Transcription: CoreDH – Podcast Episode 6

Host: Alexandra Núñez (Alex)

Guest: Glen Layne-Worthey

Disclaimer: Glen Layne-Worthey was Chair of the ADHO Executive Board until 2023. **The interview was conducted on 12.05.2023.** The current ADHO Chair is Diane Jakacki (2023-2025).

—

Intro (Hannah): Welcome to CoreDH. A podcast that explores the stories and people behind the Alliance of Digital Humanities Organizations.

(Music)

Alex: Welcome to our sixth podcast episode. My name is Alexandra Núñez (Alex) I am a communications fellow at ADHO, and this is my second conversation I have with an exciting and impressive DH personality. Today I am hosting Glen Layne-Worthey! Thank you very much, Glen, for being here! It's great to talk to you today.

Glen Layne-Worthey (Glen): Thank you very much, Alex. Nice to talk to you, too.

Alex: Before we start just a quick note: The Podcast is divided in 3 Parts and in the first one, I would like to give listeners the opportunity to get to know you, Glen, as a DH researcher a little bit better. In the second part we'll talk about one of your amazing research areas, namely, digital library collections and machine learning methods. In the last part I would like to talk a little bit more about your function and role as the chair of the Executive Board at ADHO since you are an influential figure in Digital Humanities. To the audience: You will find some timestamps which guide you directly to the different parts of this podcast episode.

So, Glen, you have an impressive resume. You are the Associate Director for Research Support Services at the HathiTrust Research Center, based at the University of Illinois at Urbana-Champaign School of Information Sciences. Previously, you were Digital Humanities Librarian at Stanford from 1997 to 2019 and founding director of the Stanford Center for Interdisciplinary Digital Research (CIDR). You received your degree in Russian children's literature from the University of California, Berkeley. You hold a number of positions in the international digital humanities community, one of which is that you have served on the Alliance of Digital Humanities Organizations (ADHO) Executive Board since 2020 and as Chair since 2021.

To the listeners you'll find a link to Glens Orcid Profile and to his projects in the description and in the transcription. Let's start: Based on your career path: **How did you get involved with Digital Humanities?**

Glen: Thank you very much Alex for that great introduction. And I'm very flattered. I should clarify, and this is important to my story that I never finished my PhD degree in Russian children's literature. That's a source of both shame and a fact of life. So, I'll talk about that a little bit. The way I got started in DH was that I was studying at University of California Berkeley in Russian literature, as you said. And I needed to get a job. My first child was about to be born. I didn't have health insurance and a really wonderful job came open in the library as a library cataloger, using my knowledge of Slavic languages in eastern European cultures. So, I took that job, got my health insurance, had my baby, but also discovered an entirely new field for myself. A new set of fields first was the library world that introduced me to the concepts of metadata and computers. Oddly enough, at that time, computers were not ubiquitous in people's homes. There were, of course, there were personal computers. This

was the mid-nineties, but they weren't really, they weren't even close to as widespread as they were now. However, in the library, we used computers for absolutely everything for cataloging, for management of the collection, etcetera, etcetera. While I was there, I… through a colleague, I became acquainted with text encoding and was introduced to the text encoding initiative. And what at the time was called the <u>Berkeley Finding Aid</u> Initiative, which later became <u>Encoded Archival Description (EAD)</u>. So that's an XML standard for archival work. And I never looked back. I just loved the work so much. Um and I did continue with my phd program, but because of my job, my family situation and my job situation had me take the first opportunity that I could um professional opportunity. I accepted a job at Stanford doing more or less what I would do for the next 20 years. I had always intended to finish my dissertation. I just got too busy doing other things and it wasn't all that relevant to my work, but I may still finish it all these 30 years later. This is important because it also set me on the path to an alternate academic career, something that I have found to be really deeply meaningful. Um I, yeah, I, I love working in alternate career paths. I think the tenure track path is a great one as well. And many, as we all know, many of our favorite digital humanities people are on the tenure track, but many are not. And that's a really important and I would claim a really good and unique feature of our field of digital humanities that as a field, it's really open to people with PhDs and without PhDs people working in teaching positions and library positions and curatorial positions, all sorts of all sorts of things that are not really found in purely academic research communities. Um So I think, II I don't, I don't mind having taken this path at all. In fact, I found it really, really meaningful.

Alex: Ok. Well, the second question is deeply connected with the first one, I think, **why do you conduct research in the field of digital humanities and what drives you?**

Glen: Well, now I think it was by chance, but it may have been by my own design. In addition to studying Russian literature and English literature as an undergraduate, I was also a physics major for a long time and I've always wanted to bridge what seemed to me like kind of artificial gaps between the hard sciences and the humanities. I sort of, I chose those majors, not because there were any good jobs in them. Certainly there aren't. Um, but because I wanted, uh naively, I wanted to sort of understand everything I wanted to understand humans and, and culture and the universe. And so I chose these things to study and, and enjoyed them all. So when, um when I got the job in the library and using computers for intellectual work, um, was a big part of that job. I found that I had an aptitude for them. I was never a great coder or anything, but I wasn't afraid of them and they, they seemed cool. Um, at that time, computers were a pretty niche interest for most humanities people. Uh But my experience in the library and my experience with um thinking of texts as data and text, as metadata, these were all new concepts to me, but they sort of fit into my world view. What drives me now, as I said, I've gone, I've taken the library career path and most of my current work is really in close support of other people's research topics. It's controversial and for good reason to say that the librarianship is a service profession. But for me, in particular, it's been most meaningful and inspirational for me to be a participant, not a servant, but a participant in other people's research. And that's something that's really crucial in DH and I found that my, my ability to think in a computational way while also understanding humanities or having experience as a humanity scholar has been really appreciated by, by people in our field who either are beginning and don't have knowledge of one or the other or as a way of communicating with highly technical people who haven't thought about the humanities. I think the bridging feature is something that my background made me well suited for and seems to be an important role for us.

Alex: And now I'm also interested in your thoughts and perception of the DH community**. What does digital humanities stand for? What characterizes the scientific DH community for you?**

Glen: I think the most salient characteristics are a sense of community, a sense of collaboration, a sense of innovation and new ways of thinking and a particular open mindedness. And I think all of these things uh reflect what I just said about my own career in DH. I think that one of the best definitions of DH, I mean, like many fields we're always defining ourselves as a community of practice. And that community part, I mean, the practical part is, is important as well. But the community part is what I would like to speak to. So I think DH as a community has a stronger identity than many other professional communities. Certainly not all of them. Collaboration is also important because of the breadth of what we do and the, and the different worlds from which we draw our tools and practices and theories. So I think all of these things build on each other, the need for collaboration and the desire for community. I also really like that Digital Humanities people are explicitly trying to think about old questions in new ways. I think thinking about things in old ways is absolutely wonderful as well. But I do like in our community that we are experimental and open minded. And then once again, I can just say about the community, uh an open mindedness in the community is not only intellectual but also cultural and organizational and personal, social. That's what allows people like me who have incomplete PhDs or a kind of different career track or not it allows us to be treated as equals in the community. That's I think that's really, really crucial and it's certainly something that's kept me happy in our community.

Alex: Thank you very much for your openness. And this is, this was a great, great statement. So, let's dive a little bit deeper into one of your great fields of research, you are part of the AEOLIAN network team. And there have been some exciting workshops and talks over the past year, the interdisciplinary AEOLIAN Network focuses on the uses of artificial intelligence for digitized material from the cultural sector to make it accessible to users and offers exciting workshops. And to be honest, I was really impressed when I discovered your research in the AEOLIAN because as a linguist, I don't have the area of library collection so much in my mind when I think about machine learning, especially because chatbots are very hyped in the media right now. Anyway, in my research, I have read some of your papers that have tested applied machine learning techniques on digital library collections. And I'd like to ask you a few questions about this. **So, in what ways can machine learning techniques be applied to digital library collections and archives? And what are the user benefits?**

Glen: Thank you for that question. And thanks for mentioning the AEOLIAN network and some of the papers that I've now co-authored. I actually want to start with some pretty serious caveats. I am, it sounds like your anti hype as far as chatbots and things like that go. I think uh pushing back against the hype is absolutely crucial for people in our field. And I certainly am constitutionally opposed to all of that hype. I don't think they're the answer to everything. I think there's more reason to be cautious of them than to be enthusiastic. On the other hand, I do work closely with a lot of people who are better at using machine learning and artificial intelligence and I trust that they, that the things that they do are meaningful. As for my co-authored papers, this is a curious thing. I'm gonna try not to sound falsely modest because it's not false at all. My contributions to these papers come from my collaboration as a co-author, most probably all of the research that you mentioned, I have co-authored with graduate students that I work with, now that's been just a true joy, to be honest, they have contributed the vast majority of the intellectual content. And I feel lucky just to have my name on the papers with them. So, I do actually, I mean, I do work and I and I talk things through with them and I edit many of them are not native speakers of English, something that I'm very sympathetic with. So I'm mainly a co-author, so it would be just false to present myself as an expert in machine learning even though I do, I mean, it's a privilege to have this great grant funded project, the AEOLIAN network, but it's a highly collaborative project. Also a fairly skeptical one. We've been really um careful to … in those workshops and in our publications we've been careful to be cautious about the claims that we make to be um again anti hype. We do celebrate the many things that machine learning gives us. So in answer to your

question, what does it actually give us? What are the positive things? I think the biggest thing that it allows us to do is to deal with the sheer scale of modern digital library collections. So my current professional identity is in the HathiTrust Research Center with the HathiTrust Digital Library, which is 17.5 million volumes, a scale that we've never as, as humans we've never dealt with before. And we're not, we don't have the capacity to do that. Machines are great at doing that. They're not so great at many other things that humans can do. So I see machine learning necessarily as a collaboration between algorithms and humans. Uh These concepts are familiar to everyone probably who listens to the podcast. Most of the time we call this the sort of machine enhanced reading, distant reading. Um but over the years as I worked closely with people like Matt Jockers, who was a very good friend and is a good friend and was a close colleague at Stanford. He called it macro analysis. My current colleague at the University of Illinois Ted Underwood calls it distant horizons, right? The ability to, to see at a distance and to and to at least have some hope of interpreting um hundreds of thousands or millions of, of long texts. Uh machine learning is absolutely essential in that.

Alex: Ok, let's focus a little bit on semantics. **To what extent can the use of machine learning help to optimize or better understand the content of the textual heritage?**

Glen: Well, in addition to that distant reading and that scope, um there are lots of things that machine learning if, if it's understood broadly can help with. So in my work, natural language processing is really important. Many definitions of machine learning include NLP, Natural Language Processing, I tend to take a broad and historical view and I think we've always had machine learning type processes. Some people would argue with that definition, which is fine, but that's how I understand it. Even optical character recognition (OCR) is um is or makes use of machine learning methods. So both of those things are hugely important. Um In the, in my current work at HathiTrust, we rely on a machine, not on a machine learning to digitize a book, but we do rely on machine learning to turn the pictures of a page into a, into a bunch of text. And then further, we rely on machine learning to um interpret, to begin to interpret that text, to tag it with part of speech and to do uh word counts and things like that. So, um even in old fashioned machine learning that, you know, some of those, many of those processes are decades old, machine learning is important. I have less experience but I work closely with people who have a lot of experience with language models. That is a task that humans can do on a small scale interpretive, I mean, you can sort of read a text and get a sense of style and you can say that this one sounds like Tolstoy, not Dostoyevsky, right, humans can do that, but we can't do that at a larger scale, machine learning methods have enabled us to do that at a larger scale with texts by authors that are less famous than Tolstoy and Dostoyevsky. So we can sort of study again, the not just the, the what we think of as the, as the tops of the mountains in literary history, but we can study the whole, the whole landscape. Yeah, and then large language models (LLMs) seem to have some promise for cultural heritage and for literary studies. I don't understand that very well. But again, people that I respect like Ted Underwood and Ben Schmidt are highly hopeful that the large language models (LLMs) if constructed carefully and thoughtfully, which they're not necessarily yet uh can have a lot to a lot to help with. And then finally, the last thing I would say about this is that many other thinkers that I respect, deeply, are highly skeptical and even fearful of that. And I respect those, um, I respect those opinions as well and I share some of them, not all of them.

Alex: Ok, and the third question touches on the area of **AI ethics** which will certainly come into sharper focus and will be a hot topic. I think in the coming years, the question relates to some challenges that may be encountered and need to be considered when using machine learning techniques and digital library collections. So, **what do we need to consider or**

what role does copyright play when applying machine learning to digital library collections and their content?

Glen: Well, this is also a controversial topic and not only intellectually controversial but legally controversial, right? I am of the strong opinion that the work that we do in modern digital libraries is and should be celebrated as a fair use of those works. So fair use and fair dealing are concepts that differ in different legal environments. But at least in, in the US and in many um many regimes, many legal regimes um digitization without permission is considered a fair use as long as a number of factors are met. And certainly what we, I mean, among those factors are things like impact of the market and whether the amount of copying or digitization is required for what we do. And in the US, the HathiTrust when it began was subject to a very, very long lawsuit by the Authors Guild of the United States. And we won that lawsuit in a really resounding and heartwarming way. And when I present the history of, of HathiTrust and the history of what we do, that's an essential, an essential document. So, um I absolutely respect copyright, the rights of people who have created works to do what they want with those works. But I don't believe that any creator has the right to tell us how we can make use of that work, how we can read that work. They can, they can protect their rights, they can protect their economic interests, they can protect their intellectual interests. But they can't just go in, just as they can't tell us that we are only allowed to read their book at a desk and not read it in bed or in the bathtub or on the beach. I think likewise, we shouldn't allow copyright holders to say that we can't read it with a computer or that we can't, we can't, you know, we can't underline it or make notes or cut it up and do all these things and, and what we do in machine learning and even in the less machine learning, heavy aspects of digital libraries are close equivalent to that. Now, there are plenty of things we shouldn't be doing. And I, and I, yeah, I just want to be clear that I don't support making a copy of a work and redistributing it ad nauseam without the author's permission. Um, many creators, uh, do want their work to be reused and they're open about that if a creator doesn't want that, that creator should be respected. But I don't think, um, I don't think we should necessarily restrict ourselves to the uses that the original creator had thought up. I think we do a lot of things on our own. We buy books for libraries and then we circulate them and the authors don't get any money for circulation, at least in some regimes. And we've decided as a culture that that's allowable. So likewise for machine learning, um I think um I'm sure there are ways that that could be that there are doubts that could be sown in my mind and I may entertain them and I may even agree with them. But for now, I think we're pretty safe both legally and ethically in doing what we're doing. Um not in disregard of copyright but using the safe havens of copyright such as fair use. Uh and I also think that it's important for us in the academic community to push back against commercial interests because some of the even though one wants to respect a creator's rights, one doesn't necessarily want to allow the interests of capitalism to dominate the cultural heritage world. I think that's a, I think it's important for people in our field and related fields to push back against that and say, um fair use as a protection of cultural heritage, protection against commercialization and commercial exploitation. So there, there are all sorts of, I'm glad you phrased this as an ethical question. I think it is also legal, practical, economic, and also social.

Alex: Yeah. Well, this was really enlightening. Yeah, thank you very much. Last but not least, I'm very curious right now. What DH projects are you now conducting or planning or if this fits better: What is your dream DH project? You can choose...

Glen: I can choose or as you can tell by my very, very long answers, I might do both. I answer both because they're both great questions. Some of the projects I'm working with uh working on in the HathiTrust research center have to do with identifying gaps in our collection. Our collection is the biggest digital library, academically focused digital library ever assembled. But um we, we know that there are gaps in it. We know that there are people

who are marginalized. So many, many of the gaps in our collection we don't know about because there's not a simple way to identify them, but we do have a big project that has worked closely with highly specialized research teams to identify gaps in their particular research areas. Some of those include African American writing, Native American writing, uh Spanish language writings in US libraries. These scholars have helped us to identify gaps and we are doing our best to fill those. So that's a great project. Machine learning hasn't been all that helpful there. It's really old fashioned bibliographic work. Uh and, and the hard labor of searching the collections to see what's there and what's not. We also have a project that I'm very fond of that is trying to make our data more accessible. So when we present our main data set, which is what we call extracted features it's basically a copyright friendly, a fair use friendly version of texts even that are in copyright. Uh We convert those to, to statistical data and metadata which is not subject to copyright and um but it's also not easy to use. So we have a new project to make that more accessible and more and more easy to use. Some of the examples of what we're doing with those two projects are um I just want to cite a few specific examples from good colleagues. One is looking for examples of a modern fictional genre in this case, the Black Fantastic finding examples of that or characteristics of that writing in older nonfiction writings. This is work that has been done in collaboration with Clarissa West-White at Betham Cookman University, Seretha Williams at Augusta University and Maryemma Graham in the University of Kansas and the History of Black Writing. It's a great research team and just a really fun and thoughtful and open minded project. Another set of projects that I work on with our PhD students in HathiTrust are testing the limits of current digitization practices and how those are impacted through how those impact the outcomes of machine learning methods or other digital methods. For example, how, how bad can uncorrected ORC be before downstream tasks are damaged? Or how can we extract meaningful information from page images, meaningful information being not only text but also images and captions and tables and graphs. These are hard problems that machine learning is helping us with and that really impact how much we can do with our huge digitization project. And since you offered me the chance to talk about my dream DH project, um I still have one unfinished dissertation chapter that I mentioned at the beginning that has to do with Leo Tolstoy's pedagogical writings. We, we know him for his novels mainly and then for his short stories, some of his philosophical essays, but he has a whole period in his life where he was a school teacher and it was immensely important to him and nobody, I mean, people have studied that of course, because he is Tolstoy but I think I have an interesting way to study it, but I haven't. And I think there's a DH method in there that can help me study similarity and difference, whether in style or topic or word choice or so, maybe, maybe someday I'll do that as a DH project and finish my PhD.

Alex: Thank you very much, Glen, for these great insights into your amazing research and future projects. And finally, I would like to broaden the focus a bit and talk about ADHO and your role and function as chair of the Executive Board of ADHO. So, tell me about the history of ADHO and why it was created.

Glen: Well, I think the, I think its origins are reflected in what I said is the, the most important aspects of the, of the profession of, of DH and that is community and collaboration. ADHO began, it was founded in 2005 or its founding protocols were finalized in 2005. And it officially started in 2006, but it was specifically meant to combine two communities that had arisen on their own somewhat organically and then discovered each other. One of those is the Association for Literary and Linguistic Computing (ALLC), which was founded in 1973 in the UK and included many people in continental Europe. In 1978 the Association for Computers in the Humanities (ACH) was founded. These two communities, um they have their own histories. But around the late seventies, they discovered that they were working on very, very similar problems, and began going to each other's conferences, that is people in the associations, not the associations themselves. In the early 2000 or maybe late 1990ies, they

began having joint conferences and then decided to combine forces rather than compete. So, um there was an economic question at the time, each of those organizations, by that time had a professional journal, but it was felt that there that um people were not likely to subscribe to two journals or belong to two associations. So they combined forces. Uh so as to collaborate and not compete. Um those two were quickly joined by the Canadian Association for Digital Humanities (CSDH/SCHN) and uh in just the space of less than 20 years, they are now 11 consistent organizations that have joined ADHO but I think it's, it's really comes down to creating a global and international scholarly community, a community of practice and to share resources and to share knowledge and to share a community.

Alex: Thanks for, thanks for giving us the history of ADHO. Now, I would like to ask **why should DH organizations reach out to or be part of ADHO and what does ADHO offer?**

Glen: I think it's precisely those same qualities of community collaboration, connection. Uh There's some amount of resource sharing. We all know that the global uh situation, the global economy is unequal for us. Um, I have found that people in our community are very, very uh open and enthusiastic about the idea of, um we shouldn't call it wealth redistribution. That's a bad word in some context. But we're all, we're very uh willing to share what we have with those who have less uh these resources come in all, all sorts of forms, time and money being the main ones. I think uh I think an organization that doesn't yet belong to ADHO but is working in our space. I just think there's a greater advantage in making connections with more people than there is in, in trying to re-invent the field on its own. As new organizations have come into ADHO, we've certainly found that each one brings its own new set of concerns and practices and values and those have enriched the community. So from ADHOs point of view, we want all organizations all around the world to join us because we want to be more diverse. We want to be more representative of the world. I suspect you're gonna ask me about challenges next. So I won't go into that here. But uh being more inclusive of the world is a way to, it has its own challenges, but it's a way to make our community better. And I think by joining a community that's continually getting better. Other DH organizations will benefit.

Alex: Well, you build a good bridge um because life presents us with new challenges every day, both as individuals and at the international organization level transfer to ADHO. What kinds of challenges does ADHO face at the moment and why?

Glen: I think all the challenges that I can that I have seen come from really positive factors. So the overarching challenge is growing pains. So this organization of, of just a friendly organization of two people, two organizations headed by two people who were very friendly with each other. That was a pretty simple thing to, to agree on, they agreed on things, they talked to their, to their colleagues on, on the two boards and the deal was done. Now that we have 11 organizations and we want to have equal representation. We want the smaller organizations to have a voice that's not drowned out by the larger organizations. That's a difficult thing organizationally. So the bureaucracy that we have um has increased somewhat. Listeners may know that a few years ago, we had a major reorganization of ADHO precisely because of the number of organizations that made the old method unwieldy the old way we counted members of the board and who had how many votes and who had, who got which fund distribution of funds. So we've thought to make that more streamlined and more equitable, but it was, it's been a hard reorganization. Um I think it probably comes from the fact that we're still very much a volunteer organization. Volunteerism or the lack of volunteers, the lack of time that all of us who do volunteer have is a continuing problem. And it also makes the reorganization harder. I mean, if we, if we could afford a management consultant to come in and tell us how we should do things and then a professional to, to implement all those changes, it might be easier, but none of us has the resources to do that. Um I, I think I do like the amateur nature of our, of our organization, the volunteer nature. It

means that the people who do the work are genuinely committed to it. They, generally, don't seem to have ulterior motives that could damage a working relationship. They all are very open hearted and open minded. And um yeah, so that challenge is one that I think will continue. I, we do continually need new people to join. ADHO, by the way, has just published a call for nominations for several positions and there will be another call in a couple of months as people rotate out as their terms end. My term (EB Chair: 2021-2023) is ending in a few months after a very, very long stretch. Uh I would love to find I would love not only to find, but to have the community find, um, someone who could take my place and um, yeah, that will be a great thing. There is someone to take my place immediately. That's Diane Jakacki. Uh, but we'll need someone to take her place as Chair-Elect of the Executive Board. So it's always a challenge to find people to do these things. I think it's a combination of everyone being busy and a combination of people being shy about putting themselves forward. Um It, it may seem like a careerist move to declare that you um that your leadership material, but I would encourage people to take a step back and just realize that you're joining a group of like-minded people who are committed to the field. Yeah, and I think there's another challenge in perception. I know from being in the community for a very long time that many people think of us as a mysterious bureaucratic organization that is hard to break into and hard to get out of what you want. Um I think we, I think we're guilty of some of the, some of the effects of those opinions, but I don't think the opinions themselves are true. In fact we are the community, we write the protocols ourselves and we negotiate with our colleagues to make sure that they're equitable and good and to and to implement them. So it's um yes, of course, it's a bureaucracy and of course, there are some things that happen behind closed doors but the doors are generally very, very, very open.

Alex: Ok and the last question that addresses your motivation as the Chair of the Executive Board at ADHO: What **drives you as the Chair of the Executive Board of ADHO? And where do you see the future of ADHO in a few years?**

Glen: Working in various roles has for me just been a highlight of my career as much as I like the work of digital humanities. I like the humanities and I like the digital, as I've mentioned. It's actually working with these dedicated people who are, many of them have been pioneers in the field and many of them still are pioneers in the field, new pioneers. As I said, by virtue of the fact that we're all volunteers. We're also all very, very dedicated to what we're doing and we care about our colleagues. I mentioned at the beginning how important it's been for me to have a non-tenure track, a non-standard uh academic career and to be respected in that role. So I think that the ability for many to, regardless of what your role in the academy is, to be in a position of leadership and in a contributing position is a great thing. And that's, yeah, that's, that's made my work really, really meaningful. I also think this should be, I think this needs to be emphasized. It seems natural to me, but I know that not everyone knows about it. The Digital Humanities seems to many to be a new powerhouse in the academy and it seems like it has lots of resources and has lots of attention and it does have an increasing amount of resources and attention, but it certainly didn't start out that way. Many of the people who worked in the early days of ADHO were in positions that didn't get respect at their universities and they discovered fellow travelers at other universities, similarly intelligent, similarly dedicated, similarly innovative, similarly scholarly but they weren't getting the respect they needed in their universities. In ADHO, they were getting that respect. And I think that has created a community that in general is non-judgmental and non-hierarchical. I think those kinds of marginalized origins are something that we should be proud of and cling to and um and yeah, keep as a standard for the community. The growth is great. It's great that there's interest from all over the world. It's great that our organization is becoming big and complicated. All those things are challenges. But II, I think one way that we deal with those challenges is by always being true to our roots of yeah, of embracing our marginal, our early marginalization uh not letting ourselves be carried away by shiny new things or Silicon

Valley level funding or any of those things. I think, keeping ourselves simple and open and collaborative is the best future I can hope for.

Alex: Wow. OK, thank you very much. This was a great closing statement. Glen, It has been a great pleasure to talk to you and I'm very grateful for your time and insights into your DH research, but also into ADHO.

Glen: Thank you very much, Alex. Thanks for being a good communicator and a good question asker and a good conversationalist. It's been a pleasure.

Alex: Thank you and also to the audience. Thank you very much for your time and your interest. See you soon.

Outro (Hannah): This podcast is a production of the Alliance of Digital Humanities Organizations (ADHO) who promotes and supports digital research and teaching across all arts and humanities disciplines. We act as a community-based advisory force and support excellence in research, publication, collaboration, and training. Learn more: https://adho.org/

—

More information about Glen Layne-Worthey:

- **University of Illinois Urbana-Champaign, School of Information Sciences:** https://ischool.illinois.edu/people/glen-layne-worthey
- **ORCID profile:** https://orcid.org/0000-0003-2785-0040

His Research:

- **HathiTrust Research Center:** https://www.hathitrust.org/about/research-center/
- **AOLIAN Network:** https://www.aeolian-network.net/

Selected Papers that were mentioned during the interview:

- Parulian, N.N., Dubnicek, R., Worthey, G., Evans, D.J., Walsh, J.A. and Downie, J.S. (2022). Uncovering Black Fantastic: Piloting A Word Feature Analysis a Jiang, M., Dubnicek, R. C., Worthey, G., Underwood, T., & Stephen Downie, J. (2022). A prototype Gutenberg-HathiTrust sentence-level parallel corpus for OCR error analysis: pilot investigations. In *JCDL 2022 - Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2022* (pp. 1-5). Article 45 (Proceedings of the ACM/IEEE Joint Conference on Digital Libraries). Association for Computing Machinery. https://doi.org/10.1145/3529372.3533298nd Machine Learning Approach for Genre Classification. Proceedings of the Association for Information Science and Technology, 59: 242-250. https://doi.org/10.1002/pra2.620

- Jiang, M., Dubnicek, R. C., Worthey, G., Underwood, T., & Stephen Downie, J. (2022). A prototype Gutenberg-HathiTrust sentence-level parallel corpus for OCR error analysis: pilot investigations. In *JCDL 2022 - Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2022* (pp. 1-5). Article 45 (Proceedings of the ACM/IEEE Joint Conference on Digital Libraries). Association for Computing Machinery. https://doi.org/10.1145/3529372.3533298

- Hu, Y., Layne-Worthey, G., Martaus, A., Downie, J.S., Diesner, J. (2023). Research with User-Generated Book Review Data: Legal and Ethical Pitfalls and Contextualized Mitigations. In Sserwanga, I., et al. *Information for a Better World: Normality, Virtuality, Physicality, Inclusivity.* iConference 2023. Lecture Notes in Computer Science, vol 13971. Springer, Cham. https://doi.org/10.1007/978-3-031-28035-1_13