
2014

An Information Theory of Copyright Law

Jeanne C. Fromer

Follow this and additional works at: <https://scholarlycommons.law.emory.edu/elj>

Recommended Citation

Jeanne C. Fromer, *An Information Theory of Copyright Law*, 64 Emory L. Rev. 71 (2014).
Available at: <https://scholarlycommons.law.emory.edu/elj/vol64/iss1/2>

This Article is brought to you for free and open access by Emory Law Scholarly Commons. It has been accepted for inclusion in Emory Law Journal by an authorized editor of Emory Law Scholarly Commons. For more information, please contact law-scholarly-commons@emory.edu.

AN INFORMATION THEORY OF COPYRIGHT LAW[†]

Jeanne C. Fromer^{*}

ABSTRACT

The dominant American theory of copyright law is utilitarian, in offering the incentive of limited copyright protection to creators to generate material that is valuable to society. Less settled is the question of the sorts of works that copyright law seeks to encourage: Ever more copyrightable creations? Only some that are artistically worthy? What makes a work valuable to society? This Article seeks to answer important aspects of these questions by examining them through the lens of information theory, a branch of applied mathematics that quantifies information and suggests optimal ways to transmit it. Using these concepts, this Article proposes that what makes expressive works valuable to society is that they make a contribution in at least one of two principal ways: by using that expression to communicate knowledge—be it systematic, factual, or cultural—and by conveying expression that is enjoyable in and of itself. Information theory sheds light on how copyright law can spur these valuable works. In undertaking this analysis, this Article explores the implications for the central doctrines of copyright law, including copyrightability, the idea-expression distinction, infringement, and fair use. In this context, this Article also considers whether we want distinct creators communicating these valuable types of information or whether it is optimal to unify particular communications of information in a single creator.

[†] Technically, the title of this Article ought to be “An Information-Theory Theory of Copyright Law.” However, one “Theory” was dropped to compress the message.

^{*} Professor of Law, New York University School of Law; Co-Director, Engelberg Center on Innovation Law & Policy. For their informative comments, I am grateful to Amy Adler, Arnaud Ajdler, Barton Beebe, Oren Bracha, Christopher Buccafusco, Irene Calboli, Sujit Choudhry, Charles Colman, Adam Cox, Samuel Estreicher, Brett Frischmann, James Grimmelmann, Laura Heymann, Kevin Hickey, Daniel Hulsebosch, Roberta Kwall, Mark McKenna, Michael Levine, Jake Linford, Jonathan Masur, Allison Orr Larsen, Daphna Renan, Betsy Rosenblatt, Zahr Said, Catherine Sharkey, Christopher Sprigman, James Stern, Alan Sykes, Rebecca Tushnet, Jeremy Waldron, Stephen Yelderman, Christopher Yoo, and Kenji Yoshino, and participants in workshops at the New York University School of Law, the William and Mary Law School, the Thirteenth Annual Intellectual Property Scholars Conference, and the 2014 Works in Progress in Intellectual Property Conference. I thank Christoffer Stromstedt for excellent research assistance. I also gratefully acknowledge support from the Filomen D’Agostino and Max E. Greenberg Research Fund.

INTRODUCTION	73
I. COPYRIGHT LAW	74
II. INFORMATION THEORY	76
III. AN INFORMATION THEORY OF COPYRIGHT LAW	83
A. <i>Noisy Expression Containing Knowledge</i>	85
B. <i>Expression Generating a Noisy Discussion</i>	92
IV. APPLICATIONS IN COPYRIGHT LAW	97
A. <i>Idea-Expression Distinction</i>	97
B. <i>Originality and Independent Creation</i>	102
1. <i>Independent Creation</i>	103
2. <i>Modicum of Creativity</i>	105
C. <i>Copyright Infringement</i>	107
1. <i>Should Any Copying Be Infringement</i>	108
2. <i>Substantial Similarity and Derivative Works</i>	109
D. <i>Fair Use</i>	113
E. <i>Dissemination</i>	117
1. <i>First Sale Doctrine</i>	118
2. <i>Secondary Liability</i>	119
3. <i>Search Engines</i>	121
F. <i>Preservation</i>	123
1. <i>Works in Multiple Formats</i>	124
2. <i>Deposit</i>	126
G. <i>Error Correcting Codes</i>	127
CONCLUSION	127

INTRODUCTION

The dominant American theory of copyright law is utilitarian, in offering the incentive of limited copyright protection to creators to generate material that is valuable to society. Less settled is the question of the sorts of works that copyright law seeks to encourage: What makes a work valuable to society? Does copyright law just want ever more expressive creations? Does it want to encourage only some that are artistically worthy? Does it want to stimulate as many works as possible without regard to these considerations? Or does it want to encourage something else? This Article seeks to answer important aspects of these questions by examining them through the lens of information theory, a branch of applied mathematics that quantifies information and suggests optimal ways to transmit it. Using these concepts, this Article proposes that what makes expressive works valuable to society is that they make a contribution in at least one of two principal ways: by using that expression to communicate knowledge—be it systematic, factual, or cultural—and by conveying expression that is enjoyable in and of itself. Information theory sheds light on how copyright law can spur these valuable works.

Works that are valuable due to the knowledge they contain are typically noisily expressive. Accordingly, many of the central doctrines of copyright law promote encoding this knowledge in redundant forms so as to transcend this noise and accomplish copyright law's goals of transmitting this knowledge. In that context, this Article considers whether rights to create these redundant forms ought to lie with the first author or more broadly with the public.

For works that are valuable due to their contributions to expression, it is clearly important that the expression itself be disseminated. Frequently underpinning this sort of expression is mystery, or at least varying views, as to the expression's meaning. Different audiences—varied in context or over time—might locate miscellaneous understandings in this expression. These diverse and possibly evolving readings are intimately tied to the value of the initial expression itself. As such, these readings are valuable for society as well. In that vein, copyright law ought to encourage both this valuable expression and the ability to work it over, repurpose it, and interpret it. By doing so, copyright law can encourage that the underlying expression promote a continuing conversation. It can do so by allowing the underlying expression to be used redundantly—to attach to varying meanings and interpretations—thereby creating a unified conversation.

Part I provides an overview of copyright law and theory. Part II introduces information theory. Part III then weaves information theory into copyright law. Part IV works through a number of central areas of copyright law that can (and often already do) accommodate an information theory of copyright law.

I. COPYRIGHT LAW

American copyright law protects “original works of authorship fixed in any tangible medium of expression,” including literary works, sound recordings, movies, and computer software code.¹ To obtain copyright protection, authors need only create a qualifying work.² Protection vests in authors without any formalities like registration.³ A copyright holder receives the exclusive right to reproduce the work, distribute copies of it, and prepare derivative works, among other things,⁴ typically until seventy years after the author’s death.⁵ Copyright protection extends to the expression of particular ideas rather than to the ideas themselves.⁶ Yet protection actually reaches well beyond the literal work to works that are copied and substantially similar,⁷ “else a plagiarist would escape by immaterial variations.”⁸

According to the Supreme Court, Congress, and many legal scholars, utilitarianism is the dominant purpose of American copyright law.⁹ According to utilitarian theory, copyright law provides the incentive of exclusive rights for a limited duration to authors to motivate them to create culturally valuable works.¹⁰ Without this incentive, the theory goes, authors might not invest the time, energy, and money necessary to create these works because they might

¹ 17 U.S.C. § 102(a) (2012); *Apple Computer, Inc. v. Franklin Computer Corp.*, 714 F.2d 1240 (3d Cir. 1983) (clarifying that computer software code is a literary work under copyright law).

² *See* 17 U.S.C. § 102(a) (requiring only that a work be fixed in “any tangible medium of expression” to be copyrightable). There is no requirement that a work be published to be protected.

³ Registration of a protected work with the Copyright Office is permissive. *Id.* § 408(a). To bring an infringement action, though, a copyright holder must in the ordinary case first have registered the copyright with the Copyright Office. *Id.* § 411(a).

⁴ *Id.* § 106.

⁵ *Id.* § 302(a).

⁶ *See id.* § 102(b); *Nichols v. Universal Pictures Corp.*, 45 F.2d 119, 121 (2d Cir. 1930).

⁷ *Corwin v. Walt Disney Co.*, 475 F.3d 1239, 1253 (11th Cir. 2007); *Whitehead v. Paramount Pictures Corp.*, 53 F. Supp. 2d 38, 45–46 (D.D.C. 1999).

⁸ *Nichols*, 45 F.2d at 121.

⁹ *See Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 558 (1985); 122 CONG. REC. 2834–35 (1976) (statement of Sen. McClellan); Shyamkrishna Balganes, *Foreseeability and Copyright Incentives*, 122 HARV. L. REV. 1569, 1576–77 (2009); William M. Landes & Richard A. Posner, *An Economic Analysis of Copyright Law*, 18 J. LEGAL STUD. 325, 326 (1989).

¹⁰ Stewart E. Sterk, *Rhetoric and Reality in Copyright Law*, 94 MICH. L. REV. 1197, 1197 (1996).

be copied cheaply and easily by free riders, eliminating authors' ability to profit from their works.¹¹

Utilitarianism aligns fluently with (and is frequently justified by) the U.S. Constitution's grant of power to Congress "[t]o promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries."¹² Most utilitarians understand social welfare to be maximized—in the context of copyright law—by the creation of ever more artistic works.¹³ For example, Paul Goldstein posits without further explanation that "[t]he aim of copyright law is to direct investment toward the production of abundant information" and "the widest possible variety of literary, musical and artistic expression."¹⁴ In a previous work, I suggest that, given what society values as artistic creativity, copyright law seeks to encourage works in which "an author[']s identification of] subjective emotional themes or ideas [have been] transform[ed] into artistic expression."¹⁵ Both of these explanations align with copyright law's low threshold for protectability.¹⁶ Others, like William Fisher, employ a broader understanding of welfare: that intellectual property protection ought "to help foster the achievement of a just and attractive culture."¹⁷ Which sorts of works a utilitarian copyright law ought to encourage remains an unsettled issue, although many of these views overlap with one another.

Pursuant to utilitarianism, the rights conferred by copyright laws are designed to be limited in time and scope.¹⁸ The reason for providing copyright protection to creators is to encourage them to produce socially valuable works, thereby maximizing social welfare.¹⁹ If the provided rights are exceedingly

¹¹ See Alina Ng, *The Author's Rights in Literary and Artistic Works*, 9 J. MARSHALL REV. INTELL. PROP. L. 453, 453 (2009); Symposium, *The Constitutionality of Copyright Term Extension: How Long Is Too Long?*, 18 CARDOZO ARTS & ENT. L.J. 651, 676 (2000) (statement of Wendy Gordon).

¹² U.S. CONST. art. I, § 8, cl. 8.

¹³ See 1 PAUL GOLDSTEIN, *COPYRIGHT: PRINCIPLES, LAW AND PRACTICE* § 2.2.1, at 63–64 (1989).

¹⁴ *Id.*

¹⁵ Jeanne C. Fromer, *A Psychology of Intellectual Property*, 104 NW. U. L. REV. 1441, 1492 (2010).

¹⁶ See GOLDSTEIN, *supra* note 13, at § 2.2.1, at 64; Fromer, *supra* note 15, at 1492–1501.

¹⁷ William Fisher, *Theories of Intellectual Property*, in *NEW ESSAYS IN THE LEGAL AND POLITICAL THEORY OF PROPERTY* 168, 172 (Stephen R. Munzer ed., 2001); accord MADHAVI SUNDER, *FROM GOODS TO A GOOD LIFE: INTELLECTUAL PROPERTY AND GLOBAL JUSTICE* 8–11 (2012) (arguing that participation, livelihood, and shared meaning ought to play a role in shaping intellectual property laws).

¹⁸ See Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 TEX. L. REV. 989, 997 (1997).

¹⁹ Ralph S. Brown, *Eligibility for Copyright Protection: A Search for Principled Standards*, 70 MINN. L. REV. 579, 592–96 (1985).

extensive, society would be hurt (and social welfare diminished).²⁰ For one thing, exclusive rights in intellectual property can prevent competition in protected works, thereby allowing the rightsholder to charge a premium for access and ultimately limiting these valuable works' diffusion to society at large.²¹ For another, given that knowledge is frequently cumulative, society benefits when subsequent creators are not prevented from building on previous artistic creations to generate new works.²² For these reasons, copyright law ensures both that the works it protects fall into the public domain in due course and that third parties are free to use protected works for certain socially valuable purposes.²³

At bottom, a utilitarian theory of copyright law rests on the premise that the benefit to society of creators crafting valuable works offsets the costs to society of the incentives the law offers to creators.²⁴

With this overview of copyright law and theory, I now turn to a discussion of information theory.

II. INFORMATION THEORY

In 1948, Claude Shannon, a mathematician and electrical engineer working at AT&T Bell Laboratories, published a foundational paper on a mathematical

²⁰ See Lemley, *supra* note 18, at 996–97.

²¹ See *id.*

²² See *id.* at 997–98.

²³ See *id.* at 999.

²⁴ *Id.* at 996–97. Despite the dominance of utilitarian thinking in American copyright law, scholars also proffer other theories to justify copyright protection. These theories are typically grounded in the notion of natural or moral rights that authors and inventors deserve by virtue of having created their works. See, e.g., Balganes, *supra* note 9, at 1576–77; Brown, *supra* note 19, at 589–90. Moral-rights theories typically come in two flavors: labor-desert and personhood. For more on the labor-desert theory, see ROBERT P. MERGES, JUSTIFYING INTELLECTUAL PROPERTY 31–67 (2011); Wendy J. Gordon, *A Property Right in Self-Expression: Equality and Individualism in the Natural Law of Intellectual Property*, 102 YALE L.J. 1533, 1540–83 (1993); Justin Hughes, *The Philosophy of Intellectual Property*, 77 GEO. L.J. 287, 296–330 (1988). For more on personhood theory, see Lawrence C. Becker, *Deserving To Own Intellectual Property*, 68 CHI.-KENT L. REV. 609 (1993); Hughes, *supra*, at 330–65; Margaret Jane Radin, *Property and Personhood*, 34 STAN. L. REV. 957 (1982). I argue in a prior work that these theories can form the basis of expressive incentives for creators in a utilitarian system. See Jeanne C. Fromer, *Expressive Incentives in Intellectual Property*, 98 VA. L. REV. 1745 (2012). That said, in this Article, I focus not on moral rights but on the utilitarian justification for copyright law, owing to its dominance and constitutional basis. Nonetheless, it might be possible to view an information theory of copyright law as described below, see *infra* Parts III–IV, as grounded in a form of moral rights for the audience of existing works.

theory of information and communications.²⁵ He sought to explain how much information could be sent per second over communication channels, and how best to encode messages to transmit them over these channels.²⁶ Shannon's research proved critical for the dawning information age, in which all sorts of messages would be sent over a variety of channels, like computer and telephone networks, and with the goal that transmission be efficient and robust.²⁷ Information theory also found applications in diverse fields, such as finance and physics.²⁸ Most relevant for copyright law, Shannon's information theory covers all sorts of communications, such as written and spoken language, musical sounds, and films.²⁹

In its most basic form, information is encoded and communicated in a series of bits, a bit being a binary digit that can have only one of two values (zero or one).³⁰ Information theory posits a message source (such as an author), a variety of possible messages that the source can produce at any given time, a channel to transmit this and other messages, and a recipient of this message.³¹ To take one concrete example, an English sentence might be the message that will be sent. The author of this sentence might encode it for transmission as a series of bits in a number of different ways. It might be written up by breaking down the sentence into its alphabetical characters, so that each letter is represented by a different series of bits (such as the character "a" being "00000," "b" being "00001," "c" being "00010," "d" being "00011," and so

²⁵ Claude E. Shannon, *The Mathematical Theory of Communication* (versions 1 & 2), 27 BELL SYS. TECHNICAL J. 379, 623 (1948), reprinted in CLAUDE E. SHANNON & WARREN WEAVER, *THE MATHEMATICAL THEORY OF COMMUNICATION* 29 (10th prtg. 1964) (reprinting this paper with minor alterations from original). See generally Graham P. Collins, *Claude E. Shannon: Founder of Information Theory*, SCI. AM. (Oct. 14, 2002), <http://www.scientificamerican.com/article/claude-e-shannon-founder/> (providing biographical information regarding Shannon).

²⁶ See THOMAS M. COVER & JOY A. THOMAS, *ELEMENTS OF INFORMATION THEORY* 1 (2d ed. 2006). See generally Warren Weaver, *Recent Contributions to the Mathematical Theory of Communication*, in SHANNON & WEAVER, *supra* note 25, at 1 (providing overview of leading communication theory works). For an accessible explanation of information theory, see generally JAMES GLEICK, *THE INFORMATION: A HISTORY, A THEORY, A FLOOD* (2011).

²⁷ See generally COVER & THOMAS, *supra* note 26, at 1, 103–241, 427–611 (describing applications of Shannon's information theory to computer science and network theory).

²⁸ See generally *id.* at 409–25, 613–56 (discussing application of information theory to entropy and investment portfolios).

²⁹ See JOHN R. PIERCE, *AN INTRODUCTION TO INFORMATION THEORY: SYMBOLS, SIGNALS & NOISE* 8–9, 64 (2d rev. ed. 1980). But cf. *id.* at 107–24 (debating whether the English language can be studied through the lens of information theory). Shannon's work and its progeny apply to such an array of contexts because information theory addresses "the accurate representation of one thing by another," an issue central to many endeavors. *Id.* at 64–65.

³⁰ See *id.* at 80.

³¹ See *id.* at 23.

forth to encode all twenty-six letters). As another of many possibilities, the sentence might be encoded by breaking down the sentence into its words, so that each word is represented by a different series of bits (for example, “the” being “00000000,” “person” being “00000001,” “information” being “00000010,” and so on to encode all known English words). In the first example, the possible symbols are English alphabet letters. In the second, they are known English words. As might be apparent, each symbol in the second example will be longer than in the first: There are only twenty-six total symbols (English letters) to encode in the first example, but there are at least 470,000 symbols (English words) to encode in the second example.³² Therefore, more bits must be allocated per symbol in the second example than in the first to allow for all of the possibilities.

As one information theorist explains, “how much information it is possible to send over a [channel] depends not only on how fast one can send successive symbols . . . over the [channel] but also on how many different symbols . . . one has available to choose among.”³³ To an information theorist, “information is a measure of one’s freedom of choice when one selects a message.”³⁴ Information theory measures information based on the number of symbols in a message multiplied by the number of bits needed to encode each symbol, or the total number of bits needed to encode a message.³⁵ The average number of bits per symbol needed to encode it is called the “entropy” rate.³⁶ For example,

³² See *How Many Words Are There in English?*, MERRIAM-WEBSTER, http://www.merriam-webster.com/help/faq/total_words.htm (last visited Sept. 2, 2014) (“*Webster’s Third New International Dictionary*, Unabridged, together with its 1993 Addenda Section, includes some 470,000 entries. *The Oxford English Dictionary*, Second Edition, reports that it includes a similar number.”). If all 470,000 words were to be encoded using the same number of bits, nineteen bits would be needed to encode all of them (because $\log_2 470,000 = 18.84$), whereas only five are needed to encode the twenty-six letters of the alphabet (because $\log_2 26 = 4.7$).

³³ PIERCE, *supra* note 29, at 28.

³⁴ Weaver, *supra* note 26, at 9.

³⁵ See PIERCE, *supra* note 29, at 39–40 (“ H , the information of the message, [i]s the logarithm of the number of possible sequences of symbols which might have been selected and showed that $H = n \log s$. Here n is the number of symbols selected, and s is the number of different symbols in the set from which symbols are selected.”); see also ROBERT ASH, INFORMATION THEORY 1–26 (1965). This calculation rests on the assumptions that successive symbols are chosen independently from one another and symbols are equally likely to be chosen. PIERCE, *supra* note 29, at 40. This assumption is not true of English prose, for example, because there are differing statistical likelihoods that certain words or letters follow the preceding ones. See *id.* at 61–62.

³⁶ PIERCE, *supra* note 29, at 80–81; see also COVER & THOMAS, *supra* note 26, at 13–55. Although it might not seem intuitive, there is a connection between entropy in this context and in physics. As Pierce noted:

[S]uppose we regard a source of information as to what state a system is in as a message source. The information-theory entropy of this source is a measure of the energy needed to transmit a

the entropy of a single toss of a (fair) coin is one bit, whereas two tosses have an entropy of two bits.

In the examples given above as to an English sentence, assuming transmission speeds are constant, the greater number of bits per symbol indicates that it would take longer to transmit the word symbols than the letter symbols. That is not to say, however, that the total information in the second message would take longer to transmit than in the first: It might be otherwise if there are fewer symbols in the second message (even with a greater number of bits per symbol).³⁷ In fact, Shannon and other information theorists showed that a word-symbol encoding of the English language would occupy less information than a letter-symbol encoding.³⁸

The examples given until now suggest that each symbol in a message will have the same number of bits. This need not be. Indeed, Shannon worked out ways to encode messages efficiently, and others—particularly David Huffman—developed even more efficient encodings with message symbols of varying lengths, which rely on knowing the possible symbols and the probability that each would be chosen.³⁹ In fact, one of the goals of information theory is to construct encodings that are as efficient, or minimal, as possible, so as to communicate as efficiently as possible.⁴⁰ As information theorists Thomas Cover and Joy Thomas explain, “[d]ata compression can be

message from the source in the presence of the thermal noise which is necessarily present in the system. The energy used in transmitting such a message is as great as the increase in free energy due to the reduction in physical entropy which the message brings about.

PIERCE, *supra* note 29, at 207. A more general way to see entropy is based on the probability of any symbol being chosen combined with the amount of information needed to encode it. *See id.* at 81–88. For example, using only the 8,727 most frequently used English words, Claude Shannon found the entropy to be 9.14 bits per word. *Id.* at 87–88. In this sense, “the number of binary digits required to transmit a message is just the entropy in bits per symbol times the number of symbols.” *Id.* at 90. For more mathematical equations for entropy, see *id.* at 90–94. Andrey Kolmogorov and others set out an alternative notion to measure information: “that the complexity of a string of data can be defined by the length of the shortest binary computer program for computing the string.” COVER & THOMAS, *supra* note 26, at 3 (discussing Kolmogorov complexity, and stating that it “is approximately equal to the Shannon entropy H if the sequence is drawn at random from a distribution that has entropy H ”).

³⁷ Mathematically speaking, “if we send symbols (successive current values) at a constant rate, the speed of transmission, W , is related to m , the number of different symbols or current values available, by $W = K \log m$,” wherein K is “a constant whose value depends on how many successive current values are sent each second.” PIERCE, *supra* note 29, at 36.

³⁸ *See id.* at 74–75 (assuming that 16,384 words are encoded using 14 bits that “we will use on the average 27.5 binary digits per word [average 5.5 characters per word], while in encoding the message word by word we need only 14 binary digits per word”).

³⁹ *See* COVER & THOMAS, *supra* note 26, at 103–58; PIERCE, *supra* note 29, at 94–95.

⁴⁰ PIERCE, *supra* note 29, at 127.

achieved by assigning short descriptions to the most frequent outcomes of the data source, and necessarily longer descriptions to the less frequent outcomes.”⁴¹

Encoding and decoding messages becomes yet more complicated when the communication channel over which messages are being conveyed is noisy, in that the channel corrupts, loses, or otherwise changes the message being sent. In that situation, the recipient can receive a message that is different from the one the sender passed along, as when there are crackles or inaudible transmissions over a telephone, radio, or computer, or a picture overlaid with “snow” on a television set.⁴²

One costly example of noise in message transmission occurred in 1887—and eventually made its way to litigation at the U.S. Supreme Court⁴³—when Frank Primrose, a Pennsylvania wool dealer, sent a telegraphic message to his agent in Kansas.⁴⁴ Primrose transmitted to his agent a message that he had bought 500,000 pounds of wool using the code word “bay,” a term which they had agreed had this meaning.⁴⁵ Somehow, his agent received a message to “buy” 500,000 pounds of wool, which the agent then proceeded to try to do, erroneously, purchasing almost 300,000 pounds of wool at a cost of \$20,000.⁴⁶ The difference between the correct and erroneous message was an “a” and “u,” or in Morse code, a difference of one dot.⁴⁷ Angered, Primrose sued the telegraph company.⁴⁸

Generally speaking, noise in transmission can result from malfunctions in the message transmitter or from interference caused by noise or signals external to the transmitter.⁴⁹ One might try to reduce noise by inventing yet better transmitters.⁵⁰ However, Claude Shannon anticipated—correctly—that better transmitters would not eliminate the problem of noise, both internal and

⁴¹ COVER & THOMAS, *supra* note 26, at 103.

⁴² PIERCE, *supra* note 29, at 145–46.

⁴³ *Primrose v. W. Union Tel. Co.*, 154 U.S. 1, 1 (1894).

⁴⁴ This story is also retold in GLEICK, *supra* note 26, at 158.

⁴⁵ *Primrose*, 154 U.S. at 4.

⁴⁶ *Id.* at 6.

⁴⁷ *Id.* at 5.

⁴⁸ The Supreme Court allowed recovery for Primrose only in the amount of \$1.15, the cost of sending the telegraph, due to a contractual waiver of liability to which Primrose had agreed when he had sent the telegraph (in addition to a discounted contractual price to repeat a telegraphed message to reduce the possibility of noisy transmission). *Id.* at 4, 34.

⁴⁹ See PIERCE, *supra* note 29, at 147.

⁵⁰ See *id.* at 146.

external to the transmitter.⁵¹ Operating under an assumption of noise, Shannon nonetheless showed that it is possible to encode messages and transmit them accurately and relatively efficiently.⁵² That is, Shannon had the insight that one can treat noise by structuring the message sent to detect and alleviate noise in the received message.⁵³

The key to noise detection and correction by a message recipient is redundancy in a message.⁵⁴ Redundancy in this context means that there is some “fraction of the message [that] is unnecessary (and hence repetitive or redundant) in the sense that if it were missing the message would still be essentially complete, or at least could be completed.”⁵⁵ That is, if the sent message contains usefully redundant information, the recipient ought to be able to use the redundancies to detect errors in transmission and, better yet, fix them. For example, if a sender seeks to transmit the result of a series of ten coin tosses of heads and tails, in which the bit “0” represents heads and the bit “1” represents tails, the sender’s message might be “0010111101.” However, noise on the communications channel might cause the recipient to receive the incorrect message of “0000111101.” The sender might instead relay each coin toss twice in a row, with the message now containing twenty bits, so that the recipient can detect an erroneous noisy transmission on the basis that a pair of symbols would not match (and the assumption that there is an error in only one of that pair). In this example, the sender’s message, doubled up, would be “00001100111111110011,” and the recipient would receive the noisy message “00000100111111110011.” Because the fifth and sixth bit received would not match, the recipient would know that an error occurred in that segment of the message (but that an error was unlikely in the other matched segments). The recipient, however, would not know how to figure out whether the third coin toss was actually heads or tails. Yet more redundancy would accomplish that goal. Were the sender to encode each coin toss in triplicate rather than in duplicate, the recipient could make the fair assumption that when the three bits allotted to a particular coin toss were not identical, the one that occurred twice in the received message is the right one.⁵⁶

⁵¹ See *id.* at 42, 146.

⁵² See *id.*

⁵³ See *id.* at 146. See generally Shannon, *supra* note 25, at 65–80 (discussing channel capacity and noise).

⁵⁴ See ASH, *supra* note 35, at 27–45, 87–168 (providing an overview of noiseless coding and error correcting code theory).

⁵⁵ Weaver, *supra* note 26, at 13.

⁵⁶ See PIERCE, *supra* note 29, at 149–50.

Relatedly, returning to the example of the English language, Shannon estimated that it has a redundancy of about fifty percent to alleviate noise.⁵⁷ For instance, eliminating a certain half of the letters from an English sentence has no effect on its comprehensibility: “If u cn rd ths”⁵⁸

It ought to be intuitive that efficient encodings are antithetical to redundancy. Errors can be particularly hard to detect and correct in efficient encodings. Efficient encodings combined with noise can thus lead to gross misunderstandings of a sent message.⁵⁹ And conversely, redundancy, particularly of the sort exemplified above, can be costly by decreasing the rate of transmission significantly—to one-half or one-third in the examples above.⁶⁰

Shannon proved that there exists an encoding for transmission over noisy channels in which errors will be at most arbitrarily small, and the encoding would be significantly more efficient than in the examples given above.⁶¹ Later information theorists derived relatively efficient systems of redundancy—including check bits, block coding, and convolutional codes—that might be inserted into codes to check or correct errors in transmission.⁶² As one information theorist explains, “the whole problem of efficient and error-free communication turns out to be that of removing from messages the somewhat inefficient redundancy which they have and then adding redundancy of the right sort in order to allow correction of errors made in transmission.”⁶³

In addition to coming up with ways to reduce errors in communication resulting from noise in transmission, the message recipient must know how to decode the received message to understand its meaning. A sent message full of

⁵⁷ Shannon, *supra* note 25, at 25–26.

⁵⁸ GLEICK, *supra* note 26, at 217. Many forms of information similarly contain redundancies, such as DNA’s redundancy in its coding to tolerate errors and music’s “unvarying tempo, uniform timbre, just a brief melodic pattern, a word, repeated over and over with slight variations till the final bars.” *Id.* at 297, 352.

⁵⁹ See PIERCE, *supra* note 29, at 149.

⁶⁰ See *id.* at 150.

⁶¹ As Pierce explained:

Let a discrete channel have a capacity C and a discrete source the entropy per second H . If $H < C$ there exists a coding system such that the output of the source can be transmitted over the channel with an arbitrarily small frequency of errors (or an arbitrarily small equivocation). If $H > C$ it is possible to encode the source so that the equivocation is less than $H - C + \epsilon$, where ϵ is arbitrarily small. There is no method of encoding which gives an equivocation less than $H - C$.

See *id.* at 156; see also COVER & THOMAS, *supra* note 26, at 183–241.

⁶² See PIERCE, *supra* note 29, at 159–62.

⁶³ *Id.* at 164.

0's and 1's has no meaning unless the recipient knows whether to attribute, say, "heads" to "0" and "tails" to "1," or "black" to "0" and "white" to "1," or "a" to "00000," "b" to "00001," and so forth. The code might not be entirely clear. According to engineer John Pierce, "in language, the listings may overlap. And one person's code book may have different entries from another's, which is sure to cause confusion."⁶⁴ In these cases, there is semantic noise in a message.⁶⁵ Semantic information theory extends the fundamentals of information theory to include in their measure of information the recipient's uncertainty after receiving a message or the likelihood that the message is true.⁶⁶ Just as in information theory generally, semantically redundant encodings can help alleviate semantic noise.⁶⁷ Additionally, when a shared knowledge base is known, semantic compression of messages can be achieved.⁶⁸ More generally, the fundamental concepts of information theory are extendable to a semantic information theory.

III. AN INFORMATION THEORY OF COPYRIGHT LAW

This Part weaves information theory as just described into the utilitarian framework of copyright law. It seeks to address the critical question lingering in the dominant explanation of copyright law's goal: the production of an abundance of expressive works.⁶⁹ Copyright's low and precise standard of originality does just that.⁷⁰ But is it true that society needs an abundance of works? One might imagine that some smaller subset that is of sufficient quality, however "quality" is defined, could be preferable.⁷¹ Or even if one

⁶⁴ *Id.* at 118.

⁶⁵ See GLEICK, *supra* note 26, at 349 ("The sender of a message can never fully know his recipient's mental code book. Two lights in a window might mean nothing or might mean 'The British come by sea.' Every poem is a message, different for every reader.").

⁶⁶ See, e.g., GUY JUMARIE, *RELATIVE INFORMATION: THEORIES AND APPLICATIONS* 46 (1990); D.M. MacKay, *The Place of 'Meaning' in the Theory of Information*, in *INFORMATION THEORY: PAPERS READ AT A SYMPOSIUM ON 'INFORMATION THEORY' HELD AT THE ROYAL INSTITUTION, LONDON, SEPTEMBER 12TH TO 16TH 1955*, 215, 215–19 (Colin Cherry ed., 1956); RUDOLF CARNAP & YEHOSHUA BAR-HILLEL, *MASS. INST. OF TECH., RESEARCH LAB. OF ELEC., TECHNICAL REPORT NO. 247, AN OUTLINE OF A THEORY OF SEMANTIC INFORMATION* (1952), available at <http://www.survivor99.com/lcg/information/CARNAP-HILLEL.pdf>. Claude Shannon, however, purposely excluded semantics from his information theory. See Shannon, *supra* note 25, at 3 ("[S]emantic aspects of communication are irrelevant to the engineering problem.").

⁶⁷ See, e.g., Prithwish Basu et al., *Preserving Quality of Information by Using Semantic Relationships*, 11 *PERVASIVE & MOBILE COMPUTING* 188, 190, 192, 201 (2014).

⁶⁸ See, e.g., *id.*

⁶⁹ See *supra* text accompanying notes 14–15.

⁷⁰ See *supra* text accompanying note 16.

⁷¹ Cf. Ned Snow, *The Regressing Progress Clause: Rethinking Constitutional Indifference to Harmful Content in Copyright*, 47 *U.C. DAVIS L. REV.* 1 (2013) (proposing that we ought to discriminate based on

thinks that an abundance of works is desirable, is there a logical stopping point before infinity at which enough works are enough, thereby avoiding Jorge Luis Borges's infinite Library of Babel?⁷²

This Part answers the question of which set of works copyright law should seek to encourage. It indeed ought to be an abundance of works, but of particular types and of an optimal—and less than infinite—amount. I propose that copyright law seeks in large part to encourage two aspects of expression: (1) the creation, dissemination, and preservation of systematic, factual, and cultural knowledge, and (2) beautiful or otherwise memorable pieces of expression itself. These two valuable aspects of expression can be noisily communicated, even when the expression itself being disseminated is accurately transmitted to the public. With regard to creating, disseminating, and preserving knowledge, the works in which these forms of knowledge are encoded are typically noisily expressive. And with reference to encouraging momentous expression, the expression might be accurately transmitted, but the expression's meaning or purpose is often unknown. In fact, the expression will often possess multiple meanings or purposes simultaneously, oftentimes for varying audiences. For meaning to attach to valued expression, then, a noisy conversation must emerge to interpret, contextualize, and repurpose the expression.

As discussed in this Part, these forms of noise cannot or should not be extinguished. As such, in both contexts, to unlock the two valuable aspects of expression, copyright law ought to encourage helpful forms of redundancy. For one thing, encoding the knowledge copyright law seeks to encourage in redundant forms will help to transcend noisy expression and accomplish copyright law's goals of creating, disseminating, and preserving this knowledge. For another, encouraging expression that is valued in and of itself to be used and reused in multiple contexts will help provide important multiple meanings for this expression to different audiences in varied contexts. This

content in copyright law); Barton Beebe, *Bleistein*; or Copyright Law and the Problem of Aesthetic Progress (2014) (unpublished manuscript) (on file with author) (suggesting that copyright law ought to seek to promote aesthetic progress).

⁷² See Jorge Luis Borges, *The Library of Babel*, in COLLECTED FICTIONS 112 (Andrew Hurley trans., 1998); accord GLEICK, *supra* note 26, at 373 (describing Borges's library as one in which "no knowledge can be discovered there, precisely because all knowledge *is* there, shelved side by side with all falsehood" and represents "no more perfect case of information glut"). Given an increasing amount of user-generated content in the digital age, e.g., Edward Lee, *Warming Up to User-Generated Content*, 2008 U. ILL. L. REV. 1459, society is moving closer to an infinite library.

Part deals in turn with each form of value in expression and its concomitant noise.

A. *Noisy Expression Containing Knowledge*

In conferring power on Congress to enact copyright laws, the Constitution states the goal of such power: “[t]o promote the Progress of Science.”⁷³ This phrase generally refers to the objective of encouraging the advancement of systematic, factual, and cultural knowledge.⁷⁴ Breaking it down, according to the Supreme Court, “to promote” as used in the Intellectual Property Clause means “‘to stimulate,’ ‘to encourage,’ or ‘to induce.’”⁷⁵ Most commentators understand “progress” to mean advancement.⁷⁶ “Science” as it appears in the Clause did not originally have the meaning contemporary Americans associate with it—biology, chemistry, and the like. Instead, at the time of the Constitution’s framing, science meant knowledge or learning, particularly of the kind that is systematic and of enduring value.⁷⁷ Herein, in line with the generally accepted understanding, I use the term capaciously to include not just systematic and factual knowledge, but also knowledge that is cultural (and is frequently about fictional “facts,” such as a novel’s details).⁷⁸

⁷³ U.S. CONST., art. I, § 8, cl. 8 (conferring also on Congress power “[t]o promote the Progress of . . . useful Arts”). For a summary of other (minority) views on this phrase’s contribution to the meaning of the Intellectual Property Clause, see Jeanne C. Fromer, *The Intellectual Property Clause’s External Limitations*, 61 DUKE L.J. 1329, 1338–39 (2012).

⁷⁴ I explore the meaning and extent of the Intellectual Property Clause in greater depth in Fromer, *supra* note 73; see also Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 545 (1985) (“[C]opyright is intended to increase and not to impede the harvest of knowledge.”).

⁷⁵ Goldstein v. California, 412 U.S. 546, 555 (1973).

⁷⁶ E.g., Lawrence B. Solum, *Congress’s Power To Promote the Progress of Science: Eldred v. Ashcroft*, 36 LOY. L.A. L. REV. 1, 45 (2002); Edward C. Walterscheid, *Conforming the General Welfare Clause and the Intellectual Property Clause*, 13 HARV. J.L. & TECH. 87, 93 (1999); Edward C. Walterscheid, *The Preambular Argument: The Dubious Premise of Eldred v. Ashcroft*, 44 IDEA 331, 376 (2004). As Solum explains further, “progress” can be understood either as “advancement of learning [with a] focus on the results of scientific activity” or as “encouraging the activity itself [with a] focus on the process itself.” Solum, *supra*, at 45–46. The first understanding can further mean improvement in a knowledge base’s quality or quantity. Rejecting an approach like Professor Solum’s, Professor Malla Pollack concludes on the basis of some historical linguistic evidence that “progress” means spread, diffusion, or distribution. Malla Pollack, *What Is Congress Supposed to Promote?: Defining “Progress” in Article I, Section 8, Clause 8 of the United States Constitution, or Introducing the Progress Clause*, 80 NEB. L. REV. 754, 755–57, 794–95, 809 (2001).

⁷⁷ EDWARD C. WALTERSCHEID, *THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE: A STUDY IN HISTORICAL PERSPECTIVE* 125 & n.46 (2002); see also Solum, *supra* note 76, at 3 (“[T]he meaning of science that best coheres with the constitutional text and the original understanding can be glossed as systematic knowledge or learning of enduring value.”).

⁷⁸ This broad understanding fits with the precise wording of the Intellectual Property Clause, which does not talk merely about promoting science, but about promoting the progress of science, a broader goal dedicated to advancing knowledge. See *supra* note 76. But see Ned Snow, *The Meaning of Science in the Copyright*

The first national copyright law, enacted in 1790,⁷⁹ confirms this understanding. The law's purpose stated that it was "[a]n Act for the encouragement of learning."⁸⁰ Moreover, in line with this purpose, the law's coverage extended only to books, maps, and charts.⁸¹ All three types of subject matter categorically advance systematic and factual knowledge, while books also tend to advance cultural knowledge.

For example, protection for books that convey factual knowledge (such as encyclopedias or biographies) helps advance societal knowledge of the facts contained therein. Maps and charts similarly advance societal knowledge by conveying their depicted facts. Fiction books, too, can convey knowledge. For one thing, the fictional world created by the author is a second-order form of knowledge (knowledge about the fictional world crafted by the author), which might be valuable knowledge for society for further analysis and dissection. Additionally, fictional works typically convey various themes—be they emotional or intellectual—and readers will locate knowledge within them (even if their individual readings differ).⁸² Finally, books that analyze and comment on prior books similarly can help to highlight the knowledge that these commentators have located in the prior books.

The constitutional objective of "promot[ing] the Progress of Science" fits neatly within the larger scheme of utilitarianism that sits at the base of copyright law.⁸³ As it currently stands, copyright law seeks to encourage authors to produce "original works of authorship," works that are valuable to society.⁸⁴ A significant subset of valuable authored works is that which advances society's knowledge systematically, factually, and culturally.⁸⁵ To promote the advancement of knowledge, copyright law's utilitarian aims with regard to these works must be threefold: to promote these works' creation,

Clause, 2013 BYU L. REV. 259, 263 (contending that the original meaning of "Science" in the Clause is "a system of knowledge comprising distinct branches of study").

⁷⁹ Act of May 31, 1790, ch. 15, 1 Stat. 124.

⁸⁰ *Id.* The British copyright law on which the American one is based and previous state copyright laws had similar articulated goals. See generally BRUCE W. BUGBEE, GENESIS OF AMERICAN PATENT AND COPYRIGHT LAW (1967) (providing historical account of the origins of copyright and patent statutes).

⁸¹ 1 Stat. at 124.

⁸² See *infra* Part III.B.

⁸³ See *supra* Part I (discussing the theory of utilitarianism).

⁸⁴ 17 U.S.C. § 102 (2012); see *supra* Part I.

⁸⁵ Of course, copyright law might also seek to promote other types of valuable authored works, perhaps such as those that purely move people and promote aesthetic progress apart from contributions to knowledge broadly construed. See Beebe, *supra* note 71 (exploring whether copyright law should seek to promote aesthetic progress). I take up this category of works in the following section.

dissemination, and preservation.⁸⁶ The pursuit of these three goals together ensures that society can glean the important knowledge contained in these works and do so over a long period of time. Society can then enjoy these works, learn from them, and build on them in creating future valuable works.⁸⁷

Given that an important basis of copyright law, then, is to encourage the creation, dissemination, and preservation of valuable works, copyright law ought to be concerned with ensuring that that which makes those works valuable—knowledge, broadly construed—is contained in the works being encouraged. That is, copyright law ought to encourage works to contain systematic, factual, and cultural knowledge, which can then be disseminated and preserved for society's use.

As information theory suggests, the most effective way to encourage transmission of this knowledge is to encode it clearly and efficiently so that it might be properly decoded by recipients and to then ensure wide dissemination and preservation.⁸⁸ To communicate this knowledge clearly, information theory indicates that this knowledge ought to be encoded concisely and in a way that recipients would comprehend.⁸⁹ For an author seeking to communicate, say, the basic tenets of molecular biology, the most direct way to do that would be to list them as concisely and clearly as possible without extra verbiage. Or an author who wants to convey the loss that she feels upon the death of her father might list the emotions she feels directly and crisply. Or a creator who wants to film a movie about a group of five self-centered friends might have each of them directly communicate that they are self-centered. More generally, works could most effectively and efficiently communicate the knowledge their author seeks to convey by robotically listing this knowledge.

In this light, one would think that copyright law would encourage to-the-point authored works that resemble these examples, and concomitantly, that we would see many works of this mechanically concise type. However, copyright law does no such thing, as will be discussed in greater detail in the next Part, and we do not frequently see works of this type, particularly within the purview of copyright law.

⁸⁶ See *Eldred v. Ashcroft*, 537 U.S. 186, 244, 266 (2003) (Breyer, J., dissenting).

⁸⁷ See *supra* Part I.

⁸⁸ See *supra* Part II.

⁸⁹ See *supra* Part II.

There are two important reasons why not: First Amendment values and people's hedonic preference for expressive works. Consider First Amendment values. Although there are many theories underpinning the First Amendment's guarantee of freedom of speech,⁹⁰ arguably the most prominent—if only popularly—is Justice Holmes's view of a marketplace of ideas to encourage the truth to prevail:

[W]hen men have realized that time has upset many fighting faiths, they may come to believe even more than they believe the very foundations of their own conduct that the ultimate good desired is better reached by free trade in ideas—that the best test of truth is the power of the thought to get itself accepted in the competition of the market, and that truth is the only ground upon which their wishes safely can be carried out. That at any rate is the theory of our Constitution. It is an experiment, as all life is an experiment.⁹¹

This theory has retained a central role among First Amendment values (despite critiques that there might not be “truths” to discover or that free speech might not get us to these truths).⁹² The marketplace-of-ideas view suggests that ideas ought to be discussed and debated from multiple vantage points.⁹³ Most relevantly, it suggests that the core nuggets building up to a viewpoint—the individual facts and the basic concepts that are woven together—ought to be left freely available to any and all to use in their speech.⁹⁴ Therefore, and as explained in greater detail below, copyright law does not allow the proprietization of these basic components, and we tend not to see such to-the-point authored works, as they are unprotectable.⁹⁵

Additionally, people have a hedonic preference for works that are more expressive than laundry lists of facts and basic ideas. As I explore in a previous work, both authors and consumers of authored works enjoy the expressive twists and turns in authored works.⁹⁶ Central to our society's appreciation of

⁹⁰ See 5 RONALD D. ROTUNDA & JOHN E. NOWAK, TREATISE ON CONSTITUTIONAL LAW: SUBSTANCE AND PROCEDURE § 20.6 (2014); 1 RODNEY A. SMOLLA, SMOLLA & NIMMER ON FREEDOM OF SPEECH § 2.3 (2014); Jason Mazzone, *Speech and Reciprocity: A Theory of the First Amendment*, 34 CONN. L. REV. 405, 407–17 (2002).

⁹¹ *Abrams v. United States*, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting).

⁹² Mazzone, *supra* note 87, at 408; see also Kent Greenawalt, *Free Speech Justifications*, 89 COLUM. L. REV. 119 (1989).

⁹³ See, e.g., Thomas W. Joo, *True Believer*, 45 U.C. DAVIS L. REV. 1823, 1828–29 (2012).

⁹⁴ See *infra* Part IV.I (discussing the justification for the idea-expression distinction in copyright law). Jack Balkin's democratic-culture theory also justifies this freedom. See *infra* text accompanying notes 133–37.

⁹⁵ See *infra* Part IV.I.

⁹⁶ See Fromer, *supra* note 15.

artistic creativity is a “component subjective or emotional problem that [an author] carried through to the final product” using personalized expression.⁹⁷ Given the emphasis on this particular characteristic of artistic creativity, authors typically like creating expressively, and society enjoys consuming expressive works.⁹⁸

Given this hedonic preference, it is easy to see that, in addition to copyright law encouraging transmission of knowledge, it ought to promote interesting expression.⁹⁹ It is likely for this reason that Alan Durham invokes information theory to argue that works that convey much information—in that they are spontaneous and unpredictable forms of expression—should be afforded copyright protection, whereas those that do not—because they are redundant and predictable—should not.¹⁰⁰

Nonetheless, instead of focusing on the encouragement of enjoyable variations in expression, this section focuses on copyright’s important goal of encouraging the production, transmission, and preservation of knowledge, broadly construed.¹⁰¹ This focus leads to the opposite stance as Durham’s: that copyright law is and ought to be concerned with encouraging all sorts of redundancies so as to disseminate and preserve systematic, factual, and cultural knowledge.¹⁰²

⁹⁷ *Id.* at 1494.

⁹⁸ *See id.* at 1474–79.

⁹⁹ *See supra* note 85.

¹⁰⁰ Alan L. Durham, *Copyright and Information Theory: Toward an Alternative Model of “Authorship”*, 2004 BYU L. REV. 69, 124; cf. Robert A. Heverly, *The Information Semicommons*, 18 BERKELEY TECH. L.J. 1127, 1148 (2003) (mentioning different definitions of “information,” including Shannon’s, as a precursor to analyzing what kind of property information is). There have been discussions of information costs in copyright and patent law with regard to delineating information about creations. *See* Jeanne C. Fromer, *Claiming Intellectual Property*, 76 U. CHI. L. REV. 719 (2009) (developing a two-dimensional taxonomy for patent and copyright claiming systems and analyzing how the claiming system can affect innovation); Clarisa Long, *Information Costs in Patent and Copyright*, 90 VA. L. REV. 465, 499 (2004) (describing the differences between patent and copyright law with respect to how each handles the information asymmetry between owners and observers); Henry E. Smith, *Intellectual Property as Property: Delineating Entitlements in Information*, 116 YALE L.J. 1742, 1745–46 (2007) (contrasting copyright and patent claiming rights and the governance regimes generated as a result of those rights). These analyses of information costs relate directly to the costs of delineating property rights in creations and do not otherwise speak to the intersection of information theory and the sorts of creations copyright law directly seeks to encourage.

¹⁰¹ The following section also focuses on conversations about and reuses of particular pieces of expression that are enjoyable in and of themselves. *See infra* Part III.B.

¹⁰² A similar observation, in opposition to Durham’s, is true with regard to redundancy in expression—not just in knowledge—something I take up in the next section.

That is, authored works of the sort we want to encourage frequently have the commendable purpose of transmitting knowledge of one kind or another. But other forces—First Amendment values and a hedonic preference for expressive works—push these works away from comprising a concise and clear articulation of the knowledge that they are transmitting. Instead, due to their noisy expressiveness, these works can easily obscure the underlying knowledge the author might also be trying to transmit in the work.¹⁰³ In addition, consumers of the authored work might have different “codebooks” than the author, causing a misinterpretation of or failure to see the knowledge underlying the work.¹⁰⁴ In essence, authors themselves introduce noise—via expressiveness—into their works in ways that cloak the knowledge they intend to convey and might in fact otherwise communicate in a clearer fashion.

Information theory speaks to precisely this situation. It suggests that noise in a message transmission can be reduced, if not entirely overcome, by introducing redundancy into the message.¹⁰⁵ That is, so that consumers can reliably decode the messages of knowledge that authors possess and convey wrapped up in noise, it is crucial for that knowledge to be communicated in redundant ways. Redundancy can occur within a particular work or through repetition of information across multiple works.¹⁰⁶ These redundancies make it more likely that consumers will gain access to the transmitted knowledge. In fact, oral cultures absorbed this lesson centuries ago by realizing the need to encode works, such as Homer’s *Odyssey*, with redundancy so that they could be remembered and transmitted without loss of the encoded information.¹⁰⁷ Similarly, copyright law ought to be concerned with ensuring that useful redundancies occur in copyrighted works to help society decode the knowledge embedded in these works. As I discuss in the next Part, copyright law in fact implements these sorts of redundancies all throughout.¹⁰⁸

¹⁰³ Although printing errors and the like are known to happen, it is less likely, however, that errors will occur in transmitting the message than errors resulting from semantic noise. *But cf.* Devin Coldewey, *Copier Conundrum: Xerox Machines Swap Numbers During Scans*, CNBC (Aug. 7, 2013, 10:44 AM), <http://www.cnbc.com/id/100945451> (indicating how a design flaw caused Xerox devices to alter numbers randomly in scanned documents).

¹⁰⁴ See *supra* text accompanying notes 64–66.

¹⁰⁵ See *supra* Part II.

¹⁰⁶ See *supra* text accompanying notes 51–65 (discussing redundancy’s role in information theory).

¹⁰⁷ GLEICK, *supra* note 26, at 34–35.

¹⁰⁸ In focusing on other aspects of information theory than that addressed here, David Opderbeck has some skepticism about the relevance of information theory to making information policy. See David W. Opderbeck, *Deconstructing Jefferson’s Candle: Towards a Critical Realist Approach to Cultural Environmentalism and Information Policy*, 49 JURIMETRICS J. 203 (2009) (arguing that information theory has been implicitly adopted by those attempting to reconcile traditional views of intellectual property with

Furthermore, given that preservation of this knowledge over space and time is important, ensuring dissemination and endurance of authored messages is important. These goals too can be accomplished by implementing various redundancies in copyright law—which it already does in some ways—as I discuss in the next Part.

All in all, given copyright's goal of encouraging the production, dissemination, and preservation of knowledge, information theory suggests that copyright law ought to be encouraging useful forms of redundancy to boost the viability of these goals.¹⁰⁹

postmodern views deemphasizing the author's centrality). Opderbeck reasons that information theory suggests that information is non-rivalrous, has no meaning until it is interpreted, and frequently ought to be left to an open "commons." *Id.* at 223–27. He then advocates as a better basis for information policy a critical realist view that information is the infrastructure of all human communities. *Id.* at 227–43.

¹⁰⁹ A few others connect information theory and other forms of intellectual property. Dan Burk writes that patent law's distinction between inventive products and processes breaks down when it comes to biotechnology, precisely because molecules are "channels for informational transfer processes." Dan L. Burk, *The Problem of Process in Biotechnology*, 43 HOUS. L. REV. 561, 563, 584–88 (2006). Peter Junger relies on information theory to argue that computer software should not be patentable, because it processes and produces information. Peter D. Junger, Manuscript, *You Can't Patent Software: Patenting Software Is Wrong*, 58 CASE W. RES. L. REV. 333, 334–35 (2008). Deven Desai draws on information theory to argue that trademark law ought to reconceptualize trademarks as channels through which many actors can transmit messages to enhance information flow instead of as messages controlled solely by trademark holders. Deven R. Desai, *Response: An Information Approach to Trademarks*, 100 GEO. L.J. 2119 (2012). Other discussions of information theory crop up occasionally in analysis of other legal subjects, such as cyberlaw, communications law, contract law, property law, and First Amendment Law. See Derek E. Bambauer, *Conundrum*, 96 MINN. L. REV. 584, 587 (2011) (taking an information-based approach to cybersecurity, by "focus[ing] on access and alteration of data and on guaranteeing its integrity" to suggest that redundancy in networks and storage improve the ability to withstand cyberattacks); Yochai Benkler, *Some Economics of Wireless Communications*, 16 HARV. J.L. & TECH. 25, 30 (2002) (arguing in favor of developing open wireless networks, and relying on information theory to show that even weak signals can benefit from processing and cooperation gains to challenge property-based communications networks); George S. Geis, *Automating Contract Law*, 83 N.Y.U. L. REV. 450, 463 (2008) (arguing that the insights of information theory might be used as a basis for automating empirical analysis of contract law); Lyrissa Lidsky, *Public Forum 2.0*, 91 B.U. L. REV. 1975, 2016–17 (2011) (maintaining that information theory is a linear model of communication, and as such, is over-simplistic and this model has had a detrimental effect on Supreme Court jurisprudence on public forum doctrine); Andrea M. Matwyshyn, *Hidden Engines of Destruction: The Reasonable Expectation of Code Safety and the Duty To Warn in Digital Products*, 62 FLA. L. REV. 109, 136–37 (2010) (relying on information theory to argue for a reasonable expectation of code safety, based on the knowledge imbalance between consumers and creators and operators of code); Henry E. Smith, *Modularity in Contracts: Boilerplate and Information Flow*, 104 MICH. L. REV. 1175 (2006) (reasoning that contractual provisions work as modules, falling on a spectrum from intensive—carrying much information and being heavily context-bound—to extensive—carrying little information and fitting into almost any context); Henry E. Smith, *The Language of Property: Form, Context, and Audience*, 55 STAN. L. REV. 1105, 1108 (2003) (relying on the insight that less expected messages carry more information for a spectrum of information intensiveness, ranging from intensive—information-heavy but with high processing costs for third-party or non-intended audiences—on one end and extensive—less

B. *Expression Generating a Noisy Discussion*

Redundancy of knowledge by encoding it in multiple forms of expression is not the only form of redundancy that copyright law ought to encourage. As I suggest herein, copyright law also ought to promote redundancies in expression that is valued in and of itself. Doing so will help generate a cacophony of meanings, contextual understandings, and reinterpretations of particular pieces of expression in ways that allow meaning to attach to this expression, which might otherwise be hard to pierce.

Given that society frequently values expression for its own sake—be it beautifully expressed, poignantly communicated, or otherwise—copyright law ought to be encouraging that form of expression in the first place.¹¹⁰ Examples abound, including poetry, abstract painting, photography, and classical music. As Alan Durham points out, expression that is valued for its expressiveness—like that just described—tends to be unpredictable or unexpected (possessing high informational content).¹¹¹ Durham argues that this sort of expression is what copyright law ought to encourage.¹¹² That much is true enough.

But that is just the beginning of what copyright law ought to be encouraging with regard to expression that is valuable in and of itself, not the end of it. Once this valuable expression exists, people will consume it. Some of those that find the expression to be momentous or interesting will seek to understand what this expression means. They might try to suss out what the author intended the expression to mean, or not infrequently, they will find meaning different than what the author might have intended the expression to mean.¹¹³

information-packed but with lower processing costs for a general audience—on the other, for a framework of how to communicate property entitlements).

¹¹⁰ See *supra* text accompanying notes 96–100.

¹¹¹ Durham, *supra* note 100; see *supra* text accompanying note 100. Of course, not all unpredictable expression—such as completely meaningless expression that does not follow syntactic rules—will be considered by society to be valuable. But the expression that society values for its expressiveness will typically be of the unpredictable sort.

¹¹² Durham, *supra* note 100.

¹¹³ Of course, there might be authors who intend merely to transmit to recipients whatever the recipient chooses to understand in the work. Cf. Amy B. Cohen, *Copyright Law and the Myth of Objectivity: The Idea-Expression Dichotomy and the Inevitability of Artistic Value Judgments*, 66 IND. L.J. 175, 184–87 (1990) (discussing different theories of art); Fromer, *supra* note 100, at 789–90 (discussing the complications an institutional theory of art makes for requiring authors to characterize their creations); Alfred C. Yen, *Copyright Opinions and Aesthetic Theory*, 71 S. CAL. L. REV. 247, 258–60 (1998) (providing an overview of the institutional definitions of art and noting its advantages for explaining modern art).

Take the example of Lewis Carroll's "Jabberwocky," a poem that seems to be nothing more than its expression: "'Twas brillig, and the slithy toves / Did gyre and gimble in the wabe; / All mimsy were the borogoves, / And the mome raths outgrabe . . .'"¹¹⁴ In Carroll's novel *Through the Looking-Glass and What Alice Found There*, the character Humpty Dumpty explains the verse to Alice to mean something along the lines of, "At four o'clock, lithe and slimy badger-like creatures go round and round like a gyroscope and make holes like a gimblet in the grass plot around a sundial. All flimsy and miserable were the thin birds with feathers sticking out all around, and the lost green pigs made a noise between bellowing and whistling, with a sneeze in the middle."¹¹⁵ That explanation has not stopped varying audiences at different times from conferring another meaning on the poem, sometimes with a particular purpose in mind. For instance, the religiously oriented Continental Historical Society states that the poem (and Humpty Dumpty's explication of it too) is in fact about seduction and illicit love, apparently as a tactic to establish that the author of the verse was pious.¹¹⁶ By contrast, a scholar of English romanticism argues for infusing the poem with "literary experience," thereby understanding the poem to be "a parody of the medieval quest-romance in balladic verse. . . . implicated with the myths of heroes, dragon-slayers, and the cycles of nature and human life."¹¹⁷ And a poetry explanation guide for students seeking to show them the relevance of canonical poetry to their everyday lives states that the poem "is about facing your demons" in a world "filled with all sorts of unknowns."¹¹⁸ Other artists have borrowed from "Jabberwocky," such as some members of the Monty Python comedy group, which filmed a movie titled "Jabberwocky" about "[a] young peasant, with no interest in adventure or fortune, [being] mistaken as the kingdom's only hope when a horrible monster threatens the countryside."¹¹⁹ Moreover, third parties have translated

¹¹⁴ LEWIS CARROLL, *THROUGH THE LOOKING-GLASS AND WHAT ALICE FOUND THERE* 21 (1897).

¹¹⁵ See *id.* at 126–29.

¹¹⁶ See John Pennington, *Reader Response and Fantasy Literature: The Uses and Abuses of Interpretation in Queen Victoria's Alice in Wonderland*, in *FUNCTIONS OF THE FANTASTIC: SELECTED ESSAYS FROM THE THIRTEENTH INTERNATIONAL CONFERENCE ON THE FANTASTIC IN THE ARTS* 55, 60–61 (Joe Sanders ed., 1995) (citing Continental Historical Society, *Queen Victoria's Through the Looking-Glass* (1986) (unpublished manuscript)) (contextualizing this interpretation as being made by a religious-minded group interested in advocating for Queen Victoria's authorship of *Alice in Wonderland* and her piety).

¹¹⁷ HAZARD ADAMS, *ACADEMIC CHILD: A MEMOIR* 57 (2008).

¹¹⁸ *Jabberwocky*, SHMOOP, <http://www.shmoop.com/jabberwocky/> (last visited Sept. 2, 2014).

¹¹⁹ *Jabberwocky* (1977), IMDB, <http://www.imdb.com/title/tt0076221/> (last visited Sept. 2, 2014).

“Jabberwocky” into other languages, such as German and French.¹²⁰ These translators had to choose which aspects of the original poem to preserve: its tense, its imagery conveyed even through nonsense words, and so forth.¹²¹ For example, one French translator chose to convert the tense from past to present to avoid awkward verb conjugations that would destroy the original’s rhythm and sought to find French nonsense words that conveyed similar imagery to French speakers as the English did to English speakers (such as “lubricilleux” for “slithy”).¹²² In light of these varied understandings and uses of “Jabberwocky,” it is perhaps not surprising that the U.S. Supreme Court has indicated specifically that this poem is “unquestionably shielded” by the First Amendment as protected expression that “‘communicat[es] ideas.’”¹²³

More generally, as the influential literary theory of reader-response criticism posits, different audiences locate various meanings in expression.¹²⁴ They might do so because of their particular circumstances: their goals, their time period, or their individual backgrounds.¹²⁵ They might be individual readers¹²⁶—like the romanticism scholar referenced above—or a collection of readers—an interpretive community such as the Continental Historical Society—who, as Stanley Fish posits, share linguistic or other background elements that constrain the group to read particular expression similarly in certain ways.¹²⁷ Furthermore, they also—like the “Jabberwocky” filmmakers—might reuse or make reference to this expression as a way to comment on, criticize, or repurpose it.¹²⁸ Just as money and citizenship have value

¹²⁰ See DOUGLAS R. HOFSTADTER, GÖDEL, ESCHER, BACH: AN ETERNAL GOLDEN BRAID 366–68 (1979) (providing German translation by Robert Scott entitled “Der Jammerwoch,” and a French translation by Frank L. Warrin entitled “Le Jaseroque”).

¹²¹ On the complicated question of how translations can seek to remain faithful to their original text, see DOUGLAS R. HOFSTADTER, LE TON BEAU DE MAROT: IN PRAISE OF THE MUSIC OF LANGUAGE (1998).

¹²² See HOFSTADTER, *supra* note 120, at 366–68.

¹²³ Hurley v. Irish-Am. Gay, Lesbian & Bisexual Grp. of Boston, Inc., 515 U.S. 557, 569 (1995) (quoting W. Va. Bd. of Ed. v. Barnette, 319 U.S. 624, 632 (1943)).

¹²⁴ See, e.g., Jane P. Tompkins, *An Introduction to Reader Response Criticism*, in *READER-RESPONSE CRITICISM: FROM FORMALISM TO POST-STRUCTURALISM* ix, ix (Jane P. Tompkins ed., 1980) (“Reader-response critics would argue that a poem cannot be understood apart from its results. Its ‘effects,’ psychological and otherwise, are essential to any accurate description of its meaning, since that meaning has no effective existence outside of its realization in the mind of a reader.”).

¹²⁵ See, e.g., *id.*

¹²⁶ See *id.*

¹²⁷ STANLEY FISH, IS THERE A TEXT IN THIS CLASS 147–74 (1980); accord Laura A. Heymann, *Everything Is Transformative: Fair Use and Reader Response*, 31 COLUM. J.L. & ARTS 445, 455–56 (2008).

¹²⁸ See, e.g., Amy M. Adler, *Against Moral Rights*, 97 CAL. L. REV. 263, 279–94 (2009) (arguing that there is artistic value in destroying and changing works of art); cf. *Cariou v. Prince*, 714 F.3d 694, 707–08 (2d

principally as social constructions,¹²⁹ expression obtains value from its social construction in specific ways by particular groups or individuals. In this way, expression gains value through its traction with society at large, thereby having more purchase.

Furthermore, as Jack Balkin cogently observes in the closely related context of the First Amendment, “[t]he purpose of freedom of speech . . . is to promote a democratic culture,” one “in which individuals have a fair opportunity to participate in the forms of meaning making that constitute them as individuals.”¹³⁰ He reasons that viewed this way, the freedom of expression enhances “individual liberty” and “collective self-governance,” by providing everyone—not just elites—with the ability to participate in and shape culture.¹³¹ In constructing this theory, Balkin appreciates that speakers will need to rely on preexisting materials in critical ways:

Freedom of speech is appropriate because it draws on existing cultural resources; it builds on cultural materials that lay to hand. Dissenters draw on what they dislike in order to criticize it; artists borrow from previous examples and build on artistic conventions; even casual conversation draws on common topics and expressions. People participate in culture through building on what they find in culture and innovating with it, modifying it, and turning it to their purposes. Freedom of speech is the ability to do that. In a democratic culture people are free to appropriate elements of culture that lay to hand, criticize them, build upon them, and create something new that is added to the mix of culture and its resources.¹³²

Balkin’s reasoning extends to copyright law.¹³³ Reframing this section’s observations in the context of information theory, society enjoys—and

Cir. 2013) (emphasizing, in the context of a fair-use inquiry, the great value in works that use existing works in transformative ways, be it through commentary, criticism, parody, or news reporting).

¹²⁹ See STEVEN PINKER, *THE BLANK SLATE: THE MODERN DENIAL OF HUMAN NATURE* 202 (2002); JOHN R. SEARLE, *THE CONSTRUCTION OF SOCIAL REALITY* 32 (1995). See generally PETER L. BERGER & THOMAS LUCKMANN, *THE SOCIAL CONSTRUCTION OF REALITY: A TREATISE IN THE SOCIOLOGY OF KNOWLEDGE* (1967) (providing a classic assessment of knowledge across society).

¹³⁰ Jack M. Balkin, *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, 79 N.Y.U. L. REV. 1, 3 (2004).

¹³¹ *Id.* at 3–4.

¹³² *Id.* at 4–5.

¹³³ It is unsurprising that Balkin’s reasoning applies to copyright law too. Neil Netanel draws out important connections between the two legal areas:

[C]opyright is in essence a state measure that uses market institutions to enhance the democratic character of civil society. In supporting a market for authors’ works, copyright serves two democracy-enhancing functions. The first is a production function. Copyright provides an

copyright law is well-placed to encourage—many forms of expression with high informational content (in that such expression communicates in an unpredictable or unexpected way).¹³⁴ This expression is valuable not solely when readers quietly and personally enjoy these expressive works, but also when they reuse that expression to discuss its significance and meaning or repurpose that expression.¹³⁵ These further uses require that the valuable information conveyed in the underlying work—the expression itself—be used again and again in subsequent works of interpretation and repurposing.¹³⁶ Taken all together, these subsequent works represent a noisy and valuable conversation about the underlying expression. Each of these works uses the expression redundantly to signal its part in this overarching “conversation” about the expression. This conversation is indubitably an important component of copyright law’s constitutional goal of “promot[ing] the Progress of Science.”¹³⁷ As a result, then, copyright law ought to permit and encourage the redundant reuse of valuable expression to work over the expression’s meaning(s) and significance and to repurpose it in various contexts. The expression itself provides the thread that connects this “conversation” across different interpretations and uses.

In sum, so as to encourage the creation, dissemination, and preservation of the two essential aspects of expression—the underlying knowledge it contains and the expression in and of itself—copyright law ought to encourage those valuable aspects themselves to be encoded and transmitted redundantly. I now turn to a discussion of how copyright law can (and often does already) implement these critical redundancies.

incentive for creative expression on a wide array of political, social, and aesthetic issues, thus bolstering the discursive foundations for democratic culture and civic association. The second function is structural. Copyright supports a sector of creative and communicative activity that is relatively free from reliance on state subsidy, elite patronage, and cultural hierarchy. The democratic paradigm requires that copyright protection be sufficiently strong to ensure support for copyright’s production and structural functions. But at the same time, it would accord authors a *limited* proprietary entitlement, designed to make room for—and, indeed, to encourage—many transformative and educative uses of existing works.

Neil Weinstock Netanel, *Copyright and a Democratic Civil Society*, 106 YALE L.J. 283, 288 (1996).

¹³⁴ See *supra* text accompanying notes 111–12.

¹³⁵ Cf. Edgar Allen Poe, *Sarah Margaret Fuller*, GODEY’S LADY’S BOOK, Aug. 1846, at 72, 74 (“The soul is a cypher, in the sense of a cryptograph; and the shorter a cryptograph is, the more difficulty there is in comprehension.”).

¹³⁶ Cf. Margaret Chon, *Postmodern “Progress”: Reconsidering the Copyright and Patent Power*, 43 DEPAUL L. REV. 97, 102 (1993) (interpreting Congress’s constitutional power to craft intellectual property laws to be, in part, “the increasing recognition of a plurality of claims and meanings”).

¹³⁷ U.S. CONST. art. I, § 8, cl. 8; see *supra* Part I.

IV. APPLICATIONS IN COPYRIGHT LAW

This Part discusses central copyright doctrines and explains how they mesh with the general information theory of copyright law set out above. It addresses, in turn, the idea-expression distinction and related doctrines, originality and independent creation, the copyright holder's exclusive rights to reproduce the work and to prepare derivative works, fair use, first sale doctrine, secondary liability, treatment of search engines, encouragement of works appearing in multiple formats, and deposit of copyrighted material. It also addresses possible extensions to copyright law, such as ways to correct errors in transmissions, much like the error-correcting codes in information theory. Importantly interspersed with these issues is the question whether the first author, the public, or follow-on authors are best placed to have particular rights.

A. *Idea-Expression Distinction*

Of all of copyright law's particulars, the idea-expression distinction and related doctrines most directly suggest copyright law's goal of transmitting redundant information so as to promote the creation, dissemination, and preservation of knowledge.

As discussed above, copyright law protects original works of authorship fixed in a tangible medium of expression.¹³⁸ Even if it is an original work of authorship, however, copyright law excludes from protection "any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work."¹³⁹ Similarly, facts are not copyrightable.¹⁴⁰ Rather, copyright protection extends to the expression of particular ideas or facts rather than the ideas or facts themselves.¹⁴¹ For example, the expression in a play about star-crossed lovers would be copyrightable, but the idea of star-crossed lovers would not.¹⁴² And a biography of the Hindenburg disaster might be copyrightable as to its

¹³⁸ See *supra* text accompanying notes 1–8.

¹³⁹ 17 U.S.C. § 102(b) (2012); accord *Nichols v. Universal Pictures Corp.*, 45 F.2d 119, 121 (2d Cir. 1930).

¹⁴⁰ *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 344 (1991) (holding, however, that compilations of fact might be copyrightable). Courts sometimes attribute the lack of copyrightability of facts to them not being original to the author. See *id.* at 345.

¹⁴¹ See 17 U.S.C. § 102(b).

¹⁴² See *Nichols*, 45 F.2d at 121.

expression but not as to the underlying facts the author assembles into the book.¹⁴³

Related to the idea-expression distinction is the merger doctrine. When there is only one or a very limited number of ways to express an idea, it is as if the idea and expression have merged, rendering the expression just as uncopyrightable as the idea.¹⁴⁴ Also, *scènes à faire*—“incidents, characters or settings which, as a practical matter, are indispensable or standard in the treatment of a given topic”—such as buildings, lampposts, and pedestrians in a drawing of New York City blocks, are not protected under copyright law.¹⁴⁵

The exclusion of ideas, facts, *scènes à faire*, and the like from copyright protection has a purpose that relates to the discussion above about First Amendment values.¹⁴⁶ The exclusion emphasizes that some elements of authored works—their ideas and facts—belong to the public domain.¹⁴⁷ In fact, the Supreme Court notes that the idea-expression distinction “strikes a definitional balance between the First Amendment and the Copyright Act by permitting free communication of facts while still protecting an author’s expression.”¹⁴⁸ That is, the basic building blocks of expression ought to be left freely available for anyone to use.¹⁴⁹ It would be both inefficient and unfair to grant rights in these basic components that so many authors will need just because one person happened to employ them first. Doing otherwise would ultimately be detrimental to generating a robust body of authored works.¹⁵⁰

¹⁴³ See *Hoehling v. Universal City Studios, Inc.*, 618 F.2d 972, 978 (2d Cir. 1980).

¹⁴⁴ See *Morrissey v. Procter & Gamble Co.*, 379 F.2d 675, 679 (1st Cir. 1967). Some courts do not find the expression in these cases uncopyrightable, but instead declare the copyright in these works to be “thin,” in that only virtually identical copies might infringe. See, e.g., *Johnson Controls, Inc. v. Phoenix Control Sys., Inc.*, 886 F.2d 1173, 1175 (9th Cir. 1989).

¹⁴⁵ *Steinberg v. Columbia Pictures Indus., Inc.*, 663 F. Supp. 706, 713 (S.D.N.Y. 1987) (quoting *Walker v. Time Life Films, Inc.*, 615 F. Supp. 430, 436 (S.D.N.Y. 1985)).

¹⁴⁶ See *supra* text accompanying notes 90–95.

¹⁴⁷ See *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 556 (1985).

¹⁴⁸ *Id.* (quoting *Harper & Row Publishers, Inc. v. Nation Enters.*, 723 F.2d 195, 203 (2d Cir. 1983) (internal quotation marks omitted)).

¹⁴⁹ Peter Lee, *The Evolution of Intellectual Infrastructure*, 83 WASH. L. REV. 39, 63–64 (2008).

¹⁵⁰ Cf. Margaret Chon, *Sticky Knowledge and Copyright*, 2011 WIS. L. REV. 177, 180–81 (arguing that there is “sticky knowledge”—“aspects of [unencoded] knowledge that have a stubbornly and sometimes irreducibly social dimension”—unprotected by copyright law, and therefore available in the public domain). In addition, for the functional categories of excluded material (such as processes), the exclusion delineates a boundary line between that which is protectable by copyright—expression—and that which is protectable by patent—inventions. See *Baker v. Selden*, 101 U.S. 99, 102–03 (1879); Pamela Samuelson, *Why Copyright Law Excludes Systems and Processes from the Scope of Its Protection*, 85 TEX. L. REV. 1921, 1926 (2007).

This collection of doctrines leaves certain aspects of authored material freely available for all authors to draw upon in crafting their works. It liberates precisely those aspects—ideas, facts, stock elements—that might readily get buried under the noise of expression in any one particular work.¹⁵¹ By leaving these valuable building blocks in the public domain, copyright law allows their redundant use again and again across different works. Copyright law goes yet further by encouraging this redundant use: It offers protection for subsequent forms of expression in works containing ideas or facts that have previously been expressed (even in copyrighted material),¹⁵² providing further incentive to create works that contain expression conveying this unprotected material.¹⁵³ All in all, valuable ideas, facts, stock elements, themes, and the like—things copyright law would like to have produced and disseminated widely¹⁵⁴—can be and are used again and again throughout different works. The idea-expression distinction encourages redundancy—weaving knowledge into multiple expressive works—so that the knowledge might ultimately and collectively cut through the noise of the expressive works in which it is encoded. Audiences are more likely to decode the valuable knowledge embodied in multiple noisy works than in just one noisy work, both because exposure to it in multiple works maximizes the chance of decoding it properly and because audience members can choose the precise works that speak to them most clearly to decipher the encoded knowledge.

Relatedly, copyright's merger and *scènes à faire* doctrines suggest that there are certain forms of expression that are shorthand for a complex set of ideas and themes that ought to be in the public domain so that they might be drawn upon again and again. These doctrines connect to the two valuable aspects of expression that copyright law seeks to promote: knowledge embodied in expression and expression that is valuable itself.¹⁵⁵ By employing these doctrines, copyright law emphasizes that when there is a singular (or nearly singular) efficient way of communicating a complex message, it cannot be roped off for one author alone.¹⁵⁶ Rather, these efficient encodings are left

¹⁵¹ See *supra* Part III.A.

¹⁵² See *F.A. Davis Co. v. Wolters Kluwer Health, Inc.*, 413 F. Supp. 2d 507, 512 (E.D. Pa. 2005) (“[I]f an idea can only be expressed in one way, that one way will never be subject to copyright protection; if an idea can be expressed in many ways, those many ways may be copyrighted.”).

¹⁵³ See *supra* Part I (discussing how American copyright law has a utilitarian basis).

¹⁵⁴ See *supra* Part III.A.

¹⁵⁵ See *supra* Part III.A.

¹⁵⁶ See *supra* text accompanying notes 39–41 (discussing data compression and designing efficient codes).

free to all authors so that consumers of authored works conveying these particular types of messages will have the benefit of easily decoding the intricate set of ideas and themes that copyright law wants them to acquire. By so doing, copyright law furthers the cause of transmitting valuable knowledge, by ensuring that no author has exclusive rights to use it if there is only one way (or very few ways) of expressing it.

The foregoing analysis also suggests that courts might have gone astray in generally treating fictional “facts” (facts about a work of fiction, such as the various magic spells in the *Harry Potter* novels) as original expression that might not be copied.¹⁵⁷ When these fictional facts are presented in compendia analyzing them as facts about the internal world of fiction, information theory’s notion of redundancy might similarly suggest that they ought to be freely available, as if they were facts about the real world.¹⁵⁸

The merger doctrine also advances copyright’s goal of generating a noisy discussion about valuable expression. When expression itself is valued, it is frequently on the basis that there is no better way to communicate whatever is being expressed. Valued expression’s meaning will often be disputed and picked over,¹⁵⁹ but one thing upon which these conversants would generally agree is that the expression itself is inseparable from its meaning.¹⁶⁰ That is, there is no better way to communicate the expression’s message than with that expression. As philosopher of aesthetics Roger Scruton puts it:

Suppose you ask me what is the content of Van Gogh’s famous painting of the yellow chair. What exactly does it *mean*? you ask:

¹⁵⁷ See, e.g., *Castle Rock Entm’t, Inc. v. Carol Publ’g Grp., Inc.*, 150 F.3d 132, 135 (2d Cir. 1998) (suit by copyright owners of *Seinfeld* television series against author of trivia book about the series); *Warner Bros. Entm’t Inc. v. RDR Books*, 575 F. Supp. 2d 513, 517 (S.D.N.Y. 2008) (suit by copyright owners of *Harry Potter* book series against author of encyclopedia for the series).

¹⁵⁸ Cf. Irene Segal Ayers, Comment, *The “Facts” of Cultural Reality: Redrawing the Line Between Fact and Expression in Copyright Law*, 67 U. CIN. L. REV. 563, 584–86 (1999) (suggesting that uses of fictional facts might qualify as fair use of copyrighted material). But cf. Alan L. Durham, *Speaking of the World: Fact, Opinion and the Originality Standard of Copyright*, 33 ARIZ. ST. L.J. 791, 824 n.228 (2001) (“[Fictional] ‘facts’ are both original to the author and inseparable from the author’s expression.”).

¹⁵⁹ See *supra* Part III.B.

¹⁶⁰ See, e.g., CHARLES ALTIERI, *PAINTERLY ABSTRACTION IN MODERNIST AMERICAN POETRY* 408–09 (1989) (making this observation with regard to visual art); I.A. RICHARDS, *COLERIDGE ON IMAGINATION* 198 (1934) (making this observation with regard to poetry). Relatedly, one federal court ruling in a case involving a copyright infringement claim of a photograph reasoned how it might be impossible to disentangle idea from expression in the context of photographs because the distinction is meaningless there. *Mannion v. Coors Brewing Co.*, 377 F. Supp. 2d 444, 455–61 (S.D.N.Y. 2005). That court further explained that “every observer will have a different interpretation” of the photograph and “it is not clear that there is any real distinction between the idea in a work of art and its expression.” *Id.*

what am I supposed to *understand*, about this chair, or about the world, from looking at this picture? . . . I am likely to argue that this painting is saying something special about this particular chair, and also about the world as seen through the image of this chair. I might try to put my thoughts and feelings into words. ‘It is an invitation to see the life that spreads from people into all their products, the way in which life radiates from the meanest things, so that nothing is at rest, all is becoming.’ But couldn’t he have written that message at the bottom of the canvas? Why does he need a chair to communicate a thought like that? I am likely to respond that my words are only a gesture; that the real meaning of the painting is *bound up with, inseparable from*, the image—that it resides in the very shapes and colours of the chair, is inseparable from Van Gogh’s distinctive style, and cannot be translated completely into another idiom.¹⁶¹

That does not mean that nothing can be said about the expression’s meaning. As Scruton continues, “For the most part you can say much about the meaning of a poem, a painting—even a work of music. But what you say will not explain the particular intensity of meaning which makes the work of art into the irreplaceable vehicle of its content.”¹⁶²

Thus, when expression itself is especially valued, copyright law would be ill-advised to allow the author to control that expression from being used in subsequent works to comment on or critique the author’s expression. The merger doctrine suggests why: Protecting expression that has no precise “synonym” ropes off that expression from commentary, criticism, or reuse in another context.¹⁶³ Copyright law would thus be well-advised to extend notions of the merger doctrine to this scenario to bar findings of infringement when expression is reused in these ways.¹⁶⁴

¹⁶¹ ROGER SCRUTON, *BEAUTY* 109–10 (2009).

¹⁶² *Id.* at 113.

¹⁶³ *Cf.* RICHARD MOON, *THE CONSTITUTIONAL PROTECTION OF FREEDOM OF EXPRESSION* 45 (2000) (“[R]estriction of a particular form of expression always affects the opportunity to communicate some messages more than others. . . . A restriction on a particular form of expression must be understood as a restriction on meaning, even if the purpose of the restriction is not to prevent the communication of a particular message.”).

¹⁶⁴ *Cf.* Julie E. Cohen, *Copyright as Property in the Post-Industrial Economy: A Research Agenda*, 2011 *WIS. L. REV.* 141, 148 (“A copyright regime that works to enable the production of big-budget Hollywood movies and long-running television series is not a bad thing. Mass culture . . . gives us things to talk about with one another, to celebrate or criticize, and to define ourselves against.”). As a practical matter, copyright law might want to reach this conclusion only for certain very valued forms of expression rather than all expression. Otherwise, copyright law would protect no expression under application of this reasoning.

All in all, these doctrines suggest that there is value in leaving certain material to the public to use when we want it used redundantly. This suggestion is particularly apt when it is preferable that there be multiple communicators of this material, rather than just the original author. If we want the knowledge that is encoded in noisily expressive works to be transmitted effectively, we likely want multiple authors encoding that same knowledge in different noisy ways (namely, authors' particular styles). That is likely to maximize the decoding of that knowledge by the public, particularly given that different people or groups are likely to find different forms of expression most natural to understand and learn from.¹⁶⁵ Similarly, to have a noisy discussion about valuable expression's meaning, we want multiple authors using that expression. Were it otherwise, the conversation would likely be monotonic and would lack enough noise.

B. Originality and Independent Creation

Information theory can also explain copyright law's requirement of originality and the infringement defense of independent creation. Recall that copyright law safeguards "original works of authorship fixed in any tangible medium of expression, now known or later developed," including literary works, sound recordings, movies, and computer software code.¹⁶⁶ The Supreme Court's most recent formulation of the originality requirement occurred in *Feist Publications, Inc. v. Rural Telephone Service Co.*,¹⁶⁷ a case involving the copyrightability of a local telephone directory listing names in alphabetical order along with their corresponding towns and telephone numbers.¹⁶⁸ The *Feist* Court held that "[o]riginal, as the term is used in copyright, means only that the work was independently created by the author (as opposed to copied from other works), and that it possesses at least some minimal degree of creativity."¹⁶⁹ The requisite level of creativity, according to the Supreme Court, "is extremely low; even a slight amount will suffice."¹⁷⁰ A

¹⁶⁵ Cf. Jeanne C. Fromer, *Patent Disclosure*, 94 IOWA L. REV. 539, 573–74 (2009) (suggesting how disclosures in patent documents can be improved by "presenting the most useful pieces of technical information about an invention for technologists in multiple modalities," particularly because different readers learn best through distinct modalities).

¹⁶⁶ 17 U.S.C. § 102(a) (2012).

¹⁶⁷ 499 U.S. 340 (1991).

¹⁶⁸ *Id.* at 342.

¹⁶⁹ *Id.* at 345.

¹⁷⁰ *Id.* Some older decisions had reasoned otherwise, finding that copyright ought to be bestowed only on very creative works, of the type that "require[] genius for [their] construction." *Jollie v. Jacques*, 13 F. Cas. 910, 913 (C.C.S.D.N.Y. 1850) (No. 7437).

work must merely evidence “intellectual production, . . . thought, and conception.”¹⁷¹ Originality does not match up to a requirement of true novelty; a minimally creative work is protectable even if there is a nearly identical work, so long as the other work was not copied.¹⁷² As Judge Learned Hand observed, “[I]f by some magic a man who had never known it were to compose anew Keats’s Ode on a Grecian Urn, he would be an ‘author,’ and, if he copyrighted it, others might not copy that poem, though they might of course copy Keats’s.”¹⁷³ It is thus the rare work that will not meet the low threshold of originality. For example, the Court held that the telephone directory in *Feist* was insufficiently original because its factual raw data did not owe its existence to the directory creator and the selection and alphabetical arrangement of the directory entries was not creative enough.¹⁷⁴ The threshold for copyright protection is thus minimal but not absent.

Not only is independent creation a prerequisite to copyright protection, but it is also a defense against copyright infringement. A defendant will not be found to have infringed a plaintiff’s copyright if the plaintiff cannot show that the defendant actually copied from the plaintiff’s work.¹⁷⁵ If the defendant’s work was independently created, it is not an infringement, even if it is somehow identical to the plaintiff’s work.¹⁷⁶

The following sections address the two components of originality— independent creation and a modicum of creativity—in turn.

1. Independent Creation

Just as the idea-expression distinction and related doctrines allow multiple authors to reuse certain important components in copyrighted works to ensure their broader production and dissemination, so too the originality standard’s requirement of independent creation and the infringement defense of

¹⁷¹ *Feist*, 499 U.S. at 362 (quoting *Burrow-Giles Lithographic Co. v. Sarony*, 111 U.S. 53, 59–60 (1884)) (internal quotation marks omitted).

¹⁷² *Id.* at 345–46.

¹⁷³ *Sheldon v. Metro-Goldwyn Pictures Corp.*, 81 F.2d 49, 54 (2d Cir. 1936). Others might copy Keats’s poem because any copyright on it has long expired, leaving the work in the public domain. John C. O’Quinn, *Protecting Private Intellectual Property from Government Intrusion: Revisiting SmithKline and the Case for Just Compensation*, 29 PEPP. L. REV. 435, 504 n.455 (2002).

¹⁷⁴ *Feist*, 499 U.S. at 361–64. As another illustration, the Ninth Circuit held that a lamp design made up of preexisting parts was not sufficiently original to qualify for copyright protection. *See Lamps Plus, Inc. v. Seattle Lighting Fixture Co.*, 345 F.3d 1140, 1146–47 (9th Cir. 2003).

¹⁷⁵ *Peters v. West*, 692 F.3d 629, 635 (7th Cir. 2012).

¹⁷⁶ *Id.*

independent creation promote the redundant use of some expression in multiple works. They do so by respectively allowing and encouraging independent creation of works, even that which is identical to preexisting works. That is, if multiple authors independently happen to come up with the same message to convey in their works, copyright law permits each authored work to be copyrighted in its own right.¹⁷⁷ Similarly, there is no penalty of copyright infringement when the subsequent work was independently created.¹⁷⁸

The direct benefit of redundancy is apparent: the possibility of greater diffusion of the expression, by virtue of its embodiment in multiple works. There are other, perhaps less obvious, benefits.

For one thing, allowing and encouraging multiple authors to create similar—and even identical—works, if they do so independently, is desirable in relation to copyright law’s goal of “promot[ing] the Progress of Science.”¹⁷⁹ Independent reexpression of knowledge is valuable because it makes it more likely that that knowledge is true or valuable. That is, if multiple authors stumble or seize upon the same knowledge separately from one another, there is a strong chance that the knowledge is truthful or at least valuable to society.¹⁸⁰ It also might be reasonable to presume that the number of people that independently come up with a particular message is a proxy for the importance of that message being communicated broadly.¹⁸¹

Furthermore, encouragement of independent creation can be a valuable way to rid the infosphere of errors. As one example, when Ireland set out to map the country, its survey office sought to provide the surveyors with portable and

¹⁷⁷ See, e.g., *Sheldon*, 81 F.2d at 54.

¹⁷⁸ See, e.g., *Peters*, 692 F.3d at 635.

¹⁷⁹ U.S. CONST. art. I, § 8, cl. 8; see *supra* Part I.

¹⁸⁰ Cf. *United States v. Baldwin*, 54 M.J. 551, 555 (A.F. Ct. Crim. App. 2000) (“The military judge should have concentrated on the requirements of the rule for independent, direct or circumstantial evidence, corroborating the essential facts admitted in the confession sufficiently to give rise to an inference of their truth.”).

¹⁸¹ For other explanations of the independent-creation defense in copyright law, see Fromer, *supra* note 15, at 1492–93 (“With regard, then, to the requirement of independent creation, the emphasis is on the personal discovery of a subjective problem that artists express in their work.”); Long, *supra* note 100, at 525–33 (arguing that the costs of delineating copyrighted works augur for an infringement defense of independent creation). It should be noted that it could also be that the author himself or herself is an important component of understanding a particular work, perhaps in addition to other contextual information external to the work itself. Per this view, multiple similar works can be protectable because they “contain” or “convey” other information that is not redundant of the previous work.

accurate sets of logarithmic tables.¹⁸² To build these tables, the office looked to logarithmic tables previously published all over the world—from England to Germany to China—and discovered the same six errors in almost every one of those tables.¹⁸³ These tables all must have been copied from one another, leading to widespread dissemination of these errors.¹⁸⁴ Had these tables been computed independently, it is unlikely that the same errors would have cropped up so widely. As a different example, copies can introduce errors into a work when made from faulty memory, carelessness, or lack of deep understanding of the original work.¹⁸⁵ Independent creation maximizes the chance that care is taken to convey information in a work, to the benefit of society.

2. *Modicum of Creativity*

Consider now the second part of the originality requirement, that a work must contain a modicum of creativity.¹⁸⁶ Information theory can help concretize the amorphous requirement of creativity.¹⁸⁷

Take *Feist* itself. In that case, the Supreme Court ruled that the selection and alphabetical arrangement of the telephone directory entries was not creative enough to merit copyright protection.¹⁸⁸ One might wonder why it lacks this creativity when there are many other ways one might organize a telephone directory, such as randomly or by the last digit in each phone number. That is, one might argue that there was a creative choice in choosing an alphabetical ordering by last name from all of the various ways one might organize a telephone directory. In a theoretical vacuum divorced from the real world, there are surely multiple ways to organize the listings. However, in reality it would be foolish to choose an ordering that is random, by the last digit in each telephone number, or in the order in which people signed up for telephone service, because (pre-information age) society has seized upon

¹⁸² GLEICK, *supra* note 26, at 94.

¹⁸³ *Id.*

¹⁸⁴ *Id.*

¹⁸⁵ *See id.* at 94–95.

¹⁸⁶ *See supra* text accompanying notes 167–74.

¹⁸⁷ In a previous work, I discuss the notion of “creativity” in copyright law from the point of view of psychological and sociological work on creativity with regard to artistic works. *See* Fromer, *supra* note 15.

¹⁸⁸ *Feist Publ’ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 361–64 (1991).

alphabetical ordering by last name as the most efficient method of organization of a telephone directory for practical use in locating listings.¹⁸⁹

Consider another example, a taxonomy of dental procedures, upheld as copyrightable by the Seventh Circuit.¹⁹⁰ In this taxonomy, “dental procedures are classified into groups; each procedure receives a number, a short description, and a long description. For example, number 04267 has been assigned to the short description ‘guided tissue regeneration—nonresorbable barrier, per site, per tooth (includes membrane removal)’, which is classified with other surgical periodontic services.”¹⁹¹ The court ruled that the taxonomy is sufficiently creative for the originality requirement because “[d]ental procedures could be classified by complexity, or by the tools necessary to perform them, or by the parts of the mouth involved, or by the anesthesia employed, or in any of a dozen different ways.”¹⁹² It seems as if the court was stating that the taxonomy could have been reasonably organized for its purpose in a variety of ways, and the one such choice made by the plaintiff was therefore creative.

The Second Circuit reached a similar conclusion in deeming Red Book valuations of used cars to be original.¹⁹³ The Red Book “is published eight times a year, in different versions for each of three regions of the United States (as well as a version for the State of Wisconsin),” and “sets forth the editors’ projections of the values for the next six weeks of ‘average’ versions of most of the used cars (up to seven years old) sold in that region. These predicted values are set forth separately for each automobile make, model number, body style, and engine type.”¹⁹⁴ The court reasoned that these valuations were sufficiently creative:

[I]ts valuations were neither reports of historical prices nor mechanical derivations of historical prices or other data. Rather, they represented predictions by the Red Book editors of future prices estimated to cover specified geographic regions. According to [the publisher’s] evidence, these predictions were based not only on a multitude of data sources, but also on professional judgment and

¹⁸⁹ Cf. ROBERT PLANT & STEPHEN MURRELL, AN EXECUTIVE’S GUIDE TO INFORMATION TECHNOLOGY: PRINCIPLES, BUSINESS MODELS, AND TERMINOLOGY 176 (2007) (“If a major city’s . . . telephone directory imposed no useful organization upon the data it contained, . . . it would . . . be entirely useless.”).

¹⁹⁰ Am. Dental Ass’n v. Delta Dental Plans Ass’n, 126 F.3d 977 (7th Cir. 1997).

¹⁹¹ *Id.* at 977.

¹⁹² *Id.* at 979.

¹⁹³ CCC Info. Servs., Inc. v. Maclean Hunter Mkt. Reports, Inc., 44 F.3d 61, 67 (2d Cir. 1994).

¹⁹⁴ *Id.* at 63 (footnote omitted).

expertise. The testimony of one of [the publisher's] deposition witnesses indicated that fifteen considerations are weighed; among the considerations, for example, is a prediction as to how traditional competitor vehicles, as defined by [the publisher], will fare against one another in the marketplace in the coming period.¹⁹⁵

As with the dental taxonomies, the Second Circuit deemed these valuations to be sufficiently creative because of the publisher's choice of one among multiple reasonable ways in which the publisher might have calculated value.¹⁹⁶

This insight suggests why the telephone directory's alphabetical ordering lacked creativity, whereas the dental taxonomy or the used car values arguably do not. When there is a most efficient way of communicating a message (be it with regard to more factually oriented or more expressively oriented works), choosing that way to communicate lacks creativity. Information theory as intertwined with copyright law's goals suggests, then, that such communications should not be roped off for one author alone. Were it otherwise, each subsequent author would have to use slightly more and more inefficient encodings to express a particular message to the detriment of societal comprehension of this message. Conversely, when there is not a most efficient way to communicate a message, choosing one of many possible ways is a creative choice sufficient to meet the originality requirement.

C. *Copyright Infringement*

The previous sections explain important aspects of copyrightability in terms of information theory. This section considers information theory with regard to what constitutes copyright infringement. Subsection 1 addresses why any copying constitutes infringement in the first place, when information theory might seem to counsel otherwise. Subsection 2 then turns to address what copyright law's exclusive rights of reproduction and preparation of derivative works suggest with regard to information theory, and in particular, the restricted circumstances under which authors should hold rights to produce follow-on works to the exclusion of later authors.

¹⁹⁵ *Id.* at 67 (footnote omitted).

¹⁹⁶ *See id.*

1. *Should Any Copying Be Infringement?*

The discussion above of independent creation as a defense against copyright infringement suggests that it is a beneficial part of copyright law: It allows independent and redundant reuse of already existing expression, which broadens the expression's chances of dissemination and preservation.¹⁹⁷ If independently created redundancies of expression are beneficial, then why not take this argument further to suggest that any and all redundancies of expression are beneficial to the expression's dissemination and preservation, all to the advantage of society at large? That is, why should society consider any copying to be copyright infringement in the first place?¹⁹⁸

If one accepts copyright law's utilitarian premises,¹⁹⁹ however, there is good reason to think that copied material should be considered wrongful. Recall that American copyright law is thought to provide incentive to authors to create valuable works that could readily be appropriated without the law.²⁰⁰ Without the incentive of allowing authors to prevent copies, authors might not produce works in the first instance. That means copyright law must forbid copies or copying, at least to the extent that it has a negative societal impact on authors' incentives to create. Even if examining this issue through the lens of information theory suggests it would be helpful to permit redundant use of expression comprehensively, if that would undermine authors' incentives to create in the first instance, there would not be much valuable expression in the first instance for others to reuse.

As discussed above, the utilitarian view of copyright law is about giving incentives to authors to create works that are valuable to society, but not so much incentive that the works are too expensive for society to enjoy, that the works will take too long to pass into the public domain when they can be used freely, or more generally, that the author will have too much control over his works in a way that hurts society.²⁰¹ Seen in this light, information theory can help set the balance between too much and too little incentive: Copyright law ought to encourage the creation, dissemination, and preservation of works that

¹⁹⁷ See *supra* Part IV.B.1.

¹⁹⁸ This question might also be asked in another way: Why have copyright protection in the first instance?

¹⁹⁹ There are many that do not. See, e.g., Rebecca Tushnet, *Economies of Desire: Fair Use and Marketplace Assumptions*, 51 WM. & MARY L. REV. 513 (2009). *But cf.* Fromer, *supra* note 24 (positing that a more expansive understanding of utilitarianism, which extends beyond the pecuniary to incorporate expressive goals as well, would match authors' incentives better).

²⁰⁰ See *supra* Part I.

²⁰¹ See *supra* Part I.

are valuable to society in the ways discussed in Part III. Yet the incentive ought not be so great that other related valuable works do not come into existence at the same time or soon thereafter. Information theory can help establish how to draw this dividing line.

Why specifically draw the line between independently created works—which can be reused—and copied works—which cannot absent some additional defense? From the perspective of information theory, as discussed above, independently created works that are substantially similar to previously existing works are valuable in and of themselves.²⁰² The subsequent independent creator is contributing something he or she thought important to create, stemming in no part from the previously existing work. And by virtue of its independent creation, the work is more likely to contain valuable material. For these reasons, subsequent independent creations give greater heft to the original work by underscoring its vitality in a way that copies do not. Additionally, subsequent work, having been independently created, is more likely to be adding an important message on its own, something that is valuable from the perspective of information theory.²⁰³ For that reason, it would also be less likely than a copy to interfere with the first author's market for his or her work, and thus the copyright incentive. Intentional copies of protected material are less likely to contain an additional important message of value to society and would thus be more likely to interfere with the copyright incentive.²⁰⁴

2. *Substantial Similarity and Derivative Works*

In turning to the specifics of what constitutes infringement, information theory can shed light on two of the most important exclusive rights given to a copyright holder: to reproduce the copyrighted work and to prepare derivative

²⁰² See *supra* Part IV.B.1.

²⁰³ Cf. Fromer, *supra* note 15, at 1493 (“Because problem finding is integral to artistic creativity, copyright law places a greater value on rewarding authors for using their pen to convert their valuable emotional and subjective concepts into an artistic product than on making sure only one problem solution receives the prize of copyright.”).

²⁰⁴ *But cf.*, e.g., Lynne A. Greenberg, *The Art of Appropriation: Puppies, Piracy, and Post-Modernism*, 11 CARDOZO ARTS & ENT. L.J. (1992) (presenting the case that art that appropriates existing works present new messages to society than those existing in the appropriated art). Seen generally, there is another explanation for the line between copied and independently created works. Unlike the independent creator, the copier intentionally copied from the existing work, suggesting that this copier can more readily be deterred by deeming his or her actions to be infringement. This deterrence would in turn give the copyright sufficient value to provide its incentive in the first place.

works.²⁰⁵ Infringement of a copyright holder's reproduction right is found when a defendant actually copied the plaintiff's copyrighted work in a way that rises to the level of an improper misappropriation.²⁰⁶ To determine improper appropriation—particularly of the copyright holder's exclusive right of reproduction—courts ask whether two works are substantially similar.²⁰⁷

Copyright law has not clearly established whether and how the right to prepare derivative works differs from the reproduction right protecting against substantially similar works.²⁰⁸ Some courts have intimated that the two rights are more or less equivalent,²⁰⁹ while others have found the rights to be at least partially distinct.²¹⁰

Either way, both substantial similarity with regard to the reproduction right and the right to prepare derivative works are keyed to using critical—and protected²¹¹—aspects of a copyrighted work to create another. Copyright law

²⁰⁵ 17 U.S.C. § 106(1)–(2) (2012).

²⁰⁶ See, e.g., *Boisson v. Banian, Ltd.*, 273 F.3d 262, 267–68 (2d Cir. 2001).

²⁰⁷ The Second and Ninth Circuits—the major courts deciding copyright cases—take different approaches to this question. The Second Circuit generally assesses the substantial similarity of an allegedly infringing work to a copyrighted work by asking “if the ordinary observer, unless he set out to detect the disparities, would be disposed to overlook them, and regard their aesthetic appeal as the same.” *Id.* at 272 (internal quotation marks omitted). By contrast, the Ninth Circuit applies both extrinsic and intrinsic tests, and only if both show similarity is there infringement. *Cavalier v. Random House, Inc.*, 297 F.3d 815, 822 (9th Cir. 2002). The extrinsic test makes an objective comparison between the two works at issue with regard to their expressive elements, such as “articulable similarities between the plot, themes, dialogue, mood, setting, pace, characters, and sequence of events.” *Id.* (quoting *Kouf v. Walt Disney Pictures & Television*, 16 F.3d 1042, 1045 (9th Cir. 1994)). The “intrinsic test” is a “subjective comparison that focuses on whether the ordinary, reasonable audience would find the works substantially similar in the total concept and feel of the works.” *Id.* (internal quotation marks omitted). For more on whose perspective ought to be used to assess substantial similarity, see Jeanne C. Fromer & Mark A. Lemley, *The Audience in Intellectual Property Infringement*, 112 MICH. L. REV. 1251 (2014).

²⁰⁸ The copyright statute defines a “derivative work” broadly as any “work based upon one or more preexisting works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which a work may be recast, transformed, or adapted.” 17 U.S.C. § 101 (2012).

²⁰⁹ See, e.g., *Castle Rock Entm't, Inc. v. Carol Publ'g Grp., Inc.*, 150 F.3d 132, 143 n.9 (2d Cir. 1998) (“[I]f the secondary work sufficiently transforms the expression of the original work such that the two works cease to be substantially similar, then the secondary work is not a derivative work and, for that matter, does not infringe the copyright of the original work.”).

²¹⁰ See, e.g., *Warner Bros. Entm't Inc. v. RDR Books*, 575 F. Supp. 2d 513, 538–39 (S.D.N.Y. 2008) (ruling that a work is derivative of another only if it still represents the original work of authorship, but even if not, the works might nonetheless be substantially similar). See generally Pamela Samuelson, *The Quest for a Sound Conception of Copyright's Derivative Work Right*, 101 GEO. L.J. 1505 (2013) (making the case that the right to prepare derivative works is narrower in scope than is conventionally believed).

²¹¹ That would exclude any borrowing of ideas, facts, stock elements, and other unprotected material from a copyrighted work. See *supra* Part IV.A–B.

places the exclusive right to make this use with the copyright owner and deems such uses made by third parties without permission to be infringing.²¹² In so doing, copyright law appears to be encouraging authors to produce new works that relate in some significant way to—oftentimes reusing and building on core aspects of—their preexisting copyrighted work.²¹³ For example, Sylvester Stallone, copyright owner of the first *Rocky* movie, might be encouraged by his exclusive rights in the movie to make a sequel to the movie incorporating similar storylines and themes of redemption and many of the same characters with their particular interconnections, albeit in a different context.²¹⁴ Or copyright law's exclusive rights might galvanize a copyright holder in a book to produce a film based on that book.

Moreover, if the copyright holders do not want to produce these follow-on works themselves, copyright law nonetheless confers the exclusive rights to do so upon them, permitting them to license others to create these works (or refrain from any such licensing altogether).²¹⁵

Information theory provokes questions as to the value of providing this encouragement at all, and furthermore, who ought to be given this encouragement: the owner of copyright in the preexisting work or third parties who would like to create these follow-on works?

Seen through the lens of information theory, the exclusive rights of reproduction and preparation of derivative works encourage authors to take key aspects of their message in their preexisting work and make them redundant by encoding them (or recognizable variations of them) in new works. Spreading an author's message redundantly across a number of works increases the odds that the message will be decoded properly by consumers. In doing so, these exclusive rights thus help fulfill the basic aims of copyright law, of promoting the production and dissemination of factual, systematic, and cultural knowledge and of valuable expression.

²¹² See 17 U.S.C. § 106 (2012).

²¹³ See Paul Goldstein, *Derivative Rights and Derivative Works in Copyright*, 30 J. COPYRIGHT SOC'Y U.S.A. 209 (1983). Another theory suggests that the right to prepare derivative works is important to firms' learning about human capital inputs and the reduction of shirking in creative teams. See Anthony J. Casey & Andres Sawicki, *Copyright in Teams*, 80 U. CHI. L. REV. 1683 (2013).

²¹⁴ Cf. *Anderson v. Stallone*, No. 87-0592 WDKGX, 1989 WL 206431 (C.D. Cal. Apr. 25, 1989) (finding a third party's script for a sequel to *Rocky III* to be an unauthorized derivative work, to which the copyright holder in *Rocky I*, *Rocky II*, and *Rocky III* had exclusive rights).

²¹⁵ See Kelly Casey Mullally, *Blocking Copyrights Revisited*, 37 COLUM. J.L. & ARTS 57, 75 (2013).

That said, this analysis evokes a question, one which occupies copyright scholars generally: Why vest these exclusive rights in the author rather than letting a third party create (and perhaps also receive rights in) substantially similar or derivative works?²¹⁶ Were it otherwise, a third party would be encouraged to create follow-on works using key components of the author's preexisting work. And that might be preferable to the current state of law, in that the world would not have to rely solely on the original author to create or license others to create valuable follow-on works but could rely more broadly on any parties that might want to and be able to produce valuable follow-on works. In fact, patent law pursues such a model, allowing both the original inventor and third parties to invent and patent improvements that might fall within the scope of preexisting patent rights.²¹⁷

This important question can be addressed by examining this issue through the lens of information theory and asking who will do the best job of transmitting and disseminating the existing work's message—be it the underlying knowledge or valuable expression in and of itself—in these follow-on works. Vesting these exclusive rights solely in the existing work's author can be helpful on the ground that the creator arguably knows better than anyone else (non-noisily) the message that he or she sought to communicate (probably noisily) in the original work. Giving strong incentive to the original creator to develop follow-on works using important aspects of the original message is helpful to increase the odds that the message will be communicated effectively to society. Third parties do not have the same advantage. Similarly, when the author does not feel well-placed to develop a follow-on work (such

²¹⁶ See, e.g., Amy B. Cohen, *When Does a Work Infringe the Derivative Works Right of a Copyright Owner?*, 17 *CARDOZO ARTS & ENT. L.J.* 623 (1999) (suggesting that there are competing interests in the public and the original author with regard to the right to prepare derivative works, and that current copyright law does not strike the balance correctly); Goldstein, *supra* note 213 (expressing sympathy for authors possessing rights in derivative works, while also seeing the competing interest of allowing third parties to sometimes create follow-on works); Lemley, *supra* note 18 (arguing that copyright law should do away with this rule permitting only the original author to “improve on” the copyrighted work); Glynn S. Lunney, Jr., *Copyright, Derivative Works, and the Economics of Complements*, 12 *VAND. J. ENT. & TECH. L.* 779, 779 (2010) (“Where a given use, reuse, or product is a strong complement to a copyrighted work, and would, in the absence of copyright’s intervention, be available in a naturally competitive market, the copyright owner should not have the exclusive right to control such a use, reuse, or product.”); cf. Kindra Deneau, *The Historical Development and Misplaced Justification for the Derivative Work Right*, 19 *B.U. J. SCI. & TECH. L.* 68, 68 (2013) (“[I]njunctive relief for infringement is inappropriate when a derivative author has contributed substantial new creativity relative to the portion of the preexisting work infringed.”).

²¹⁷ See Lemley, *supra* note 18, at 1052. Nonetheless, the original patentee cannot make the improved invention without getting a license from the improver. See *id.* (discussing patent law’s resulting situation of “blocking patents”).

as translating the work into a language that the author does not speak or adapting a book to be a movie), the author can choose a particular person that would do the best job of communicating the author's message in the follow-on work and license that person to do it. More generally, for some categories of works, we might conclude that the author of the existing work is generally best placed to be creating valuably redundant follow-on works.

That said, there are categories of works for which the author is likely not best placed to communicate (or choose someone to communicate) the underlying message in follow-on works. One example might be existing works that society has come to understand quite differently than—and perhaps at cross purposes to—the original author's intended message. In this case, we want to make sure that third parties can create follow-on works that carry on the meaning that society has come to accept. Or in a different scenario, when the original author chooses not to create follow-on works for which society clamors, we might want to allow third parties to do so, either by permitting them to do so freely or through a compulsory-licensing scheme, unless there is a good reason to think that follow-on works will damage the existing work's message. Moreover, when a third party seeks to use a preexisting work to criticize it or purposely change its accepted meaning, the author does not occupy a privileged position as communicator (and in fact likely holds a weaker one than the third party as communicator).²¹⁸ There are likely yet more categories of works for which the author is similarly no better—and might be worse—placed than third parties to create follow-on works.

Information theory thus helps shed analytical light on why it is helpful to encourage substantially similar and derivative works as follow-on works to existing ones, and when to give rights to the follow-on works to the authors of the existing works and when to withhold them in favor of third parties.

D. Fair Use

Connected to this analysis on whether rights to prepare follow-on or related works ought to lie with the author of the underlying work is the issue of fair use in copyright law. Copyright law excuses some instances that would otherwise be infringing conduct when a third party's use of someone else's copyrighted work is deemed to be fair.²¹⁹ The copyright statute illustrates some

²¹⁸ This category connects strongly to copyright law's fair use defense, as explored in the following section.

²¹⁹ See 17 U.S.C. § 107 (2012).

such instances: “criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research.”²²⁰ A set of (nonexclusive) statutory factors must be analyzed to determine whether a particular use is fair: “the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes,” “the nature of the copyrighted work,” “the amount and substantiality of the portion used in relation to the copyrighted work as a whole,” and “the effect of the use upon the potential market for or value of the copyrighted work.”²²¹

A “fair use” carve-out has numerous justifications. Most relevantly, the fair use doctrine can stimulate the production of creative works for public consumption without undercutting the value of the original copyrighted work.²²² It does so by enabling third parties to create culturally valuable works that must borrow from the original work in some capacity in order to succeed, often transforming it.²²³ As the statute suggests,²²⁴ news reporting, critical reviews, and parodies are some prototypical examples.²²⁵ A second, partially related argument set forth by Wendy Gordon is that “fair use [ought] to permit uncompensated transfers that are socially desirable but not capable of effectuation through the market.”²²⁶ Examples include parodies that might cast an unfavorable light on an original work or uses for which the transaction costs are too great for the copyright owner to agree to a licensing arrangement.²²⁷

With these purposes in mind, fair use doctrine will sometimes allow for the reuse of key (or unimportant) pieces of knowledge encoded in others’ copyrighted works. Notwithstanding the expense of litigating fair use²²⁸ and some unpredictability in the doctrine,²²⁹ the doctrine allows this reuse in a way

²²⁰ *Id.*

²²¹ *Id.*

²²² See *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 577 (1994).

²²³ See, e.g., Pierre N. Leval, *Toward a Fair Use Standard*, 103 HARV. L. REV. 1105 (1990).

²²⁴ See *supra* text accompanying note 220.

²²⁵ See *Campbell*, 510 U.S. at 578–85 (parody); *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 561 (1985) (news reporting); *Sundeman v. Seajay Soc’y, Inc.*, 142 F.3d 194, 206 (4th Cir. 1998) (critical review).

²²⁶ Wendy J. Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and Its Predecessors*, 82 COLUM. L. REV. 1600, 1601 (1982). A third argument—less relevant here—is grounded in technology, such as allowing the intermediate copying of copyrighted software code to make a program that is interoperable with a preexisting computer or gaming system. See, e.g., *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d 1510, 1513–14 (9th Cir. 1992).

²²⁷ See Gordon, *supra* note 226.

²²⁸ E.g., Shyamkrishna Balganesh, *Copyright Infringement Markets*, 113 COLUM. L. REV. 2277 (2013).

²²⁹ Compare, e.g., Deidré A. Keller, *Recognizing the Derivative Works Right as a Moral Right: A Case Comparison and Proposal*, 63 CASE W. RES. L. REV. 511, 546 (2012) (calling fair-use doctrine “notoriously

that might ease some concerns about third parties' inability otherwise to create follow-on works to works in which they have no copyright permissions.²³⁰ In addition to the exemplary statutory categories of favored uses,²³¹ courts typically deem parodies—expressive works that borrow expression from a preexisting work as a way to comment on that work—to be fair use.²³² These important categories—journalism and news reporting, parody, criticism and commentary, and scholarship and research—coincide with the types of uses information theory would want to promote.

With regard to criticism, commentary, scholarship, and research, expression copied for these sorts of works is principally done either to excavate the knowledge therein or to contribute to a conversation on the expression's meaning. Both purposes are precisely those that coincide with what makes expression valuable to society.²³³ Given that the original author is unlikely to supply these sorts of follow-on works and that the contribution is important for society, it is sensible for copyright law to allow third parties to copy the otherwise protected expression or the knowledge contained therein. These third parties employ that expression redundantly to generate a greater conversation about the expression than the author might contribute on his or her own.²³⁴ And when this copying is done to lay bare the knowledge contained therein, it seeks to cut through the noisy expression in which knowledge is usually embedded to transmit that knowledge efficiently to the public at large.²³⁵

unpredictable”), and Jason Mazzone, *Administering Fair Use*, 51 WM. & MARY L. REV. 395, 398 (2009) (“[N]either Congress in enacting the fair use law nor the courts in applying it have supplied sufficiently clear guidance to permit individuals who wish to make use of a copyrighted work—and who also desire to avoid infringing the copyright in the work—to determine whether their proposed use is fair.”), with Michael J. Madison, *A Pattern-Oriented Approach to Fair Use*, 45 WM. & MARY L. REV. 1525, 1530 (2004) (“[A] pattern-oriented approach offers . . . a coherent method of analyzing fair use questions generally.”), and Pamela Samuelson, *Unbundling Fair Uses*, 77 FORDHAM L. REV. 2537, 2541 (2009) (“[F]air use law is both more coherent and more predictable than many commentators have perceived once one recognizes that fair use cases tend to fall into . . . policy-relevant clusters.”).

²³⁰ See *supra* text accompanying notes 216–18. But cf. James Gibson, *Risk Aversion and Rights Accretion in Intellectual Property Law*, 116 YALE L.J. 882, 884 (2007) (emphasizing how risk aversion strongly deters those who might have a strong fair-use defense from making it).

²³¹ See *supra* text accompanying note 220.

²³² See *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 580–83 (1994); see also Madison, *supra* note 229; Samuelson, *supra* note 229.

²³³ *Supra* Part III.

²³⁴ See *supra* Part III.B.

²³⁵ See *supra* Part III.A.

A similar explanation can be provided for parodies. For example, the U.S. Supreme Court held that a bawdy rap version of Roy Orbison's song, "Oh, Pretty Woman" could reasonably be construed as a parody—deserving of fair use protection—because it could "be taken as a comment on the naiveté of the original of an earlier day, as a rejection of its sentiment that ignores the ugliness of street life and the debasement that it signifies."²³⁶ The principal reason the Supreme Court thought that parody would frequently be deserving of fair use protection is that it transforms the original expression into something new and valuable.²³⁷ That is, parody borrows expression that is valuable in and of itself and confers on that expression new meanings, frequently and deliberately muddying the original author's viewpoint.²³⁸ Parodists reuse the expression from the underlying work somewhat redundantly in a way that continues a conversation on the meaning of the expression itself. The use of the underlying expression thus creates a direct link to a valuably noisy conversation on the expression and ought to be encouraged.²³⁹ Given that authors are unlikely to critique their own work in this way and that the noisy conversation about an expression's meaning and value ought to have many voices, it is sensible to permit third parties to make such uses. A similar argument can be made about promoting other transformative uses, such as some satire (expressive works that borrow expression from a preexisting work not to comment on that work but merely to create a new work with its own different meaning).²⁴⁰

In fact, in a recent fair use case about appropriation art, the Second Circuit seemed to adopt this view, stating that it would not look merely to a defendant's explanation of his reuse of a copyright owner's expression in his artwork but more capaciously to "how the artworks may 'reasonably be perceived' in order to assess their transformative nature."²⁴¹ Much like the

²³⁶ *Campbell*, 510 U.S. at 583.

²³⁷ *See id.*

²³⁸ *See* Heymann, *supra* note 127, at 449 ("[T]he relevant question should be the *degree* of transformativeness—the amount of interpretive distance that the defendant's use of the plaintiff's work creates. If that distance is significant enough to create a distinct and separate discursive community around the second work, the defendant's use is more likely to be transformative (and, perhaps, fair.); *see also* ROBERT CHAMBERS, *PARODY: THE ART THAT PLAYS WITH ART* (2010); MARGARET A. ROSE, *PARODY: ANCIENT, MODERN, AND POST-MODERN* (1993).

²³⁹ *See supra* Part III.B.

²⁴⁰ *But cf. Campbell*, 510 U.S. at 581 ("[S]atire can stand on its own two feet and so requires justification for the very act of borrowing.").

²⁴¹ *Cariou v. Prince*, 714 F.3d 694, 707 (2d Cir. 2013) (quoting *Campbell*, 510 U.S. at 582).

theory of reader-response criticism discussed above,²⁴² the Second Circuit emphasized the importance of seeing whether consumers are reacting to the defendant's expression in a way that is distinct from their reaction to the same expression by the copyright owner.²⁴³

Information theory also suggests why the transformative use of expression in journalism and news reporting ought to be considered fair. For journalism and news reporting, much expression copied therein—no matter how expressive it is—is for use as fact in a news story. As one example, when a Puerto Rican newspaper published a photographer's nearly nude photographs of a Miss Puerto Rico Universe winner in response to a growing scandal about her fitness to retain her crown in light of these photographs, the First Circuit deemed that use to be fair.²⁴⁴ Even if the photographs were otherwise expressive, they were being used factually to tell the story about the controversy.²⁴⁵ Reuse of expression to convey knowledge (or fact) is transformative, and from the perspective of information theory, it ought to be encouraged. By contrast, when a third party reuses someone's particular expression of news itself—a reuse that involves no transformation but mere co-optation, as when *The Nation* magazine took important parts of Gerald Ford's unpublished autobiography and published them as a news story²⁴⁶—the interest in permitting the reiteration of the same expression for the same purpose as the original is diminished.

E. Dissemination

Until now, the discussion of how information theory informs copyright law has focused primarily on encouraging the creation of particular sorts of valuable expression and redundancies. Yet for this expression to have any value, it must reach the public so that it can consume these works, enjoy these works, and create additional works based on valuable aspects of these

²⁴² See *supra* text accompanying notes 124–28.

²⁴³ Cf. Heymann, *supra* note 127, at 449 (“The focus [in a fair use inquiry ought not to be] on the author’s intent (although, like any statement of authorial interpretation, intent may be relevant evidence) but on the reader’s reaction.”).

²⁴⁴ *Núñez v. Caribbean Int’l News Corp.*, 235 F.3d 18, 25 (1st Cir. 2000).

²⁴⁵ See *id.* at 22 (“[The newspaper] reprinted the pictures not just to entice the buying public, but to place its news articles in context; as the district court pointed out, ‘the pictures were the story.’”).

²⁴⁶ *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 542 (1985).

works.²⁴⁷ Encouraging broad dissemination is thus central to copyright law.²⁴⁸ This section discusses copyright's major efforts to promote dissemination: the first sale doctrine, secondary liability, and treatment of search engines.

1. *First Sale Doctrine*

A major way in which copyright law promotes the broad dissemination of valuable works is through its first sale doctrine, a limitation on copyright holders' rights. Specifically, copyright owners have the exclusive right to distribute their works.²⁴⁹ This right is thought to encourage a copyright holder to distribute its works and to stop others from interfering with its marketplace in this regard.²⁵⁰ Accompanying this broad right of distribution, however, is a fear that copyright owners will use this distribution right to curtail widespread dissemination, except perhaps at prohibitive prices.²⁵¹

Enter copyright's first sale doctrine, which statutorily provides that, notwithstanding the copyright owner's distribution right, "the owner of a particular copy . . . lawfully made under [the copyright laws] . . . is entitled, without the authority of the copyright owner, to sell or otherwise dispose of the possession of that copy."²⁵² This doctrine serves to counterbalance the copyright holder's rights by permitting owners of copies of the copyright holder's work to dispose of it as they wish.²⁵³ As Aaron Perzanowski and Jason Schultz explain, the doctrine "improves both the affordability and availability of copyrighted works by fostering secondary markets for lawful copies and distribution models that operate outside of copyright holder control."²⁵⁴ Because secondhand copies of works frequently compete with new copies, the secondary market's existence encourages copyright holders to keep

²⁴⁷ See Jenny Lynn Sheridan, *Copyright's Knowledge Principle* (Drexel Univ. Earle Mack Sch. of Law, Research Paper No. 2013-W-02, 2013), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2341211 ("[A]ccess to existing knowledge is a necessary condition for the creation of new knowledge.").

²⁴⁸ See *Golan v. Holder*, 132 S. Ct. 873, 888–89 (2012) (noting that copyright law's constitutional purpose is to encourage both creation and dissemination).

²⁴⁹ 17 U.S.C. § 106(3) (2012).

²⁵⁰ See *Eldred v. Ashcroft*, 537 U.S. 186, 206–07 (2003).

²⁵¹ See, e.g., *Bobbs-Merrill Co. v. Straus*, 210 U.S. 339, 350–51 (1908) (originating the first sale doctrine as a response to a copyright holder's attempt to prohibit downstream sales of a book for less than one dollar, a more considerable sum at the time).

²⁵² 17 U.S.C. § 109(a) (2012). There are a few exceptions, such as the rental of sound recordings or computer programs. *Id.* § 109(b)(1)(A).

²⁵³ See Aaron Perzanowski & Jason Schultz, *Digital Exhaustion*, 58 UCLA L. REV. 889, 892 (2011).

²⁵⁴ *Id.* at 894; accord R. Anthony Reese, *The First Sale Doctrine in the Era of Digital Networks*, 44 B.C. L. REV. 577, 585–610 (2003).

the prices for new copies reasonable.²⁵⁵ That is, not only does the first sale doctrine increase the dissemination of works by creating lawful secondary markets for them, but it also increases works' dissemination by pressuring copyright holders to lower their prices for new copies. In fact, the secondary market promotes price discrimination in a way that expands the dissemination of copyrighted works: those who are not willing to pay the copyright holder's price are sometimes willing to pay the secondary market price for the same, albeit used, work.²⁵⁶ The first sale doctrine thus seems to work hand in hand with the copyright owner's distribution right to maximize access to copyrighted works.

The doctrine, however, does have limitations, which seem to undermine its effect of encouraging broad dissemination of works. Most prominently, it has no application to digital copyrighted works, because in order to transfer a purchased digital copy, one technically has to make a copy of it (even if the transferor deletes that copy after transferring).²⁵⁷ This copy is infringing because the first sale doctrine serves merely as a limitation on the copyright holder's distribution right, not the reproduction right.²⁵⁸ In addition, the doctrine does not cover copyrighted works that a third party has license to use but does not own, which is typical of many digital works.²⁵⁹ If broad dissemination is a goal in the digital age, for the reasons just discussed, it is advisable to find a way to shelter secondary transfers of digital works and licensed works.²⁶⁰

2. *Secondary Liability*

Secondary liability rules provide copyright law with another avenue to promote broad dissemination. As discussed herein, those who help proliferate copies of copyrighted works without the copyright holder's authorization—by contributing to infringement, inducing infringement, or acting on behalf of an infringer—might be found secondarily liable for copyright infringement. From the perspective of information theory, copyright law ought to chart a course

²⁵⁵ See Perzanowski & Schultz, *supra* note 253, at 894.

²⁵⁶ See *id.* at 894–95. Additionally, the first sale doctrine is thought to improve copyright law's goal of preserving works, specifically by increasing the availability of out-of-print works. See *id.* at 895; *infra* Part IV.F (discussing how information theory suggests how copyright law ought to preserve works).

²⁵⁷ See Perzanowski & Schultz, *supra* note 253, at 902.

²⁵⁸ See *id.*

²⁵⁹ See *id.* at 901–02; Sheridan, *supra* note 247.

²⁶⁰ See Perzanowski & Schultz, *supra* note 253, at 892 (advocating a regime of digital exhaustion to protect secondary transfers of digital works).

between stopping those who help infringement happen and promoting the efforts of those that make transmissions of these valuable works more direct and therefore less noisy.²⁶¹

Copyright law imposes secondary liability in at least three different contexts. First, it penalizes those who contribute sufficiently to another's copyright infringement, so long as they have knowledge of the infringement and materially contribute to the infringing activity.²⁶² The Supreme Court has ruled, however, that the law will not impute knowledge of infringement to a device manufacturer if the device can be used to infringe, so long as the device is capable of substantial noninfringing uses.²⁶³ Second, the law forbids vicarious copyright infringement, which occurs when the party has the right and ability to supervise the infringing activity and also has a direct financial interest in it.²⁶⁴ Third, one who intentionally induces infringement will be deemed secondarily liable.²⁶⁵ All in all, these doctrines seek to punish—and thus deter—those who would sufficiently help another engage in copyright infringement.²⁶⁶

If not carefully constructed, however, secondary liability can over-deter those who act as conduits for material—including copyrighted material—from creating and distributing innovative dissemination platforms, to the detriment of copyright law's goals.²⁶⁷ Information theory can indicate how to walk the fine line the Supreme Court has acknowledged in secondary liability doctrine between “supporting creative pursuits through copyright protection and promoting innovation in new communication technologies by limiting the incidence of liability for copyright infringement.”²⁶⁸ Most pertinently, some preference against secondary liability ought to be given to those third parties that innovate ways to disseminate materials—including copyrighted materials—in ways that reduce undesirable noise that might otherwise surround these materials.

²⁶¹ Cf. Timothy Wu, *Copyright's Communications Policy*, 103 MICH. L. REV. 278, 279–81 (2004) (shining a light on “copyright’s poorly understood role in regulating competition among rival disseminators” by focusing on copyright’s compulsory licenses, safe harbors from infringement, and antipiracy rules).

²⁶² See *Fonovisa, Inc. v. Cherry Auction, Inc.*, 76 F.3d 259, 264 (9th Cir. 1996).

²⁶³ *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 442 (1984).

²⁶⁴ See *Fonovisa*, 76 F.3d at 261–62.

²⁶⁵ See *Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd.*, 545 U.S. 913, 930 (2005).

²⁶⁶ See *supra* Part IV.C.1 (analyzing why copyright infringement ought to be forbidden).

²⁶⁷ See *Grokster*, 545 U.S. at 928.

²⁶⁸ *Id.*

There are at least two ways in which devices or software programs could help disseminate copyrighted materials in ways that diminish undesirable forms of noise. For one thing, in today's age of information overload, filters, knowledge curation, and other organizing mechanisms are ever important.²⁶⁹ From this perspective, those who provide a forum for others to organize or discuss their knowledge—about particular factual topics or about expression—ought to be encouraged for providing effective ways to manage noisy floods of expression. Additionally, those entities that devise technologies that make transmissions of valuable expression to consumers more direct—particularly when the copyright holders of that expression seek to prevent these direct transmissions, as the music industry did for a long time with regard to digital distribution²⁷⁰—ought to be encouraged. Examples of such favored uses include devices permitting television viewers to record programs that are freely available to them at specified times for later viewing at a convenient time,²⁷¹ and software programs enabling the digital distribution of music.²⁷² Particularly, when copyright holders can provide these dissemination benefits and choose not to do so, society is hurt if others cannot provide them instead.²⁷³ Encouraging others to do so—particularly within a murky regime of secondary liability—might furthermore encourage copyright holders to negotiate mutually beneficial agreements with these innovative entities.

3. Search Engines

In this vein, the search engine is one of the most important organizers and filters of information in the digital age. In a number of recent infringement suits against search engines—frequently Google—for indexing copyrighted works, courts have held that there was no liability largely because the indexing enabled widespread and easy online access to these works.²⁷⁴ In one such case against Google for displaying thumbnail representations of copyrighted images placed on the Internet by others, the Ninth Circuit determined that Google's use was fair.²⁷⁵ In coming to its conclusion, the court relied heavily on the fact

²⁶⁹ See GLEICK, *supra* note 26, at 401–12.

²⁷⁰ See Ariel B. Taitz, Note, *Removing Road Blocks Along the Information Superhighway: Facilitating the Dissemination of New Technology by Changing the Law of Contributory Copyright Infringement*, 64 GEO. WASH. L. REV. 133 (1995).

²⁷¹ See *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 442–56 (1984).

²⁷² See *Grokster*, 545 U.S. at 919–20.

²⁷³ See, e.g., Taitz, *supra* note 270, at 135–36.

²⁷⁴ See, e.g., *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146 (9th Cir. 2007).

²⁷⁵ See *id.* at 1163–68.

that Google's use of the thumbnails was transformative.²⁷⁶ The Ninth Circuit reasoned that "[a]lthough [a copyrighted] image may have been created originally to serve an entertainment, aesthetic, or informative function, a search engine transforms the image into a pointer directing a user to a source of information."²⁷⁷ The court relied on this fact to conclude that the use was fair because by putting the copyrighted images "to a use fundamentally different than the use intended by" the copyright owner—namely Google Image Search, to locate online images based on users' queries—"Google has provided a significant benefit to the public."²⁷⁸

Similarly, in another infringement case by authors against Google for its storage and indexing of digital versions of the authors' books, the Southern District of New York determined that Google's use was fair.²⁷⁹ The court came to its conclusion in large part because Google's use of the books was highly transformative because it "digitizes books and transforms expressive text into a comprehensive word index that helps readers, scholars, researchers, and others find books."²⁸⁰ In response to user queries, Google Book Search would display short snippets of text from the books that matched those queries.²⁸¹ The court thought that "the snippets help users locate books and determine whether they may be of interest," thereby "us[ing] words for a different purpose—it uses snippets of text to act as pointers directing users to a broad selection of books."²⁸² Moreover, the court relied on the fact that Google Book Search provided greater access to books, even increasing authors' book sales:

[A] reasonable factfinder could only find that Google Books enhances the sales of books to the benefit of copyright holders. An important factor in the success of an individual title is whether it is discovered—whether potential readers learn of its existence. Google Books provides a way for authors' works to become noticed, much like traditional in-store book displays. Indeed, both librarians and their patrons use Google Books to identify books to purchase. Many authors have noted that online browsing in general and Google Books in particular helps readers find their work, thus increasing their audiences. Further, Google provides convenient links to booksellers to make it easy for a reader to order a book. In this day and age of

²⁷⁶ *Id.* at 1165.

²⁷⁷ *Id.*

²⁷⁸ *Id.* at 1168.

²⁷⁹ *See* Authors Guild, Inc. v. Google Inc., 954 F. Supp. 2d 282, 294 (S.D.N.Y. 2013).

²⁸⁰ *Id.* at 291.

²⁸¹ *Id.*

²⁸² *Id.*

online shopping, there can be no doubt but that Google Books improves books sales.²⁸³

In holding that Google was not liable for copyright infringement, the courts in both of these cases relied heavily on the increased access the public has to copyrighted works by virtue of search engines making them accessible in response to user queries. The courts thus gave permission to those who would invest the efforts to digitize and index copyrighted content on a mass scale in a way that benefits the public. In doing so, the courts put a thumb on the scale in favor of broader access to, and thus dissemination of, copyrighted materials. These search engines allow users to navigate through a glut of information to locate precisely the information they are seeking, effectively disseminating this information further and wider to those who find it of value.

In sum, an information theory of copyright law appreciates the importance of dissemination of valuable works. Analysis of a number of important areas in copyright law—the first sale doctrine, secondary liability, and the particular treatment of search engines—can help chart a course as to when to provide rights to the copyright holder to encourage dissemination and when to withhold them so that third parties can better encourage dissemination.²⁸⁴

F. Preservation

Generating and disseminating expression that is both valuable for the knowledge it contains and for the noisy conversation it generates about the expression itself is helpful over a span of time only if the valuable expression itself is preserved.²⁸⁵ Copyright law encourages this preservation in at least two important ways: by giving authors a copyright interest in their work as it would appear across multiple formats—rather than just the one in which the work

²⁸³ *Id.* at 293 (citations omitted).

²⁸⁴ Another important area of copyright law that illustrates this issue is the Digital Millennium Copyright Act's anti-circumvention provisions, which generally forbid the making of and trafficking in devices that circumvent technological protection measures that copyright holders install to protect their copyrighted works from being infringed. 17 U.S.C. § 1201 (2012). Congress adopted these measures to provide copyright owners with an additional measure of security from infringement, particularly in the digital age. *Universal City Studios, Inc. v. Corley*, 273 F.3d 429, 435 (2d Cir. 2001). Nonetheless, many courts have understood these provisions to prevent circumvention even if it were designed to access what would be a fair use of a copyrighted work. *E.g.*, *Universal City Studios, Inc. v. Reimerdes*, 111 F. Supp. 2d 294, 323–24 (S.D.N.Y. 2000). Given the centrality of fair use to promoting copyright law's goals—particularly from the perspective of information theory—the lack of a general fair use exemption to these provisions is troublesome.

²⁸⁵ *Cf.* R. Anthony Reese, *What Copyright Owes the Future*, 50 HOUS. L. REV. 287, 288, 316 (2012) (proposing that the law should require that all copyrighted—or once copyrighted—works be preserved, so that the works will be available after they enter the public domain).

first appeared—and by encouraging deposit of works in the central collection of the Library of Congress. The following sections discuss each in turn.

1. *Works in Multiple Formats*

Courts declare copyright law to be media neutral, particularly in the context of finding infringement when a work is produced in another medium than the form in which the copyrighted work has already appeared.²⁸⁶ When according rights in a particular work in multiple formats to the copyright holder, copyright law gives copyright holders incentive to reproduce their work in various formats, including those new ones that might arise during the long duration of copyright (such as digital formats). Publishing works in new formats as old ones fall out of use helps ensure that valuable works are preserved. And when copyright owners are not taking care to move their works from obsolete formats to new ones—such as with volatile film stock, fragile manuscripts, or software in outdated formats—copyright law typically sanctions third parties' reproduction of these works in newer formats as fair use.²⁸⁷ Copyright law's rules in this regard might help serve this purpose of preservation.

In some contexts, copyright law doles out rights to closely connected works to different individuals, which can also serve the purpose of preservation by giving multiple stakeholders an interest in preservation. The most direct example occurs with regard to music. Copyright law provides protection for both musical compositions and sound recordings.²⁸⁸ Thus, for any recorded song, copyright will vest in the songwriter and also in the song performer. By issuing layered protection to closely connected works, copyright law encourages redundancies in varying formats, such as sheet music and sound recordings, and doing so helps ensure that these musical works will survive in

²⁸⁶ See, e.g., *Rand McNally & Co. v. Fleet Mgmt. Sys., Inc.*, 600 F. Supp. 933, 942 (N.D. Ill. 1984) (“[I]t is clear that inputting a copyrighted work into a computer would violate the copyright holder’s exclusive rights. . . . The way that information must be formatted in order to be of use by a particular computer or program should not prevent a finding of infringement.”).

²⁸⁷ See, e.g., *Sundeman v. Seajay Soc’y, Inc.*, 142 F.3d 194, 203 (4th Cir. 1998) (holding that a library’s reproduction of an unpublished, “fragile, seventy year old original manuscript” was fair use, and “unquestionably served the ‘public benefit’ and the ‘development of art’”); H.R. REP. NO. 94-1476, at 73 (1976) (“The efforts of the Library of Congress, the American Film Institute, and other organizations to rescue and preserve [volatile film stock,] this irreplaceable contribution to our cultural life are to be applauded, and the making of duplicate copies for purposes of archival preservation certainly falls within the scope of ‘fair use.’”); cf. 17 U.S.C. § 108(c) (2012) (conferring on libraries limited rights of reproduction when a work in their possession is in an obsolete format).

²⁸⁸ 17 U.S.C. § 102(a)(2), (7).

one form or another. It also helps them be consumed by different audiences. For music, there might be those that prefer to read sheet music and those that prefer to listen to sound recordings.²⁸⁹

Another context in which the issue of works in multiple formats arises is copyright licensing. Given the extremely long period of copyright protection,²⁹⁰ there arise disputes between licensees and copyright holders as to which of them has rights to a work in a format that was not yet in existence at the time of licensing. For example, the Walt Disney Company had a licensing dispute with the copyright owner of Igor Stravinsky's composition for *The Rite of Spring* over whether Disney had a right to use the composition in a video release of its *Fantasia* film.²⁹¹ In 1939, the two parties had entered into a licensing agreement that gave Disney particular rights to the composition in "one motion picture," which they used in their 1940 release of *Fantasia*.²⁹² At the time, there were no devices to watch movies at home and nothing in the contract discussed them. But they became popular many decades later while the licensing agreement was still in place. Once that happened, Disney released *Fantasia* (containing the Stravinsky composition) on video, and the copyright owner objected on the basis that it was not covered by the license.²⁹³ Chances are that neither party contemplated the video format at the time of contracting, but the court had to rule whether it was covered under the license. The Second Circuit determined that the rights lay with Disney because "[t]he words of Disney's license are more reasonably read to include than to exclude a motion picture distributed in video format."²⁹⁴

In another example, a number of literary authors had entered into book publishing contracts with Random House well before electronic books were a plausible technology.²⁹⁵ When e-books eventually came on the scene, some of these authors conferred the right to release their books electronically on another company, Rosetta Books.²⁹⁶ Random House sought to enjoin Rosetta Books from selling these books as e-books on the basis that Random House

²⁸⁹ Cf. *supra* note 165 and accompanying text (emphasizing that different individuals absorb information better in different formats).

²⁹⁰ As noted above, copyright protection generally lasts for the lifetime of the author plus seventy years. See *supra* text accompanying note 5.

²⁹¹ *Boosey & Hawkes Music Publishers, Ltd. v. Walt Disney Co.*, 145 F.3d 481, 483 (2d Cir. 1998).

²⁹² *Id.* at 484.

²⁹³ *Id.* at 485.

²⁹⁴ *Id.* at 487.

²⁹⁵ *Random House, Inc. v. Rosetta Books LLC*, 150 F. Supp. 2d 613, 614 (S.D.N.Y. 2001).

²⁹⁶ *Id.*

already had the exclusive right to do so by virtue of its publishing agreements with the authors to “print, publish and sell the works in book form.”²⁹⁷ Nonetheless, the court interpreted this contractual language to refer only to more limited rights to publish certain hardcover trade books, leaving the authors free to contract with Rosetta Books to publish e-books.²⁹⁸

In each of these cases, without saying as much, in seeking to apply general rules of contract interpretation to novel situations decades down the road, the court favored the party that sought to release the copyrighted work in the newly available format. Whether or not this was each court’s aim, it is a laudable goal. When the contract language on long-term copyright licenses is not clear, an information theory of copyright law suggests favoring the party that is motivated to or would do the best job of publishing an existing work in a new (and often, increasingly desirable) format. Doing so provides an incentive to preserve valuable works over the long term.

2. *Deposit*

Copyright law also fosters preservation by its rules encouraging the deposit of copyrighted material in a central repository, the Library of Congress. Starting with the earliest American copyright law, deposit of the copyrighted work in a limited number of locations—then, in the district court where the author resided and with the Secretary of State—was a prerequisite to copyright protection.²⁹⁹ Although the deposit requirement has evolved over time,³⁰⁰ it still plays a role in copyright law. Currently, copyright owners must deposit “two complete copies of the best edition” of their work in the Copyright Office “for the use or disposition of the Library of Congress.”³⁰¹ Deposit is a prerequisite to an infringement suit,³⁰² and failure to deposit can subject a copyright owner to a fine.³⁰³

In sum, providing encouragement to preserve works in multiple formats as well as deposit them in a central location helps ensure that the works—and the messages encoded in them—are preserved over time.

²⁹⁷ *Id.*

²⁹⁸ *Id.* at 624.

²⁹⁹ Act of May 31, 1790, ch. 15, §§ 3–4, 1 Stat. 124, 125.

³⁰⁰ See Peter S. Menell, *Knowledge Accessibility and Preservation Policy for the Digital Age*, 44 HOUS. L. REV. 1013, 1026–39 (2007) (recounting this history).

³⁰¹ 17 U.S.C. § 407(a)–(b) (2012).

³⁰² *Torres-Negrón v. J & N Records, LLC*, 504 F.3d 151, 154 (1st Cir. 2007).

³⁰³ 17 U.S.C. § 407(d).

G. Error Correcting Codes

Until this point, some has been said about using copyright law as a way to correct errors that readers might have made about works they have consumed, particularly with regard to producing multiple forms of noisy expression to get at the knowledge contained therein from multiple perspectives. However, copyright law can and ought to think more generally about ways in which it can correct “errors” in transmitted messages, much like information theory led to the development of a variety of error correcting codes.³⁰⁴ An example of an error correcting code that now operates by norm in many—but not all—expressive disciplines is citation, or attribution.³⁰⁵ Citation is a way in which readers can be sure that another expressive work is linked to the one they are consuming, and for which the reader has another author’s description of what the cited work expresses. When a unified meaning is warranted and helpful, such as with factual knowledge, citation thus provides a method by which various expressive works can be brought together into a single decisive meaning.³⁰⁶ Similarly, when it is useful to link together various works into a noisy web of conversation about valuable expression, citation can provide the links to build that web. These are but two possibilities of thinking about implementing error correcting “codes” in some form or another in copyright law. More generally, examining this through the lens of information theory summons that possibility in copyright law.

CONCLUSION

Information theory and its notion of redundancy to cut through noise are helpful to explain what about copyrightable works is valuable and how the law ought to encourage these valuable aspects. Both the systematic, factual, and cultural knowledge found in expressive works and poignant expression itself are valuable aspects of expressive works. Copyright law thus ought to encourage both valuable aspects. For works that are valuable due to the knowledge they contain, it is critical that copyright law encourage this

³⁰⁴ See *supra* text accompanying notes 52–63.

³⁰⁵ See Fromer, *supra* note 24 (advocating attribution as a promising expressive incentive, and discussing its upsides and downsides generally); see also Christopher Jon Sprigman, Christopher Buccafusco & Zachary Burns, *What’s a Name Worth?: Experimental Tests of the Value of Attribution in Intellectual Property*, 93 B.U. L. REV. 1389 (2013).

³⁰⁶ Cf. Martin Shapiro, *Toward a Theory of Stare Decisis*, 1 J. LEGAL STUD. 125, 127–34 (1972) (reasoning that the practice of string citations in legal communication, surely informational redundancy, serves to ease the processing of legal data and also harmonizes various courts’ judicial rulings over time).

knowledge's appearance in multiple works redundantly as a way to transmit it, particularly given that the works containing this knowledge are frequently noisily expressive. For works that are valuable due to their contributions to expression, it is important that the expression itself be disseminated. Frequently underpinning this sort of expression is mystery as to the expression's meaning. Different audiences—in different contexts and also in different times—might locate varying understandings in this expression. These diverse and possibly evolving readings are intimately tied to the value of the initial expression itself. As such, they are valuable for society as well. In that vein, copyright law ought to encourage the idea that valuable expression can be worked over, repurposed, and interpreted. By doing so, copyright law can encourage the underlying expression to promote a continuing conversation. It can do so by allowing the underlying expression to be used redundantly—to attach to varying meanings and interpretations—thereby creating a unified conversation.