

**IN THE UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF NEW YORK**

THE AUTHORS GUILD, INC., et al.,

Plaintiffs,

v.

HATHITRUST, et al.,

Defendants.

Case No. 11-cv-6351(HB)

**DECLARATION OF GEORGE KERSCHER
IN SUPPORT OF MOTION FOR SUMMARY JUDGMENT**

I, George Kerscher, do hereby declare that:

Background and Qualifications

1. I am over eighteen years of age and am competent to make this Declaration.
2. I am legally blind.
3. Attached hereto as Exhibit A is a copy of my curriculum vitae.
4. I have dedicated the last 25 years to creating and promoting digital access to print documents for the blind. I received a bachelor's degree in English Education from Northeastern Illinois University in 1974 and taught special education and English in public schools from 1975 to 1985.
5. I then began working toward a master's degree in computer science at the University of Montana in 1985.

6. While working toward my master's degree, I developed the concept of computerized books for persons with print disabilities, a term I coined during the same time. A print-disabled person is someone who cannot effectively read print because of a visual, physical, perceptual, developmental, cognitive, or learning disability.

7. I developed computerized books because, as a blind master's degree candidate in computer science, I could not access even a single book I needed to complete my degree. I therefore decided to develop the technology to create such books for myself and others with print disabilities.

8. During my time as a student at University of Montana, I founded and developed Computerized Books for the Blind and Print Disabled (CBFB), through which I began creating e-books from files from publishers. In 1988, I created the first publicly available e-book, a copy of Mastering WordPerfect 5.0.

9. I did not attempt to patent the e-book technology because I wanted it to be readily available to anyone who was willing to make accessible books for the blind.

10. Ultimately, I left University of Montana without completing my degree. Because I could not obtain books relevant to my field of study, the thesis requirement for my master's degree was nearly impossible to complete. The university would not grant me thesis credit for the work I had done developing e-books. I chose instead to pursue my professional goal of improving accessibility for the broader population through CBFB.

11. Over the last twenty years, I have served on numerous panels and committees dedicated to improving the creation and distribution of electronic accessible texts for the blind. These include: The Commission on Accessible Instructional Materials in Postsecondary Education for Students with Disabilities; the National Instructional Materials Accessibility Center (NIMAC)

Advisory Committee; the U.S. National File Format Technical Panel; the World Wide Web Accessibility Initiative Steering Council; and the International Committee for Accessible Document Design.

12. On May 7, 2012 , I was one of fourteen individuals honored at the White House as a Champion of Change for leading the fields of science, technology, engineering, and math for people with disabilities.

13. I serve as the Senior Officer of Accessible Technology at Learning Ally. Learning Ally, formerly known as Recording for the Blind & Dyslexic, creates recorded copies of print materials for K-12, college and graduate students, and veterans and lifelong learners, who cannot read standard print due to blindness, visual impairment, dyslexia, or other learning disabilities. Learning Ally's collection of more than 70,000 digitally recorded textbooks and literature titles is one of the largest of its kind in the world. I have worked at Learning Ally since 1991, first as Research and Development Director from 1991-1995, and in my current position since 1995. Learning Ally is a 501(c)(3) non-profit corporation.

14. Currently, I also serve as Secretary General of the DAISY Consortium, an international association that develops, maintains and promotes international DAISY (Digital Accessible Information System) Standards for authorship and distribution, and am President of the International Digital Publishing Forum (IDPF), which is the global trade and standards organization dedicated to the development and promotion of electronic publishing and content consumption. Both of these organizations work to promote accessibility in electronic publishing.

15. Through my committee participation and my positions with the DAISY Consortium, IDPF, and Learning Ally, I have remained integrally involved with the development of electronic

books and am intimately acquainted with the issues surrounding the creation and distribution of materials in formats that are accessible to the blind.

Statement of Opinions

16. The availability of the HathiTrust Digital Library (HDL) stands to revolutionize blind students' and scholars' ability to compete with their sighted counterparts. The HDL titles I have reviewed are the most sophisticated and accessible scanned copies of print materials in a large collection I have ever seen.

17. New digital books can be readily made accessible but rarely are. Even if new books are to be made generally accessible, the expense of converting existing library collections with many highly specialized and even out-of-print books means that the type of mass digitization conducted by the HathiTrust, with complete metadata, is unlikely to ever occur again. There simply is no market for digital copies of old and out-of-print books in which only students and scholars have an interest. Publishers have not made digital copies for sale of the vast majority of the books that are available in a university library and are unlikely to do so in the future. Thus, the only way any one of these books will become available to the blind is if someone, either the HathiTrust, a disability student services (DSS) office, Learning Ally, Bookshare, or the NLS, makes an accessible copy.

18. To truly provide equal access for blind students and scholars to a university library, mass digitization of a collection like the HDL is necessary. Without this, blind students and scholars will always be limited to ad hoc access to titles they identify and request to be scanned without being able to search the library or skim materials in the way that sighted researchers can.

Without a fully digitized collection, therefore, blind researchers will never be able to compete with their sighted counterparts in academia on a level playing field.

Factual Basis for Opinions

I. Explanation of accessible digital books

19. Prior to the development of accessible digital books, the blind could access print materials only if the materials were converted to braille or if they were read by a human reader, either live or recorded. Accessible digital books that are available to sighted and blind alike are a revolutionary change for blind readers seeking access to content over either braille or human readers.

20. Although human narration was once the best access a blind reader could receive to print materials, the technology of accessible books has advanced far past the capabilities offered by human narration, making human narration alone substantially inferior to use of accessible digital books. To use a live human reader is expensive or burdensome for a family member or friend. Moreover, live readers' orations cannot be reproduced, giving the blind reader only one opportunity to hear the material. Live readers also cannot increase their speed – they are inherently limited to the pace they can reasonably read aloud. (Live readers may not be available until the wee hours the morning before a term paper is due.) Recorded human narration resolves some of these issues, like repetition and speed (and reader exhaustion), but presents its own problems. Typically, it will take six months to more than a year for a blind person to receive a requested recording of a textbook from an entity like Learning Ally. Moreover, even recorded human narration cannot be navigated like an accessible digital book and will not allow a reader to hear each character to discern spelling.

21. Today, blind readers access digital books with a screen reader or built-in text-to-speech software, both of which can output information either as a computerized vocalization of the text or as braille, through a refreshable braille pad. Unlike books narrated by human readers, accessible digital books can be read as quickly as the reader wants, or even skimmed. Further, they provide significant search and navigation capabilities, allowing readers to jump from chapter to chapter, paragraph to paragraph, and sentence to sentence, as well as to discern spelling. This allows blind readers to re-read certain sections of a work they might not grasp on the first pass, just as a sighted reader may re-read a complicated passage.

22. Not all digital information is accessible. For example, scanning a copy of print material usually results in a file in portable document format (PDF). PDFs are created essentially by taking a picture of the page. This gives a sighted person enough to read on a computer screen, but it does not allow screen reader software to recognize the text.

23. To take this next step toward accessibility, the scan must be run through optical character recognition software (OCR) and optical structural recognition software (OSR). OCR/OSR software takes a high resolution image of the page and recognizes the image of characters and even structural data like columns and images. Character recognition software looks at the characters and compares them to a database of what it knows. For example, the software will match an image of the letter “c” to image of the letter “c” in its database. The software will also check spelling, to ensure it has matched the image correctly to images of characters in known words. The OSR component will recognize word boundaries, text block boundaries, and, on occasion, headings. The software then identifies the x/y coordinates of all the characters on a page and attempts to identify the correct reading order for each page, when there are columns or images that alter the usual reading order. The OCR process also allows the text to be searched.

24. A further step called “tagging” provides additional metadata about the content, such as the existence of tables in a work or the existence of headings and other document structures. Although the OCR engine will try to add meaningful style information, no existing software can recognize document structures perfectly and this final step must be completed manually. Only materials that are originally created for digital books, or “born digital,” rather than scanned from print material do not have to be manually tagged. Tagged works provide to blind readers the closest equivalent to the experience of a sighted person reading the material in its print form, but the labor required to create them has made them very rare.

25. Accessible digital texts present a further benefit for low vision readers over human narration alone. These users often will use print and sound at the same time. They may be able to visually discern paragraphs or chapters while using sound to read characters and words. Human narration therefore is substantially inferior for low vision readers who have some usable vision.

26. Even what are commonly referred to as “audiobooks” do not provide the benefit of accessible digital books. While having Jim Dale or Stephen Fry read *Harry Potter and the Order of the Phoenix* is ideal for entertainment purposes, it does not provide equal access for academic or scholarly pursuits. The ability to access text at high-speed is crucial for students and researchers alike—accessible digital books, like those in the HDL, make high-speed access possible, where audiobooks cannot. Digitally accessible books make it possible for readers with print disabilities to “virtually” bookmark a page, to electronically jot notes in the margin, and to digitally riffle through pages to “scan” for just the right passage. While there was a time where a book read dramatically or even non-dramatically by a human was the best users with print

disabilities could hope for, advances in technology mean audiobooks do not equal (and are vastly inferior to) OCR'ed books in the modern era.

27. The DAISY Consortium and the IDPF have established standards to ensure that “born digital” material is accessible. Any digital copy of print material that is created to meet the DAISY standard will be fully accessible to the blind.

28. The IDPF develops and maintains the EPUB content publication distribution standard, which is a generally available open standard, available without royalty, for the next generation of commercial and non-commercial digital books. The standardization of a distribution file means that publishers can design their print materials using any authorship tool, convert them to an EPUB file, and then provide that file to any e-book distributor, which will be able to publish the content on whatever platform it uses.

29. The latest EPUB standard, EPUB 3, incorporates the current DAISY requirements for distribution, which ensures that all documents published using EPUB 3 that follow the accessibility guidelines will be distributed in an accessible format, unless publishers then convert the EPUB files to platforms that are themselves inaccessible.

II. Availability of accessible books in higher education

30. I spoke with the University of Michigan Library back in 2005 (before it established the HathiTrust). At that time, it had already taken proactive steps to make its digital collections accessible to users with print disabilities. Even in its early incarnation, the University of Michigan Library's accessible book platform was already enabling students and scholars with print disabilities to make unprecedented and meaningful use of the library's vast collection.

31. Since then, I have had the opportunity to review a number of the digital books in the HDL and to discuss the technical specifications of these scans with personnel from the

University of Michigan Library. The HDL scans are high resolution images that have been digitized using the most sophisticated OCR/OSR software I have ever encountered. Although images are not described and tables are not tagged, the table text is present, and the scans include the vast majority of metadata necessary to make them fully accessible. They can be navigated by chapter, page, line, and character. My understanding is that the collection encompasses close to ten million books.

32. Today, as when I was a graduate student, it is virtually impossible for blind students to conduct library research. A university's disability student services office (DSS) is responsible for scanning print materials and converting them into accessible digital copies for blind students, but the vast majority of these offices will only provide the works listed on the students' syllabi. They simply do not have the resources to create copies of books that are not required reading, and certainly not do so in a timely manner. As a practical matter, this means it is impossible for blind students to conduct independent library research. Even when a student switches classes or a professor adds a reading to the syllabus after the fact, DSS offices are often overwhelmed and unable to fill the requests. It may take weeks or even months for the student to receive the scanned materials.

33. The quality of the copies made by the DSS offices varies substantially from university to university. In the vast majority of cases, the scans will only be run through very basic OCR software, without any of the structural recognition in the HDL scans.

34. Even more significant, indexes and tables of contents are not available in an accessible format in almost any university library. Thus, blind students cannot view the index or table of contents of a book to see if it contains relevant information. In the HDL, most of the tables of contents have been manually tagged, allowing blind students to recognize them and navigate to

them with a screen reader the way a sighted person would open the book and flip to the table of contents.

35. At the universities with the best DSS offices, a graduate student may be able to provide a list of materials for research that the office then will have the capacity to digitize. The office, however, is limited to the books the student initially identifies as relevant. Blind students cannot do what sighted students do, that is, browse through many books to find the chapters or sections that are relevant.

36. At the vast majority of universities, where the DSS offices do not have the capacity to honor requests for research materials, a blind student's only option is to use a scanner in the library to scan individual books of possible interest one page at a time, listening to each, until he or she finds the tables of contents. It is an impossible task for a blind student to use a library in this way; the time it would take to complete this process prohibits blind students from completing any library research at a pace at which they can compete with their sighted peers.

37. Besides universities' DSS offices, the only accessible digital books available are those available for purchase as iBooks or Blio books, and the collections of Learning Ally, Bookshare, and the National Library Service for the Blind and Physically Handicapped (NLS). Bookshare is an initiative of the non-profit organization Benetech® that creates accessible copies of popular digital books and academic textbooks on an ad-hoc basis for people with print disabilities at no cost. NLS is an affiliate of the Library of Congress.

38. From my experience with Learning Ally, I know that each of these entities has a very limited capacity to make new books. Further, Learning Ally and the NLS focus their limited resources on particular titles with the greatest appeal. NLS focuses on novels and other current popular works. Learning Ally and Bookshare place an emphasis on K-12 education. Although

they do digitize some books for higher education, both have very limited budgets. Their collections therefore are significantly different than the HDL, which naturally has an academic focus. Learning Ally has approximately 70,000 titles in its collection, Bookshare has approximately 150,000 titles, and the NLS has approximately 20,000 titles. These include many that overlap. In total these organizations have approximately 200,000 titles available to blind readers, while the HDL has ten million.

39. The AccessText Network, a membership exchange network that is intended to facilitate and support sharing of textbooks for students with diagnosed print-related disabilities, has had limited success and has only focused on textbooks identified in the syllabi of students. The Network is intended to connect DSS offices directly with publishers to receive electronic files and facilitate the sharing of scanned copies between DSS offices at different universities. As an initial matter, the program involves voluntary participation and neither have publishers joined as expected), nor have DSS offices shared their files at the rates the founders of the network had hoped. Further, the network does not have a quality control mechanism to ensure that texts scanned by different DSS offices have the necessary structure and content. In addition, it is limited to textbooks and required items in syllabi, and therefore does not include the vast majority of titles available in a university library. Finally, the Access Text network was established because there was deemed to be no meaningful market in the blind and print-disabled community. That publishers are expected to give away the electronic files for free demonstrates that those involved do not believe there is any market for accessible books created for the blind.

40. Today, for scholars and students with print disabilities, the best promise of meaningful access to an academic library exists at the University of Michigan through the HDL. It is the kind of access, at the minimum, that should be available to all in the academy.

III. **History of failed attempts to achieve market-based access to digital text for blind readers**

41. Learning Ally struggles to find charitable funding because there simply is no market for accessible books for the blind. Learning Ally, Bookshare, and the NLS exist because of this market failure.

42. In 2007, I attended a presentation at the Annual International Technology & Persons with Disabilities Conference at California State University, Northridge, at which the Association of American Publishers announced that it had conducted a study and determined that there was no exploitable market for the creation of accessible print materials for the blind.

43. Authors and publishers have not only ignored accessibility concerns related to digital texts, but actively worked to prevent the market from reaching the blind. When Microsoft created the first commercially available e-reader device in the late 1990's, Microsoft and its competitors, Adobe, Gem Star, Sony, and others, ignored persons who are blind or print disabled. They did not build in any accessibility features that a blind person could use. While the underlying content was accessible, the user interfaces did not cater to the disabled community.

44. All of these companies indicated that the effort to make the products accessible did not justify the return on investment. From contemporaneous discussions with persons in charge of the various e-book programs or in charge of accessibility at each of these companies I learned that the choice to exclude the blind to preserve anti-piracy software was a deliberate decision. They consciously decided that the work to modify software to make it accessible to the blind was not economically worthwhile in light of the perceived small incremental addition of the blind to the market. They recognized that people with disabilities would be left out, but they were not willing to develop mechanisms for the blind to access the underlying information.

45. This trend has continued. The development of popular e-book platforms that are inaccessible, like the Amazon Kindle and the Barnes & Noble Nook, demonstrates that tech companies and publishers do not believe that there is sufficient economic benefit from making accessible books, or at least that their perceived concerns about possible piracy outweigh, from a business perspective, any monetary or societal benefits from creating accessible books.

46. Indeed, I, along with representatives from the National Federation of the Blind attempted to lobby Amazon to make the Kindle accessible, but encountered opposition from copyright owners and their allies. We met with representatives from Amazon, presented statistics concerning the market for talking e-books, and demonstrated the minimal cost associated with making both the text of the books and the menus on the Kindle accessible for people with print disabilities. But, when Amazon announced that it had released the Kindle 2 with a text-to-speech function, the Authors Guild actively opposed Amazon's policy, and Amazon capitulated, allowing individual publishers to turn off text-to-speech on the Kindle for, at their selection, all or some of their booklist.

47. Further, even when Amazon activated the text to speech function on the Kindle, it only worked for the text of the book, not the menus. Blind users therefore cannot effectively use a Kindle book. Amazon's failure to make these minimal changes in its platform demonstrates that it does not consider the blind to be a significant market.

48. New books could be made accessible with little expense to publishers. All new books are created digitally. However, the design software commonly used by publishers takes the accessible word processing files submitted by authors and converts them into an inaccessible format.

49. Because of the DAISY standards and because of partnerships, we have made some progress in building accessibility into new e-books. Adobe Indesign 6, the premier electronic publishing design software, exports into EPUB 3, which makes the basic text accessible. But, these new EPUB materials may still be made inaccessible if they are transformed for use with inaccessible platforms, such as those used on the Amazon Kindle or the Barnes and Noble Nook.

50. Given the lack of a market in the blindness community even for new popular books, and the publishers and technology companies' persistent refusal to make their products accessible to the blind, the access problems faced by blind readers with respect to academic library collections are unlikely to ever be solved unless the HathiTrust is permitted to continue providing accessible digital versions of the books in the university libraries' collections.

Conclusion

51. Based on the facts set forth above, and my experience and expertise in providing accessible books for the blind, it is my view that the HDL represents an unparalleled opportunity to achieve true equality in higher education for blind and print-disabled students and scholars; and that the opportunity to participate in education on a basis of true equality is very unlikely to arise again if the blind are denied access to the HDL.

I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct.

Dated: June 28, 2012



George Kerscher