions made in good faith based on the pre-clearance. Alternatively, the jurisdiction could define a list of prohibited speech. In either case, the determinations of acceptability may occur through a judicial or administrative process, or the jurisdiction may delegate its authority to an independent rating service.[14] A jurisdiction might even rely on a computer program to provide an initial classification of the speech at issue, and publish that classification as a pre-clearance, perhaps with a stipulation that the initial classification might be changed in the future after human review.

In the American context, the ordinary procedure follows 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.[15] It is unclear whether a regime of voluntary pre-clearance would be permissible. On the surface, it would seem that any step that would reduce the uncertainty surrounding the distribution of speech would be speech enhancing. On the margin, if a speaker could be certain that her speech was permissible, she would be more likely to utter it than if she faced the risk that it would be found to be illegal. But some

---

are ignoring here. Some rules depend upon where the person is when he or she acts, rather than the jurisdiction where the person is a citizen. If the drinking age in one state is 21, it does not matter than 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.

[14] An example would be CyberPatrol's CyberNot list. *See* CyberPatrol's *Home Page* (visited January 7, 1999) <http://www.cyberpatrol.com/ >.

[15] *See* Paris Adult Theatre I v. Slaton, 413 U.S. 49, 55 (1972) (Injunction could be used so long as adequate procedures to determine obscenity had been used).

who have considered the matter believe that if this voluntary regime became, in effect, mandatory, with speech not appearing on a pre-clearance list in effect then restricted, it would then become constitutionally suspect.[16] The constitution notwithstanding, we believe that the net effect on speech is unclear: Lower costs could lead to less chilling of speech (if it is clearer what is prohibited and what is not) but to more control on speech (if it results in greater prosecution of improper speech.)

(2) Assignments of Responsibility

The regulator's second step is to 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, are the 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. In this version of our analysis, we consider only exclusive assignments of responsibility for blocking, 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.[17] The law must therefore create an incentive for parties to produce sufficient information to determine whether access should be blocked.

The law's ordinary technique for creating incentives are either property or liability regimes. While a property regime is

---

[16] *See* Frederick Schauer, *Fear, Risk and the First Amendment: Unraveling the "Chilling Effect"*, 58 B.U.L. REV. 685, 725-29 (1978). The closest case is perhaps Bantam Books v. Sullivan, 372 U.S. 58 (1963), where the Court invalidated a "blacklist" Commission. The pre-clearance 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.

[17] Again, the sender does not know the recipient; the recipient does not know the content of the item; the intermediary does not know either. This does not mean that there would not be extreme, and therefore easy, cases. The speaker would certainly know, therefore, for some kinds of speech that it is highly likely to be permitted, or not. Banalities about the weather are fairly safe speech acts anywhere; sadistic child porn is fairly unsafe in most jurisdictions.

conceivable,[18] we focus here on a liability regime. The law can create an incentive to produce the information necessary to determine whether a exchange should be blocked by assigning liability to an actor for failing properly to block a transaction, or by setting a default rule about whether to block a transaction when there is uncertainty.[19]

We consider two such defaults.[20] Under the first default, the sender is liable if she enters a transaction without reliable indicators that in fact the transaction was legal, and that transaction is later determined to be illegal. We call this the "prohibited unless permitted" rule. Because liability turns upon on the steps taken to comply with the law, we believe it is distinct from a prior restraint.[21]

Under the second rule, the sender is liable only if she enters a transaction in the face of indicators that in fact the transaction was illegal, and that transaction is later determined to be illegal. This is the "permitted unless prohibited" rule, and it is equivalent to a rule punishing a specific intent to violate the law.[22] One

---

[18] For an excellent analysis of a property regime for dealing with the access control, see Sean Griffith, Code Solutions and Property Rules (forthcoming, Harvard Law Review 1999).

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

[20] We are not claiming at this point that either default would, for all types of speech, be constitutional under the U.S. constitution. Nor are we speaking about the burdens of proof under a particular statute. We will assume throughout that the state bears the burden for all elements of the charge. *See* Smith v. California, 361 U.S. 147 (1959). Rather than a claim about what is constitutionally possible, our defaults will help clarify the relationship between the proscription and uncertainty. Like Schauer, our objective is to further explore this relationship, and the constitutional implications of uncertainty. *See* Schauer, *supra* note 16, at 725-27.

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

[22] Specific intent is equated to acting knowingly in the Model Penal Code. The relevant section reads:

modification of this latter rule would hold the sender responsible if the sender should have known that the transaction is illegal. This would comport with a negligence standard, and we consider this change where relevant in the analysis below.

These default rules will have significant consequences if there is systematic uncertainty. 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).[23] Thus in the face of uncertainty, the default rule will be important, especially if one default is constitutionally compelled.

Our focus will be 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 it is easier for an intermediary or recipient to determine item types, and block accordingly; if people are tagged, then it is easier for an intermediary or sender to identify recipient and jurisdiction types, and block accordingly.

(3) Monitoring and (4) Enforcement

The regulator's final two steps are first, devising schemes for monitoring compliance, and second, implementing 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.

> 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 exists; 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) (Official Draft 1985).

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

The major issues for enforcement all involve the question whether the target of enforcement can easily, or cheaply, be reached. We assume there are more receivers than senders, so one might believe targeting senders would be cheaper than targeting receivers. This, however, is complicated if the sender is outside the regulating jurisdiction, making the sender sometimes legally, and if not legally, then often practically, beyond the reach of the regulating jurisdiction. The cost of enforcement against aliens, there, may mean that it is cheaper to enforce a rule against receivers than senders.

Whether there are more receivers or listeners, however, there are certainly fewer intermediaries than either. Intermediaries, as we discuss below, may therefore be the optimal target of regulation, even though they have even less information than either the sender or receiver. Again, the savings in enforcing a rule against them may be greater than the cost of their obtaining the necessary information. Thus from a social cost perspective, making them liable may be efficient.

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, 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—e.g., more control at less free speech cost. This is 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.

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. The obvious case is material that is "harmful to minors" to minors — many states require that providers of such material keep it from kids.[24] But the possibility is more general than this: Rules regulating SEC filings,

[24] Reno v. American Civil Liberties Union, 521 U.S. 844, 117 S.Ct. 2329, 2352-53 & n.1 (1997) (O'Connor, J., dissenting).

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 are costly. There is no simple way to identify the jurisdiction within which the recipient resides,[25] and no cheap way to be certain of characteristics of the individual. The rule would therefore be quite costly to a speaker — unconstitutionally costly is the suggestion of *Reno v. ACLU*, though differently costly under each of our two default rules.

Under the "prohibited unless permitted" rule, the cost is to "free speech" interests. The burden of determining eligibility is likely to present a significant chill on the speaker's speech.[26] 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 using an area-code on a telephone number as a proxy. And the sender would need to rely upon proxies from credentials (such as a credit card) to guess whether the individual is 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.[27] 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

---

[25] 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.

[26] 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* 16, at 693.

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

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.[28] But *Reno* did not consider the "harmful to minors" standard — or as described by some, the obscene-as-to-minors standard[29] — and there is no clear indication by the Supreme Court that the test would be different.

If, on the other hand, the rule is "permitted unless prohibited," the cost is 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 10 year old boy doesn't look much like a 20 year old man), in cyberspace, such facts are not self-authenticating. 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 requirement 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."

Sensitivity

Some of the burden on the sender could be reduced by certain 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 more cheaply to identify 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 18.) The changes described

---

[28] *See* Reno v. ACLU, 117 S. Ct. 2329, 1997 U.S. LEXIS 4037, 24 (1997).

[29] See, e.g., Upper Midwest Booksellers Assoc. v. City of Minneapolis, 780 F.2d 1389 (4th Cir. 1986); see also M.S. News v. Casado, 721 F.2d 1281 (10th Cir. 1982) (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. Material is obscene as to minors if it is: patently offensive, appeals to minors' prurient interest, and completely lacks socially redeeming value for minors. Reno, 117 S. Ct. at 2356 (O'Connor, J., concurring in the judgment in part and dissenting in part).

here would facilitate the sender knowing both facts at a relatively cheap cost.

The first technique relies on digital certificates.[30] In the standard model of certificates, certificates identify who someone is. They are digital objects cryptographically signed by a certificate authority. The dominant use of such certificates today is to certify the identity of the holder. This is the model, for example, of the Verisign Digital ID, which Verisign describes as a "driver's license for the Internet."[31]

But there is no reason that the same technology couldn't 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 (CA) could then certify that X is from Massachusetts, and that X is over the age of 18, without identifying who X is.[32] 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.

We can call this a "credentialling" solution.[33] 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 18. Recipients interested in receiving restricted materials will have an

---

[30] *See* A. Michael Froomkin, *The Essential Role of Trusted Third Parties in Electronic Commerce*, 75 OR. L. REV. 49 (1996).

[31] *See Verisign* (visited August 22, 1998) <http://www.verisign.com/>.

[32] 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. OF THE ACM 1030 (1985).

[33] 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.

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 will 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 is the obvious 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. 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 correctly to block exchanges that would otherwise have been permitted by default.[34]

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 would be required to ask only if the material is illegal in at least one jurisdiction. This regime would still burden somewhat those recipients living in jurisdictions where the speech was wholly legal; its viability would rest then upon the significance of that burden.[35] Alternatively, the rule could require that intermediaries

---

[34] This regime is close to the regime we discuss below when considering COPA. See *infra* text at notes 53-67.

[35] 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.

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 cost of monitoring and compliance might be lower than if the same role was being performed by the state. The intermediary's advantage is not over the primary conduct — certainly receivers are 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 jurisdiction into which speech was to be sent would be an IP map — a table that would give a rough approximation of the location of the recipient's computer.[36] 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 suffice sufficiently to segregate restrictive jurisdictions from nonrestrictive.[37]

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.

---

[36] Currently, the InterNIC maintains a database of which organization each IP address was allocated to. 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 this database be used as a starting point for developing an IP to jurisdiction mapping. *See* Philip McCrea, Bob Smart, & Mark Andrews, *Blocking Content on the Internet: A Technical Perspective*, Appendix 5 (visited August 22, 1998) <http://www.noie.gov.au/reports/blocking/index.html>.

[37] We note that already, companies such as Microsoft are using IP addresses to assure themselves that the user is within the United States, so that Microsoft does not become an "exporter" of high grade encryption technology. *See Internet Explorer Products Download*: *Microsoft Strong Encryption Products (US and Canada Only)* (visited August 23, 1998) <http://mssecure.www.conxion.com/cgi-bin/ieitar.pl>.

But there are important social costs associated with this IP-to-geography mapping. These 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 mapping 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 pre-clearance technology. While we presume that the sender knows about its speech, it may not understand the classification scheme of every legal jurisdiction. As discussed above,[38] pre-clearance of the sender's materials can eliminate the sender's uncertainty. If the pre-clearance is judicial, the cost of the clearance would still be high. It may be possible to use automated regimes to facilitate this certification, so long as the government had the right, prospectively, to change its mind about a certification.

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 version of the base jurisdiction model that we described in Equation C.[39]

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, it is the recipient who is liable 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

---

[38] See *supra* page 5.

[39] See *supra* page 5.

about the law of its jurisdiction, and about its own recipient type. But there are obvious disadvantages as well. The recipient is 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.[40] The receiver does not. This 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[41] in the very act of determining whether a particular exchange complies with the law. And if the rule is "permitted unless prohibited," then restrictions are likely to be completely ineffective.

A second problem with placing liability on the receiver is the costs of classification. There will be more receivers than senders, and thus this rule shifts the cost of classification to the many, rather than the few. This will result in either too much or too little classification. 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 are likely to push the classifier to classify not 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 enforcement generally, of course, depends upon whether the alternative targets — senders, or intermediaries — are more easily regulated. If they are primarily outside the jurisdiction regulating access, then regulating recipients may be less costly then regulating senders or intermediaries.

---

[40] 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 v. Compuserve, 776 F. Supp. 135 (S.D.N.Y. 1991). We would consider such a "sender" to be an intermediary in our analysis.

[41] 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.

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 — pre-screening — and here the same techniques for reducing the costs of pre-clearance would apply, including the use of automatic text classification and delegation of the pre-clearance powers to an independent third-party rater. As we mentioned before, however, there remains a concern about the constitutionality of even a voluntary pre-clearance regime. In the American context, despite the reduction in uncertainty, this might be 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 sufficiently (on dimensions such as sex related) that were detailed enough to infer whether it should be blocked.[42] This solution simply inverts the certificate solution — since here it is the sender that is offering a "certificate" and the receiver who is relying, while in the case above, it was the recipient providing the certificate, and the sender who was relying. The analysis is also analogous.

Under a "prohibited unless permitted" regime, the labels would convey information that the speech was permitted (e.g., no sex or hate speech). Recipients would be given immunity if they in good faith rely upon a sender's labels to determine that an 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.

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

---

[42] The labels could be expressed in PICS format (*see http://www.w3.org/PICS)* or the new RDF format (*see http://www.w3.org/RDF*), and distributed along with the items.

legal access.[43] 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. Some have argued that they would not,[44] but we believe this is a close question. To reduce the cost to senders of labeling, a government might subsidize third party rating or itself produce suggested ratings. In the United States, its ratings could not be treated as definitive[45] in such a system, but may be 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 26 carry IDs to purchase cigarettes, even though the prohibition reaches only those 18 and under). Even here, however, the requirement raises difficult questions, since it is requiring speech by the sender in the form of a label even if the underlying speech is clearly legal in the jurisdiction into which it is being sent. Thus the most restrictive jurisdiction would in effect determine whether the speaker must label.

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. There is one important asymmetry, 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 make only the sender and not the access provider responsible for any inaccuracies in the label.

---

[43] If many people voluntarily adopted a prohibited unless permitted filter, then the market demand for labels might be a sufficient incentive to 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.

[44] *See* R. Polk Wagner, *Filters and the First Amendment* (visited August 22, 1998) <http://www.pobox.com/~polk/filters.pdf>.

[45] In no case could the government's own ratings be determinative of whether speech were delivered or not, absent a judicial finding. *See* Rowan v. U.S. Post Office, 397 U.S. 728 (1970).

We also note an interesting constitutional asymmetry between requirements of senders providing labels and recipients providing certificates. It seems clear that there could be no U.S. law that required receivers not to receive any speech unless it were plain that the speech was legal. (A rule, that is, that would punish the receiver if she received speech without a clear indication that it was not illegal.) Even with respect to obscenity, that restriction would be overbroad, or fail a minimal *mens rea* analysis. It is at least arguable, however, that a law that required senders to check every digital ID before sending material would not be constitutionally invalid. Analytically these two regimes are quite similar: In both, the transaction is conditioned upon verification of its legality. But the burden on the sender is likely less constitutionally troubling than the burden on the receiver.

Beyond the constitutional issues, 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.

Intermediary responsible for blocking

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

But this is misleading. Intermediaries are a cheap target of regulation. There are fewer of them than receivers or senders, and they are typically more stable, or harder to move. Just as it is easier for the government to regulate telephone companies than it is to regulate 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 is more likely to get innovation in the methods of compliance.

In addition to a lack of information, intermediaries may have limited capabilities for implementing blocks. Blocking can either be implemented at the application layer (e.g., web page requests) or at the network layer (i.e., individual packets). 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.[46]

We consider two types of intermediaries. One type is an Internet access or service provider or an employer or a school (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.[47] 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

---

[46] For a more complete description of application layer and network layer blocking, *see* McCrea, *supra* note 37.

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

as tunneling, where the contents of one packet are wrapped inside another packet, may bypass a network layer block.[48]

Sensitivity

Given how fundamental the architectural features are that yield the conclusion that intermediaries are not in a position to control access, one might well conclude that it would be unadvisable to make any changes to increase their ability to control. However, because intermediaries are also practically easier for a jurisdiction to regulate, 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 pre-clearance 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 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 counter-acting performance gain from caching. Second, messages may be encrypted for privacy or security purposes (e.g., 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.[49] The Internet's current architecture has enabled experimentation and rapid

---

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

[49] Many corporate firewalls do 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 75 (1994).

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[50]). One final cost might come in the form of reliability. It is relatively easy for a service provider to provide multiple routers, so that network layer service is not interrupted if a router is temporarily disabled. It may be more costly to arrange for continued service if an application layer gateway is temporarily disabled.[51]

MIXED STRATEGIES: "KIDS" AND COPA

In our analysis so far, we have assumed that the strategies would either place access responsibility on receivers, senders, or intermediaries, with those responsibilities in turn creating incentives for the other actors in the system. A requirement that senders check for IDs is sufficient to create an incentive for recipients to secure IDs; a requirement that recipients block illegal content may be sufficient to create the incentive for senders to label content.

In both cases, the target of the regulation has an obvious incentive to disobey the regulation. Senders lose business if they must check for IDs; recipients lose access if they must block illegal content. Thus in both cases, the regulator must threaten a sufficient punishment to induce the target to obey the regulation.

Regulations designed to protect kids, however, are a special case.[52] If one assumes that the parent is the relevant "recipient,"

---

[50] 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 What Is ICQ?* (visited August 23, 1998) <http://www.icq.com/whatisicq.html> and *How To Use ICQ* (visited August 23, 1998) <http://www.icq.com/icqtour/new-quicktour.html>.

[51] *See* McCrea, *supra* note 37.

[52] 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" were considered. Likewise, the problem is far more difficult if the recipient always has an incentive to evade the regulation. Regulating "kids" is a special case because at

then unlike the general case, this recipient has an incentive to facilitate blocking — if indeed, blocking access is what that parent wants.[53] Sender based regulations would therefore be cheaper when the recipient has an incentive to comply with the restrictions of the sender.

This difference has a constitutional significance when one considers regulations designed to block access to kids. For consider one alternative to the regulation prescribed by Congress in CDA and COPA — a regulation that required senders to block self-identifying kids, rather than a regulation that required senders to block content unless a receiver certified he was an adult. The following hypothetical statute will suggest the idea:

1. *Kids-Mode-Browsers [KMB]:* Manufacturers of browsers will enable those browsers to browse in "kids-mode." When enabled, "kids-mode" will signal to servers that the user is a minor. The browser shall also enable password protection for non-kids-mode browsing. It shall also disable any data collection about the user of a kids-mode browser. In particular, it shall not transmit to a site any identifying data about the user.
2. *Server Responsibility:* When a server detects a kids-mode client, it shall (1) block that client from any material properly deemed "harmful to minors" and (2) refrain from collecting any identification data about the user, except data necessary to process user requests (i.e., 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 the type of browser it is. 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,"[54] and the site you were viewing

---

least sometimes, the parent has an interest to enforce the regulation.

[53] The Court in *Reno* made it clear that the relevant question is whether parents are enabled in protecting kids, not whether the state is. If a parent decides to give kids access, that decision cannot, for the range of speech being discussed here, be overridden by the state.

[54] "Cookies are a general mechanism which server side 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/served

before you switched to that site. This statute would simply require that manufacturers add one more bit of information to that set — whether the user was a kid.

Of course this statute imposes burdens, but the burdens are far less (constitutionally) significant than the burdens of CDA or COPA. The primary burden of this statute rests upon software manufacturers, though that burden would be slight, and not constitutionally suspect. The programming change required to comply would be trivial; software manufacturers have no constitutional claim against manufacturing regulations.

The statute also burdens senders, who must now discriminate on the basis of whether the client is a kid. But that burden is also trivial. Relative to the existing constitutional baseline,[55] it would be a trivial change for servers to check for the existence of a kids-mode signal.

Finally, one might believe the statute "burdens" parents — since they would have to turn the kids-mode browsing on. But the burden here is not a legal burden; the burden is the difficulty of checking a preference box. That, we believe, is far less significant than the alternatives, say, of purchasing and installing blocking software.

Now compare this hypothetical regulation to the regulation in COPA. COPA was modeled on the CDA, but regulates more narrowly than the CDA. Like the CDA, COPA puts the burden on the speaker to avoid speaking improperly to kids. But unlike the CDA, it 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. The statute is thus narrower: (1) in the breadth of speech regulated ("harmful to minors" rather than "indecent"), (2) in the scope of speakers covered (it doesn't reach noncommercial providers), and (3) in the range of the internet affected (it does not reach newsgroups, or chat rooms).

Similarly with the defenses provided to a speaker by the statute.

---

applications." Http://home.netscape.com/newsref/std/cookie_spec.html. *See generally* http://www.cookiecentral.com.

[55] *Ginsberg* implies that suppliers can be burdened to separate "harmful to minor" speech from other speech.

COPA's defenses are broader than the defenses under CDA.[56] Under COPA a provider has a defense if he "in good faith … restricted access by minors …

> 1. by requiring the use of a credit card, debit account, adult access code, or adult personal identification number;
> 2. by accepting a digital certificate that verifies age; or
> 3. by any other reasonable measures that are feasible under available technology.

Section 231(c)(2) adds immunity from prosecution to the substantive defense of §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).

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. And second, the catchall category of technologies ("by any other reasonable measures that are feasible under available technology") is broader that the parallel in the CDA. The CDA required that these other technologies be "reasonable, effective, and appropriate."[57] The court read this standard not as an ordinary tort standard, but as an absolute effectiveness requirement. COPA's test, by contrast, is 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."

These differences evince what Congress said it was doing: responding to the concerns of the Supreme Court in *ACLU v. Reno*.[58] It has followed the outline sketched by Justice O'Connor's concurrence of what she thought a constitutional regulation would be. If the Court is eager to reward legislative obedience, it might well feel itself compelled to uphold Congress' 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

---

[56] Child Online Protection Act §231(c)(1)

[57] 47 U.S.C.A. §225(e)(5)(A) (1997).

[58] 144 Cong. Rec. H9902-01, 1998 WL 694693 (Cong. Rec.).

burdensome.

In our view, this analysis is mistaken. We agree with Professor Volokh that question the Supreme Court has asked is not whether the regulation is "too burdensome."[59] That is a form of analysis restricted so far to abortion regulation,[60] and perhaps to the dormant commerce clause.[61] Rather the question the Supreme Court has asked in this context is whether the regulation is more burdensome than needed.[62] And if that question were answered by asking whether this regulation mandated the least burdensome *adult*-ID regime possible, then we believe this statute does impose the smallest adult-ID regime burden possible.

This is true because unlike CDA, COPA includes a catch-all provision that permits "any other reasonable measures that are feasible under available technology." This is a clear invocation of traditional negligence standards. It requires only the technology that is at the time reasonably effective. CDA required that the technology *be* effective; this requires only that it be reasonably effective, given the existing technology. In effect, by definition then COPA is the least burdensome *adult-ID regime*.

The *adult-ID* regime is not the only ID regime possible. As we outlined at the beginning of this section, one alternative would be, in effect, a *kids-ID* regime. By requiring that manufacturers of browser technologies enable the signaling of a user who is a minor, Congress would be facilitating the identification of kids. And by requiring that servers segregate based on that analysis, Congress would be enacting a sender based regulation. But this sender based regulation would be far less burdensome than CDA or COPA — or indeed any adult-ID regime. And, we believe, it would be just as effective.

The advantages of the kids-ID regime are many.

---

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

[60] Planned Parenthood v. Casey, 505 U.S. 833 (1992).

[61] Pike v. Bruce Church, Inc, 397 US 137 (1970).

[62] See Ginsberg v. New York, 390 U.S. 629 (1968) (upholding New York statute that required keeping of material harmful to minors from minors); Sable Communications v. FCC, 492 U.S. 115 (1989) (plurality permitting regulation of radio to protect kids).

- First, the burden of signaling that the user was a kid would be far less costly than the burden of signaling that the user was an adult. An adult-ID needs to be verified (since it is granting access which otherwise would not be permitted, there would be an incentive to cheat); a kids-ID would only blocking access; there would be no incentive to lie.
- Second, the absolute number of people burdened by the regulation is likely to be less. 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 fall on a much smaller proportion of the population, and 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.[63]
- Fourth, the greatest burden of this regulation falls, in a sense, on the cheapest cost avoider — browser manufacturers. It also avoids imposing any cost on recipients without kids. [64]
- Fifth, rather than an elaborate identification system, maintained either by companies such as AdultCheck, or by content providers, the only data that this regime would provide is the single unverified assertion that the user was 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 kids-ID than with the adult-ID system. The 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

---

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

[64] *Reno* indicates quite clearly, we believe, that the state's interest is limited to facilitating the choice by parents. 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. But the Court did not embrace this broader restriction. *Reno* at 36.

quite severe.

- Seventh, this technique would easily generalize to other kid-protective regulations. In the very same act enacting COPA, Congress enacted the Child Online Privacy Protection Act. 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 got access.

These reasons together suggest why this alternative to COPA would be less burdensome than COPA. That's one half of the Court's test. The other half requires that the alternative be "at least as effective."[65] Here again, we believe that it would.

- First, if the relevant 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 there is no reason this profile would have to be difficult to set. Indeed, it would be easier to establish this profile than to establish a profile to collect email through a browser.
- 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 it could disable the ability to build alternative profiles beyond those set. These locations then would be protected locally, while under COPA, they are protected only if the kid doesn't 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 adult check requirement. The simplest way for a child to evade COPA is to steal a credit card number. It is certainly as easy to do that as it is to crack a security provision built into a

---

[65] Reno v. American Civil Liberties Union, 117 S. Ct. 2329, 2346 (1997).

browser.[66]

These considerations suggest that the effectiveness of this alternative is at least as good as the effectiveness of 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 US web sites would seem plainly ineffective. But between the two, we believe the kids-mode browser dominates COPA, and would therefore render COPA unconstitutional.

SECONDARY CONSEQUENCES

As we said at the start, our aim in this essay is not COPA alone. Our aim is any attempt to regulate access that does not fully consider the costs. In this section, we extend our analysis of these costs, to a 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 — at least when viewed from the perspective of the values of the internet.

In our analysis so far, we have considered three techniques for regulating access — tagging the sender, tagging the recipient, 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.

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.

---

[66] 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 is could be addressed. Again, the browser manufacturers could easily segregate download locations, based on whether the browser making the request was kids-ID enabled. If it was kids-ID enabled, then the company would download a kids-ID set browser only. Alternatively, the kids-IC could be enabled in the operating system, making substitution of an adult OS for a kids OS significantly more difficult.

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. These architectures, that is, might facilitate other jurisdiction-based regulation or access control imposed on senders, beyond the narrow purposes that motivated the initial change. We might, that is, 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 citizens, both while the citizens are at home, and while they are away. A certificate rich Internet could facilitate the identification of who could be regulated by whom, or what standards could be imposed on whom. And this, in turn, could facilitate a more general regulation of behavior in cyberspace.

We might imagine the scheme that looks something like this: States would enter a compact, whereby they, as a home jurisdiction, agree to require senders, or intermediaries, within their own jurisdiction, to respect the rules of other jurisdictions, in exchange for senders, or intermediaries in other jurisdictions doing the same for the home jurisdiction. These rules would specify the restrictions imposed on citizens from a given jurisdiction, and the range of citizens for whom the restriction applies. For example, a jurisdiction might specify that citizens from it 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 upon 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 which 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. 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, a form of discrimination is being enabled by the certificate infrastructure. The discriminations are then a consequence of this certificate infrastructure.

This change in regulability, however, is not the same for every ID architecture. Obviously, the more data a certificate architecture transmits, the more regulability increases. Likewise, an ID architecture like the kids-ID would facilitate very little regulation beyond regulation protecting kids. This is a second, and we believe, compelling reason to prefer it over the adult-ID regulation, for if the government has two means available for protecting kids, we believe it should select the means that have the least significant secondary effects.

Labels and Improper Control

The other general solution that 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 pre-clearance lists. But as should be obvious, an inexpensive and widely used labeling infrastructure would have its own secondary impacts, including both the possibility of more widespread speech regulation and 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.[67]

Second, labels might be used for voluntary access controls as well as mandated access controls.[68] 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

---

[67] 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, however; and our view is that this spill over ought to be reckoned in any regulatory regime.

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

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.[69]

The availability of voluntary access controls by 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 as well may be unproblematic.

But in practice, there are a number of reasons why 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.[70]

· 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.)[71]

· 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).[72] If everyone can 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

---

[69] For a demonstration of Alta Vista's "Family Filter", using ratings from SurfWatch, *click on the AV Family Filter link at http://www.altavista.com/*. For a discussion of the implications, *see* Jonathan Weinberg, *Rating the Net*, 19 HASTINGS COMM. & ENT. L. J. 453 n.108 (1997).

[70] *See* Rikki McGinty, *Safety Online: Will It Impede Free Speech?* Media Daily, December 5, 1997.

[71] *See* Weinberg, *supra* note 70, at n. 108.

[72] *See* CASS R. SUNSTEIN, DEMOCRACY AND THE PROBLEM OF FREE SPEECH (1995).

exclusive in form or in fact. As Niva Elkin-Koren describes,[73] one great virtue of the internet is its democratization of the process for drawing categories. 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.[74]

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."[75] 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.[76]

But whether one supports labels for these secondary uses or not, we both acknowledge that the consequence of these labels is to enable this secondary use. And if one were sufficiently troubled by this secondary use — as for example the ACLU and other civil rights organizations are[77] — then this secondary consequence might well affect one's judgment about whether a law mandating kids-IDs was preferable to a world with private labels. Or in other words: If part of the motivation for private labels comes from the

---

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

[74] *See* Weinberg, *supra* note 70, at 477.

[75] *See* Lawrence Lessig, *Tyranny in the Infrastructure*, Wired, July, 1997, at 96.

[76] *See* Paul Resnick, *Filtering Information on the Internet*, SCIENTIFIC AMERICAN 62 (March 1997) and *PICS, Censorship, & Intellectual Freedom FAQ*, (Paul Resnick, ed.) (last modified January 26, 1998) <http://www.w3.org/PICS/PICS-FAQ-980126.html>.

[77] Ann Beeson & Chris Hansen, Fahrenheit 451.2: Is Cyberspace Burning? How Rating and Blocking Proposals May Torch Free Speech on the Internet http://www.aclu.org./ issues/cyber/burning.html.

need to protect kids, that motivation would be undermined if there were other ways to protect kids.

CONCLUSION

This article has proposed an abstract model of mandated access controls, and it has applied that model to one concrete case. 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 pre-clearance 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 unless permitted regime). Otherwise, there will be no natural incentive, and the government will have to require the provision of that information.

---Table 1 about here---

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 voluntarily to exclude 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 are to be responsible 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, are the most costly place to impose responsibility. On the other hand, they are the most easily regulated, since there are fewer of them, they are more stable, 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 entail both direct costs and would create secondary by-products whose value is debatable. Given that the costs of any such architectural change would be significant, it is important for governments to answer the fundamental question of how important such changes are: perhaps a lessening of governments' traditional power to control the distribution of harmful information would be preferable.

SENSITIVITY TABLE (TABLE 1)

	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 · pre-clearance, thesauri	as for sender and recipient, plus: · responsibility to assure sender /recipient compliance · use of proxies and application gateways	· pre-clearance · sender's self-rating · third-party rating
Consequences	Enables general regulability of behavior on the net based on recipient type and jurisdiction	IAPs as the state	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 ISPs 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