Florida International University

FIU Digital Commons

Works of the FIU Libraries

FIU Libraries

5-23-2024

Speaking Volumes: Transforming Archival Access Through Speech-to-Text Transcription

Rebecca Bakker Florida International University

Follow this and additional works at: https://digitalcommons.fiu.edu/glworks



Part of the Archival Science Commons

Recommended Citation

Bakker, Rebecca, "Speaking Volumes: Transforming Archival Access Through Speech-to-Text Transcription" (2024). Works of the FIU Libraries. 148. https://digitalcommons.fiu.edu/glworks/148

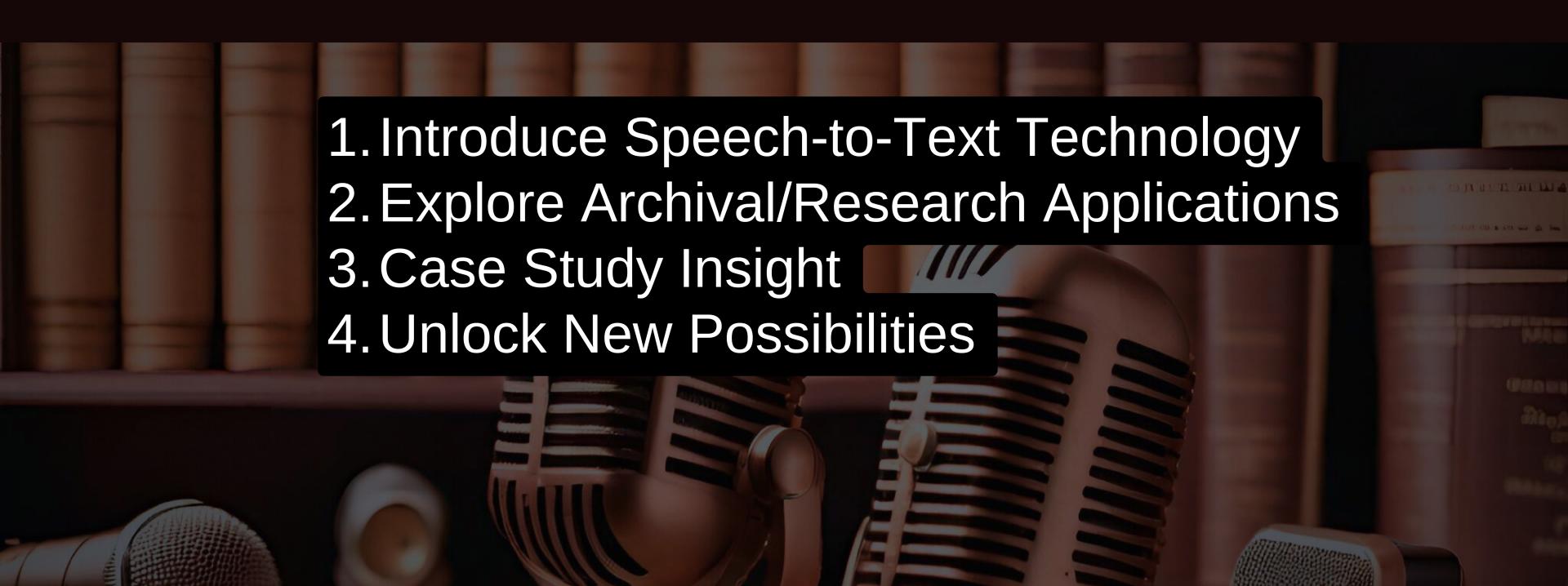
This work is brought to you for free and open access by the FIU Libraries at FIU Digital Commons. It has been accepted for inclusion in Works of the FIU Libraries by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.



Speaking Volumes: Transforming Archival Access Through Speech-to-Text Transcription

Rebecca Bakker rbakker@fiu.edu Digital Collections Center FIU Libraries

Workshop Objectives



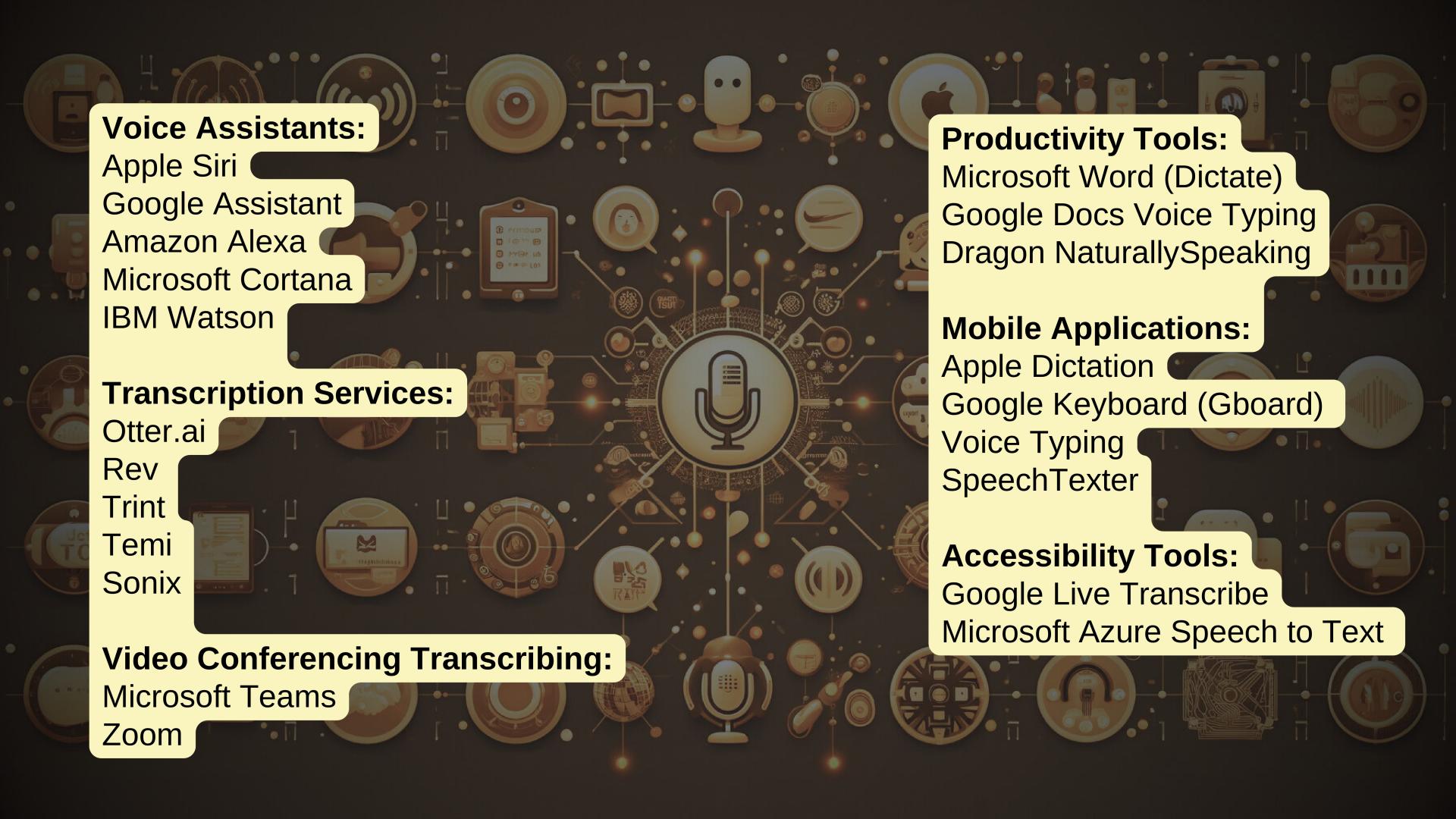
oral language processing Applications Recognized ccessibili Automatic Speech Recognition **Audio Capture** Assistants Capture

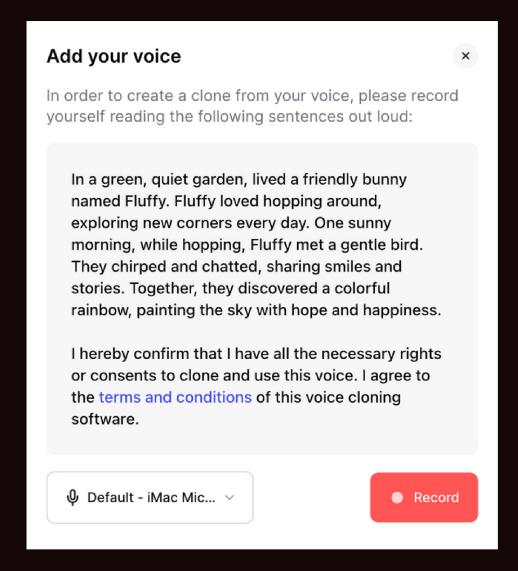
> audio transcription

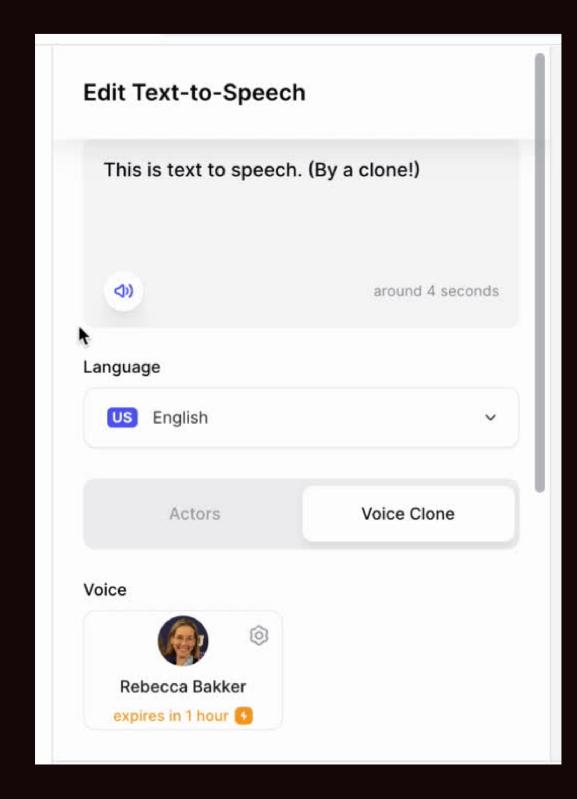
voice recognition

Speech-to-Text:

Technology that converts spoken language into written text by capturing audio, processing the sounds to identify words, and transcribing the recognized speech into text.







Text-to-speech:

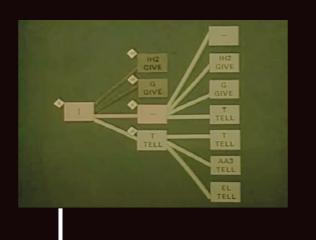
Technology that converts written text into spoken language by analyzing the text and generating synthetic speech that mimics human voice patterns.

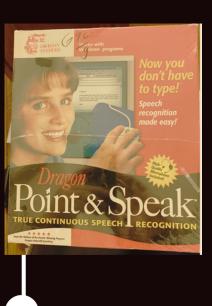


A brief history...









1952 - AUDREY (AUTOMATIC DIGIT RECOGNIZER)

Speech recognition technology developed by Bell Laboratories. It could recognize the digits zero to nine with a 90% accuracy rate.

1961 - THE IBM SHOEBOX

Could recognize and respond to 16 spoken words and digits (0-9). A significant milestone in the development of Natural Language Processing (NLP) and interactive voice-controlled computing.

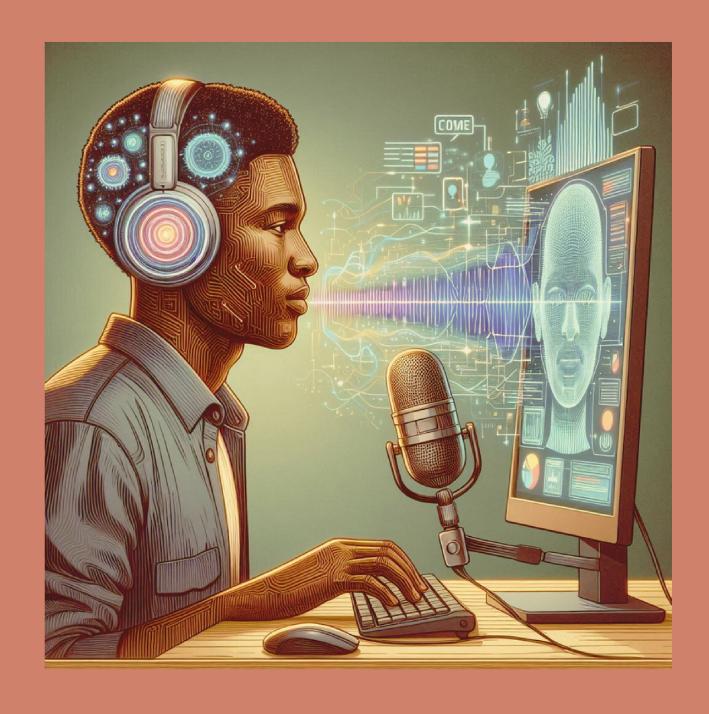
1976 - HARPY

Developed by Carnegie
Mellon University, Harpy
could recognize over 1,000
words with high accuracy,
marking a major
advancement in speech
recognition technology.

1990 - DRAGON DICTATE

Dragon Systems launched the first consumer speech recognition software, using discrete speech technology. Users could dictate text to their computers with notable accuracy by speaking one word at a time.

How Speech to Text Works



- * Audio Input
- * Preprocessing
- **Feature Extraction**
- * Acoustic Modeling
- * Language Modeling
- * Decoding
- Continuous Improvement

Feature	OCR Technology	Speech-to-Text Technology
Conversion	Converts printed text to digital text	Converts spoken words to digital text
Pattern Recognition	Identifies characters and words	Identifies phonemes and words
Preprocessing	Noise reduction, binarization, segmentation	Noise reduction, audio segmentation, feature extraction
Machine Learning and AI	Improves accuracy from text image datasets	Improves accuracy from spoken language datasets
Language and Context	Predicts and corrects text	Predicts and transcribes speech
Error Handling	Manages image quality and handwriting errors	Manages background noise and accent errors
Applications	Digitizes text, automates data entry	Transcribes speech, enables voice commands



Speech-To-Text Technology

Optical Character
Recognition
"text recognition"

Transcription of Audio/Video Content



https://digitalcommons.fiu.edu/dss-resources/22/



X Benefits:

- Accessibility
- Efficiency
- Searchability
- Engagement

* Challenges:

- Accuracy
- Bias
- Editing
- Privacy
- Context Misinterpretation
- Cost

Transcription of Audio/Video Content









Otter.ai













- * Purpose-Specific Transcription
- * Human Editor Involvement
- * Awareness of Biases
- * Cost and Time Considerations
- * Software Testing
- * Privacy

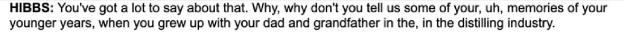


Interview with Parker Beam, August 28, 2013

Dixie Hibbs, Interviewer | 2013OH318 BIK 019







BEAM: Well, my grandfather, which is a brother to Jim Beam, which was, uh, the, uh, person I was named after, was my grandfather. Uh, his name was William Parker. So, in order to get the Parker in, and my father's name was Earl, so I am Earl Parker Beam. So, as, uh, as I recall, I remember going to Jim Beam's house when I was young with my grandfather and my father. And then they would go out on rides. And Dad would be the chauffeur. And they would, uh, ride around. And of course, they'd always take a bottle with them to, you know, to, to make the ride go more smoothly. So, in the meantime as they were riding, if they saw somebody that looked like would enjoy taking a drink out of their bottle, they would stop and, and pass the bottle around, so I've been told. I didn't always get to go on the rides though. (Hibbs laughs) So anyway, one of their favorite spots was, to ride was that they would go over to Fort Knox. And my father would drive and they'd see certain folks, soldiers walking on the, on the, uh, I guess, posts there. And they would offer them a drink and make sure they recognized that it, it was Jim Beam Bourbon. (Hibbs laughs) So that's some of the tales that, that I remember Dad talking about, uh, and Grandfather about Jim Beam. Uh, then, uh, I guess getting back to, at my part in Heaven Hill, I followed my dad around there. I believe he came from Beam to Heaven Hill in 1946. So, uh, that, that time, I was probably five years old. So, I've been walking around here on this property ever since I was big enough to walk--







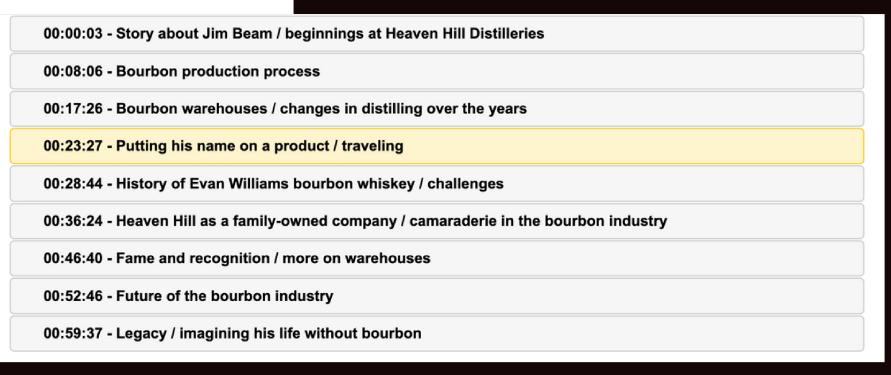
00:01:00

00:02:00

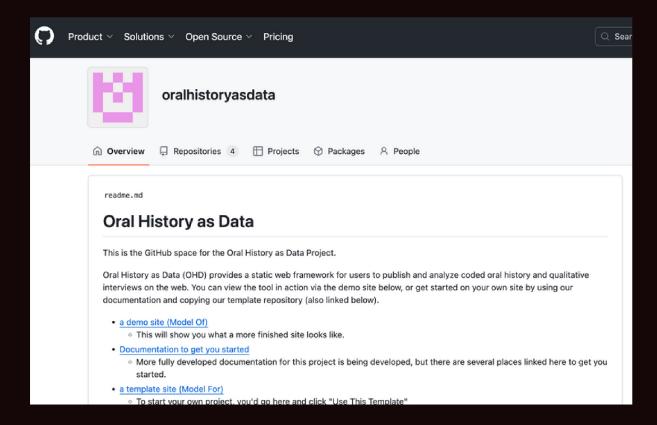
Interview with Parker Beam, August 28,

Louie B. Nunn Center for Oral History, University of Kentucky Libraries

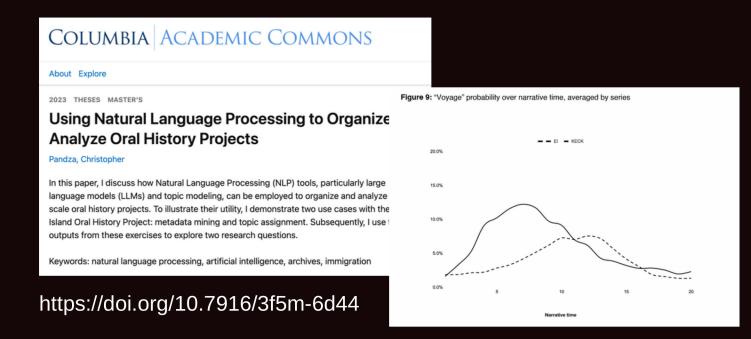
nunncenter.net



Oral History Research Trends

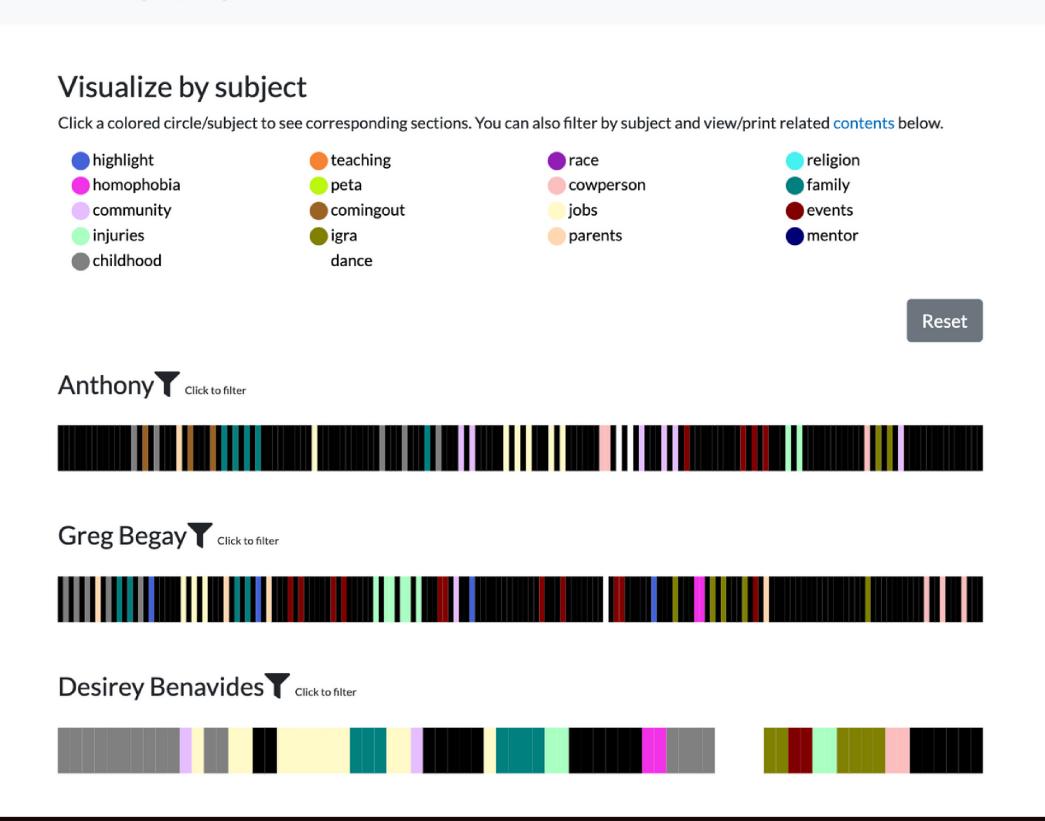


https://github.com/oralhistoryasdata



Voices of Gay Rodeo

Home About Interviews Visualizations



https://www.voicesofgayrodeo.com/visualizations.html



Oral History Association Virtual Symposium July 15-19, 2024 https://oralhistory.org/ai/

THURSDAY, JULY 18TH 10AM-12PM (CDT)

07. Making A Case for the Humans: The Ethics and Politics of Al in Oral History (Roundtable) Chair: Anna Sheftel

Participants: Crystal Mun-hye Baik, Doug Boyd, Anna Sheftel, Graham Smith, Katrina Srigley, & Stacey Zembrzycki

2PM-4PM (CDT)

08. Al's Impact on Oral History Indexing (Panel)
Chair: Steven Sielaff

Emerging Directions in Interview Processing and Collection Management: Al Auto-Transcription, Timecode Indexing, and "Multi-Dimensional Content" in Oral History Michael Frisch

Oral History Indexing through the Lens of AI Douglas Lambert

Training AI through Oral History Workflow Daniel McMullan

FRIDAY, JULY 19TH 10AM-12PM (CDT)

09. Designing Switchboard: An Open-Source Chatbot That Collects Trans* Oral Histories (Roundtable)

Chair: Molly Merryman

Participants: Alejandro T. Acierto, Ava Aviva Avnisan, Molly Merryman, Dana Middleton, & Dan Paz

2PM-4PM (CDT)

10. Old Ways, New Ways: Archiving the Heirloom Gardens Project Using Al (Roundtable) Chair: Tessa Lowinske Desmond

Participants: Hassan Adeeb, Tessa Lowinske Desmond, & LuAnna Nesbitt

MONDAY, JULY 15TH 10AM-12PM (CDT)

01. Ethical Considerations in the Age of Al (Panel)
Chair: Jennifer A. Cramer

From Contextualized Interview to Anonymous Data: Making Meaning with Oral Histories in the Age of Big Data

Mary Larson

In Support of Putting Oral History Narratives "In the Vault" - Do Humanists Need to Take a Page from "Preppers" in the Age of Artificial Intelligence? Jane Collings

World Englishes and Whisper: Testing AI and Pedagogically Reframing Oral History Transcription Workflows

Aieshah Arif, Lauren Kata, & Sweta Kumari

2PM-4PM (CDT)

02. From Capture to Engagement: Experiments in Using Al for Indexing, Named Entity Recognition, and More (Roundtable)

Chair: Zack Ellis

Participants: Bram Adams, Zack Ellis & William Mattingly

TUESDAY, JULY 16TH 10AM-12PM (CDT)

03. Al as Human Substitutes in the Oral History Process (Panel)

Chair: James Fowler

Envisioning AI as an Oral History Interviewer: The Educated Youth Oral History Project Haina Jin & Zhao Yigong

Probing a Community-Based Conversational Storytelling Agent to Document Digital Stories of Housing Insecurity Brett Halperin

How will AI Enable Oral History to Make it Possible to Converse with Our Dead Relatives? Mary Gordon

TUESDAY, JULY 16TH (CONT.) 2PM-4PM (CDT)

04. Al's Role in Processing Oral Histories (Panel) Chair: Patrick Daglaris

Curating Oral History with Al Christopher Pandza

Creating a Hybrid Workflow: Describing (and Improving) User-Generated Content Using Al Tools Virginia Millington

Powered by Al: Crafting a GPT Prototype for Oral History

Juliana Nykolaiszyn

WEDNESDAY, JULY 17TH 10AM-12PM (CDT)

05. Al and Black Voices (Roundtable) Chair: Sara Baum

Participants: Sara Baum, Destiny Crawford, Zack Ellis, Alissa Rae Funderburk, Michael Sesling, & Hanassa Wicke

2PM-4PM (CDT)

06. Using Whisper AI for Transcription to Improve Oral History Workflows (Roundtable)

Chair: Tamar Carroll

Participants: Kirk M. Anne, Benjamin Myers, Rebekah Walker, Emmarose Tabin, Caitlin McCabe, Liz Call, & Landyn Hatch

Transcribing

- Recorded audio/video to text
- Verbatim transcripts
- Closed captions
- Archival and accessibility

Convert speech to text

Dictation

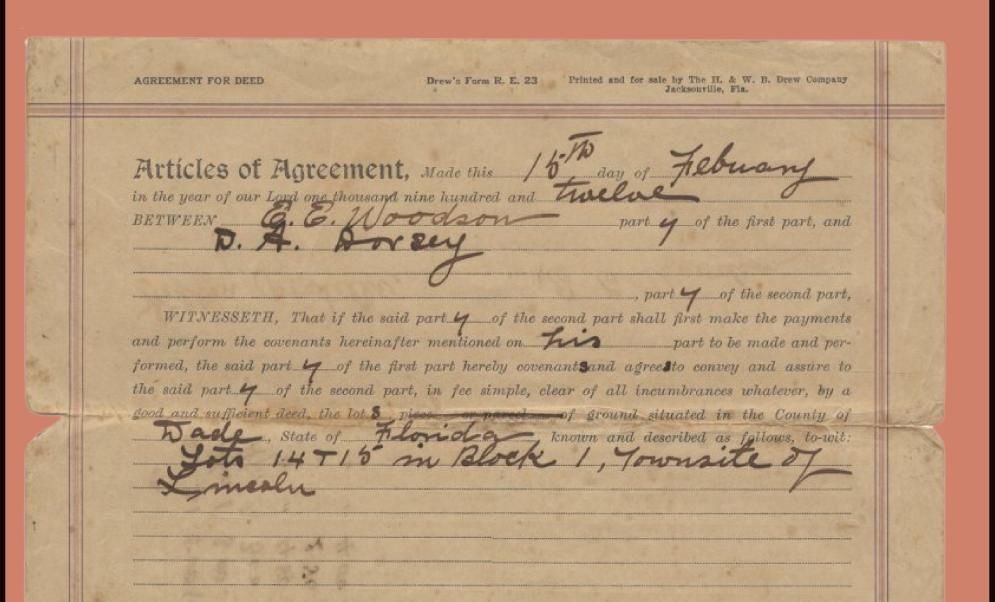
- Real-time speech to text
- Composing documents/emails
- Virtual assistant commands
- Speed and convenience
- Hands-free writing

Voice Writing



- Listen to proceedings
- Repeat spoken words into microphone
- Special Commands
- Real-time text conversion by software
- Review and edit for accuracy

Dictation for Mixed-Media Texts





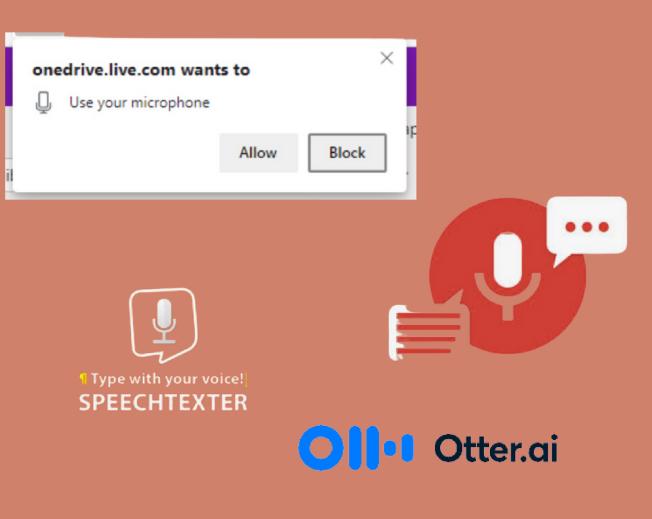
- Accessibility
- Efficiency
- Accuracy
- Close Reading
- Text analysis

* Challenges:

- Accuracy Variability
- Editing Required
- Training and Familiarity
- Software limitations

Dictation Software





Listening



- Easily available
- Advanced Development
- Voice Commands
- Adaptive Learning
- Hands-Free Operation

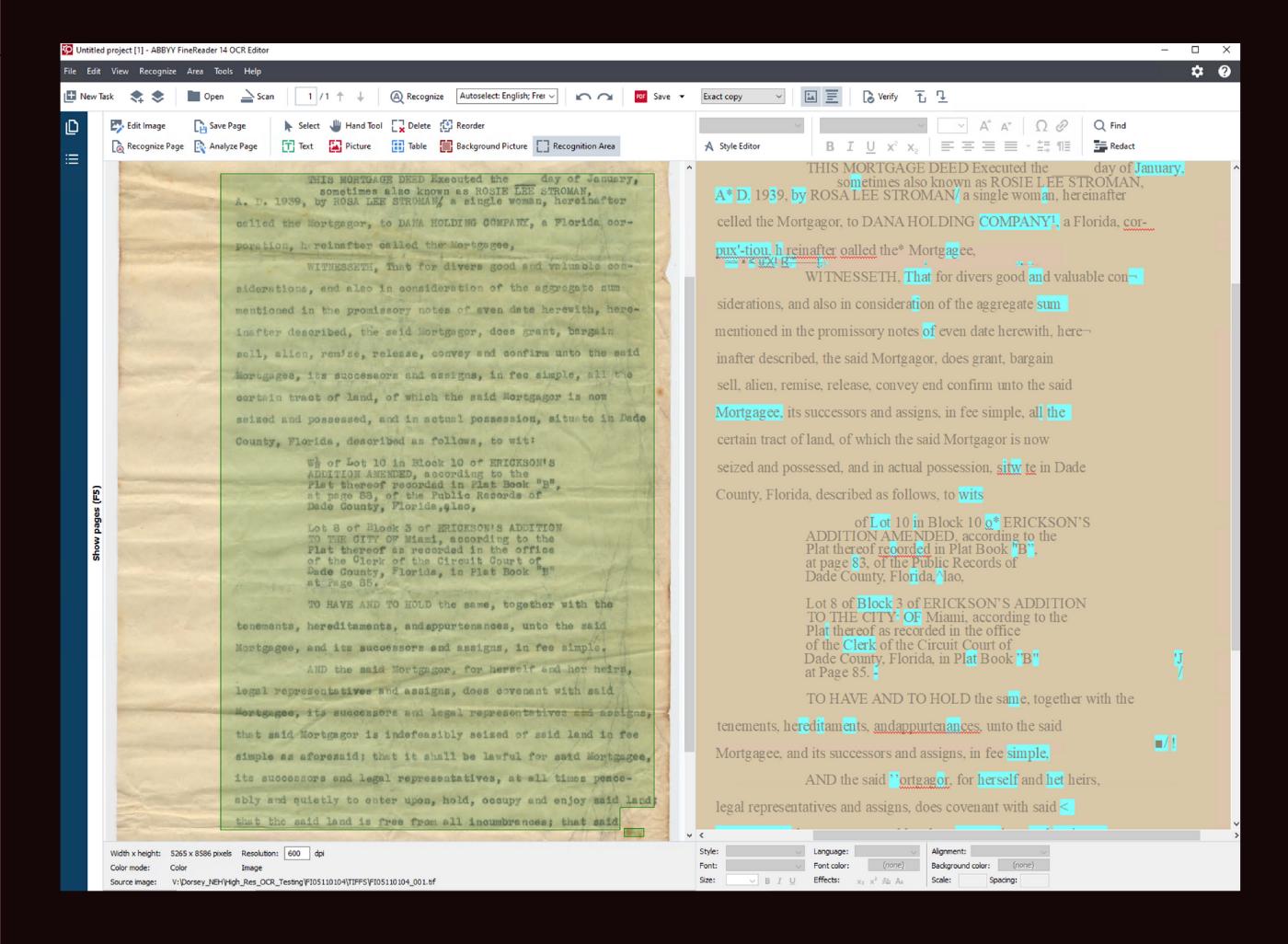
The bad:

- Learning Curve
- Bias in Recognition
- Accent, Language and Dialect Variability

X Testing:

- Accuracy
- User friendliness
- Integration and Compatibility

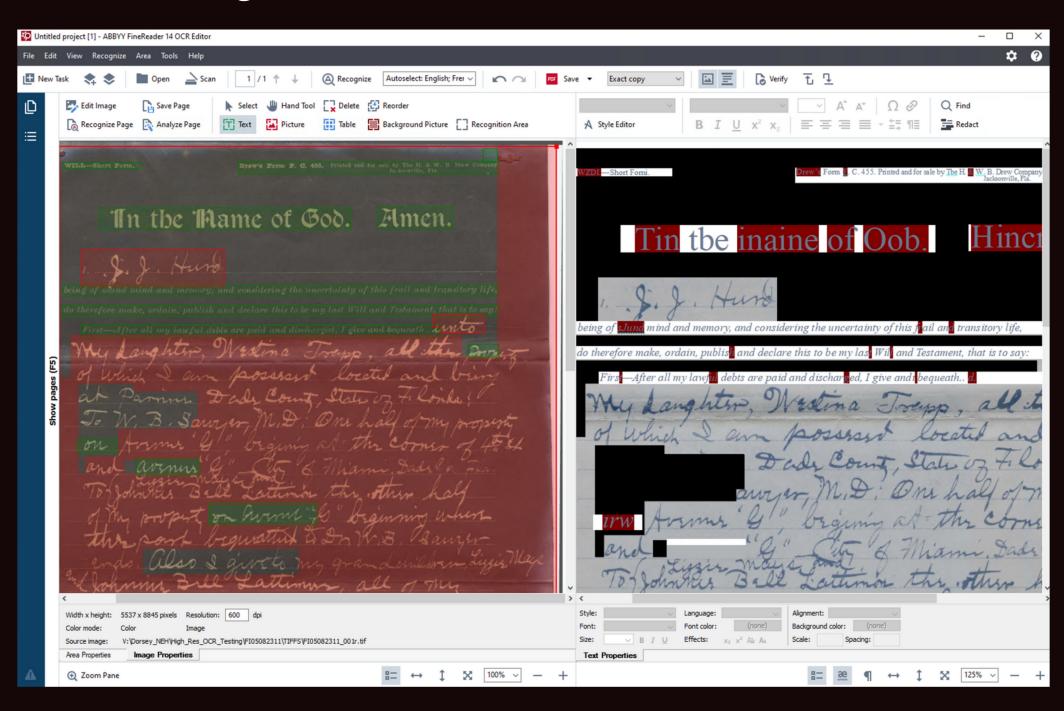
Dictation VS. OCR



Original

OCR'd

Dictation Transcribed



Will- Short Form,

Drew's Form P.C. 455. Printed and for sale by the H. & W. B. Drew Company Jacksonville, Fla. In the name of God. Amen

I. J. J. Hurd

being of sound mind and memory and considering the uncertainty of this frail and transitory life, do therefore make, ordain, publish and declare this to be my last Will and Testament, that is to say:

First - After all my lawful debts are paid and discharged, I give and bequeath unto my daughter, Westina Trapp, all the property of which I am possessed located and being at Perrine Dade County, State of Florida;

To W. B. Sawyer, M.D: One half of my property on Avenue "G", beginning at the corner of 4th St and Ave "G" city of Miami, Dade Co., Fla.

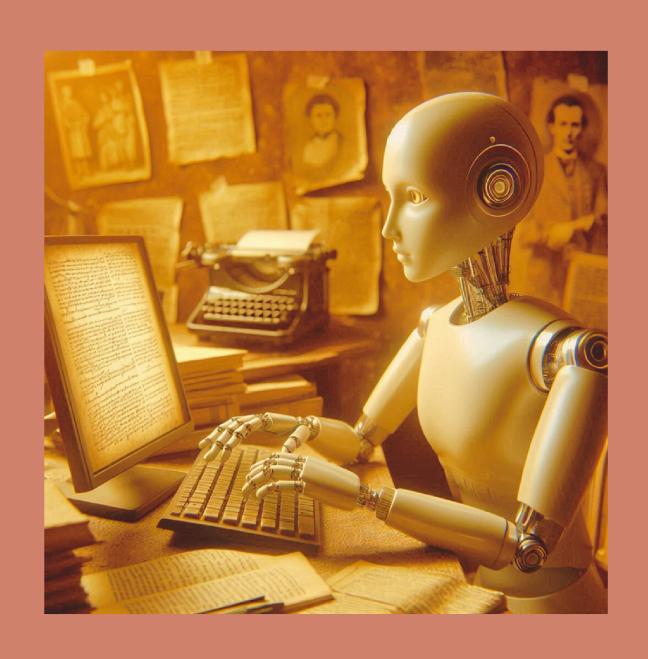
To Lizzie May and Johnny Bell Lattimer the other half of my property on avenue "G" beginning where the pond bequeathed to doctor W. B. Sawyer ends. Also I give to my grandchildren Lizzie Mae and Johnny Ball Latimer all of my property situated in the part of Miami known as Baldwin's addition.

Likewise I make, constitute an appoint Alan Stokes and John R Scott junior to be executors of this my last Will and Testament, hereby revoking all former Wills by me made. In Witness Whereof, I have hereunto subscribed my name and affixed my seal the 29th day of November in the year of our Lord one thousand nine hundred and sixteen.

J. J. Hurd his mark seal

The above partly written and printed instrument was subscribed by the said J. J. Hurd in our presence and acknowledged by him to each of us; and he at the same time declared the above instrument so subscribed to be his last Will and Testament; and we at his request, have signed our names as witness hereto in his presence and in the presence of each other, and written opposite our names our respective places of residence.

Dictation vs. Typing





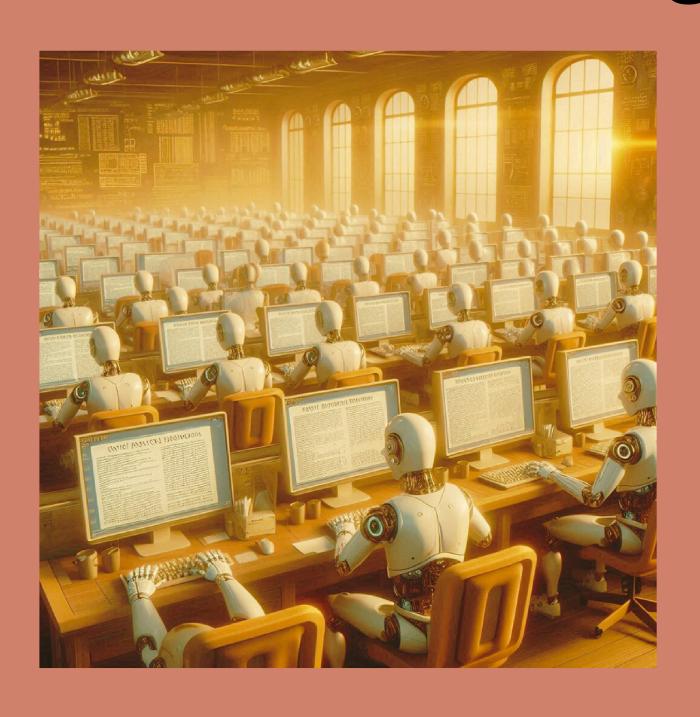
Speed and Efficiency

- Speech Read Aloud: 183
 words per minute (wpm)
- Typing: 38-40 wpm

* Accuracy

- Speech Recognition
 Software: Improved
 accuracy with modern tools
- Manual Typing: Slower, prone to human error

Dictation vs. Crowdsourcing



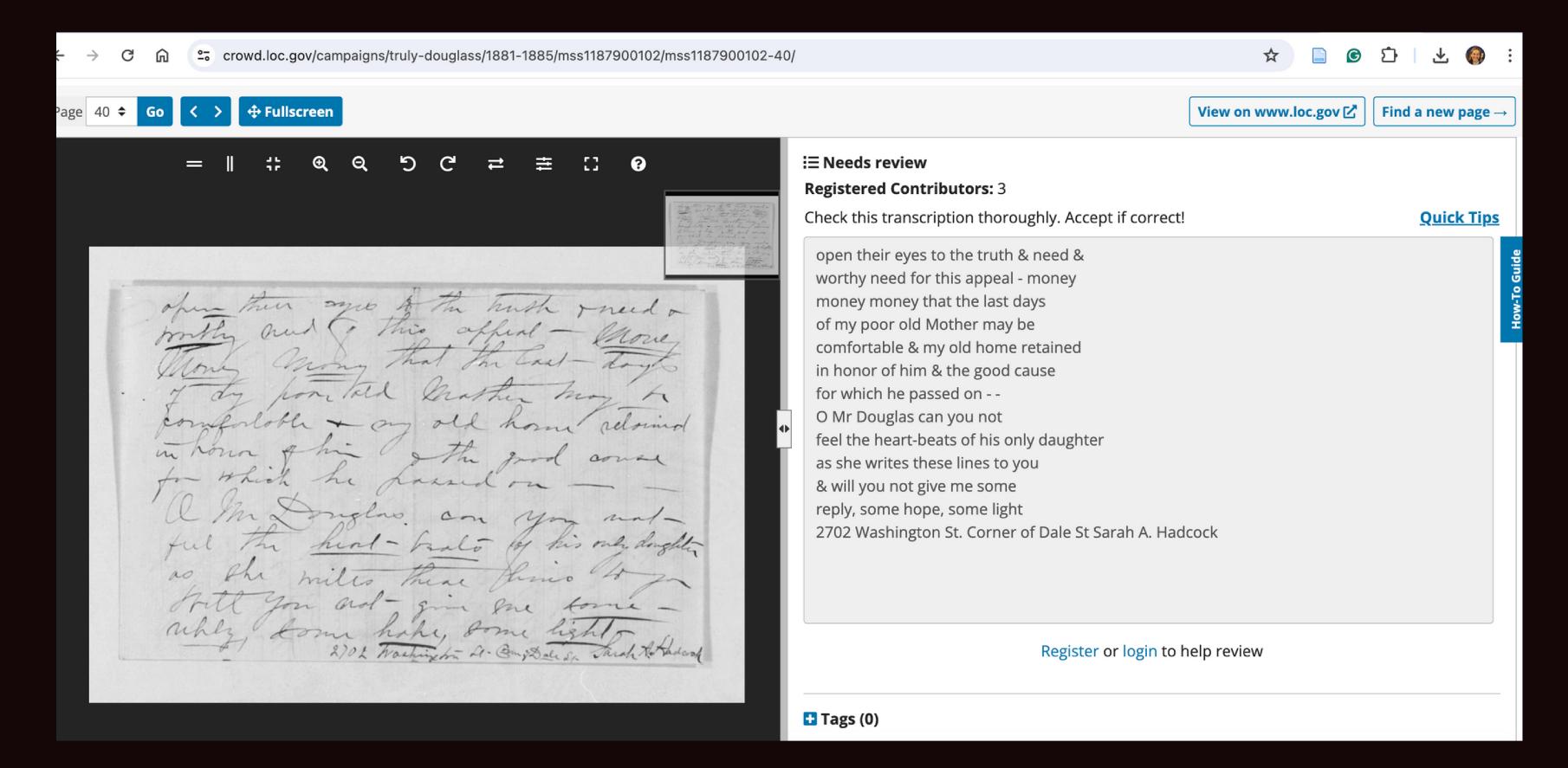
Project Management & Oversight

- Requires significant planning and coordination
- Maintaining consistent quality is challenging
- Need a platform for transcribed submissions

K Ethical Considerations

Fair labor practices

Crowdsourced Transcription



Case study: Dana Dorsey Project

www.go.fiu.edu/dorsey

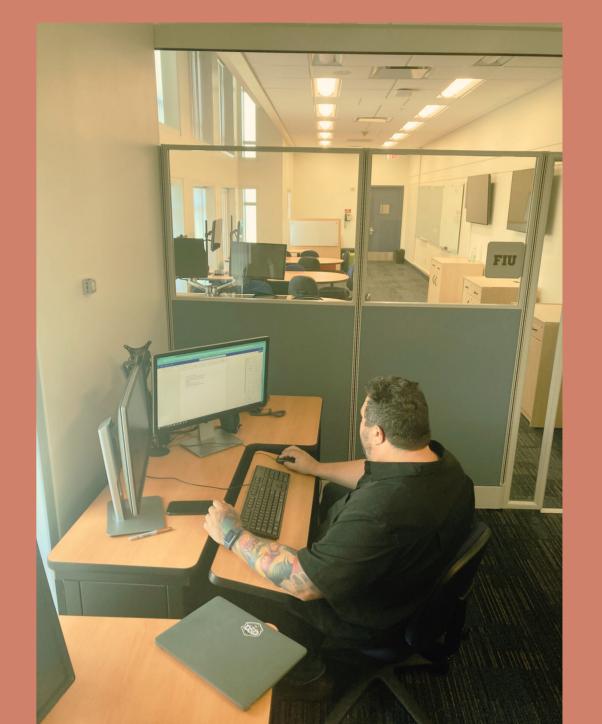


The Dana A. Dorsey Papers

Enhancing Access and Research Possibilities through Critical Engagement with Historical Data **Dana A. Dorsey**: "First Black millionaire in Miami," a successful businessman (1868-1940). **NEH Project Goals**: Enhance access to Dana A. Dorsey's papers, highlight community building, make stories of Black lives more discoverable, and understand the roles of civic groups, trade groups, and community alliances.

- Collection consists of warranty deeds, mortgages, legal documents, and correspondence.
- Documents date from around 1900 through 1940.

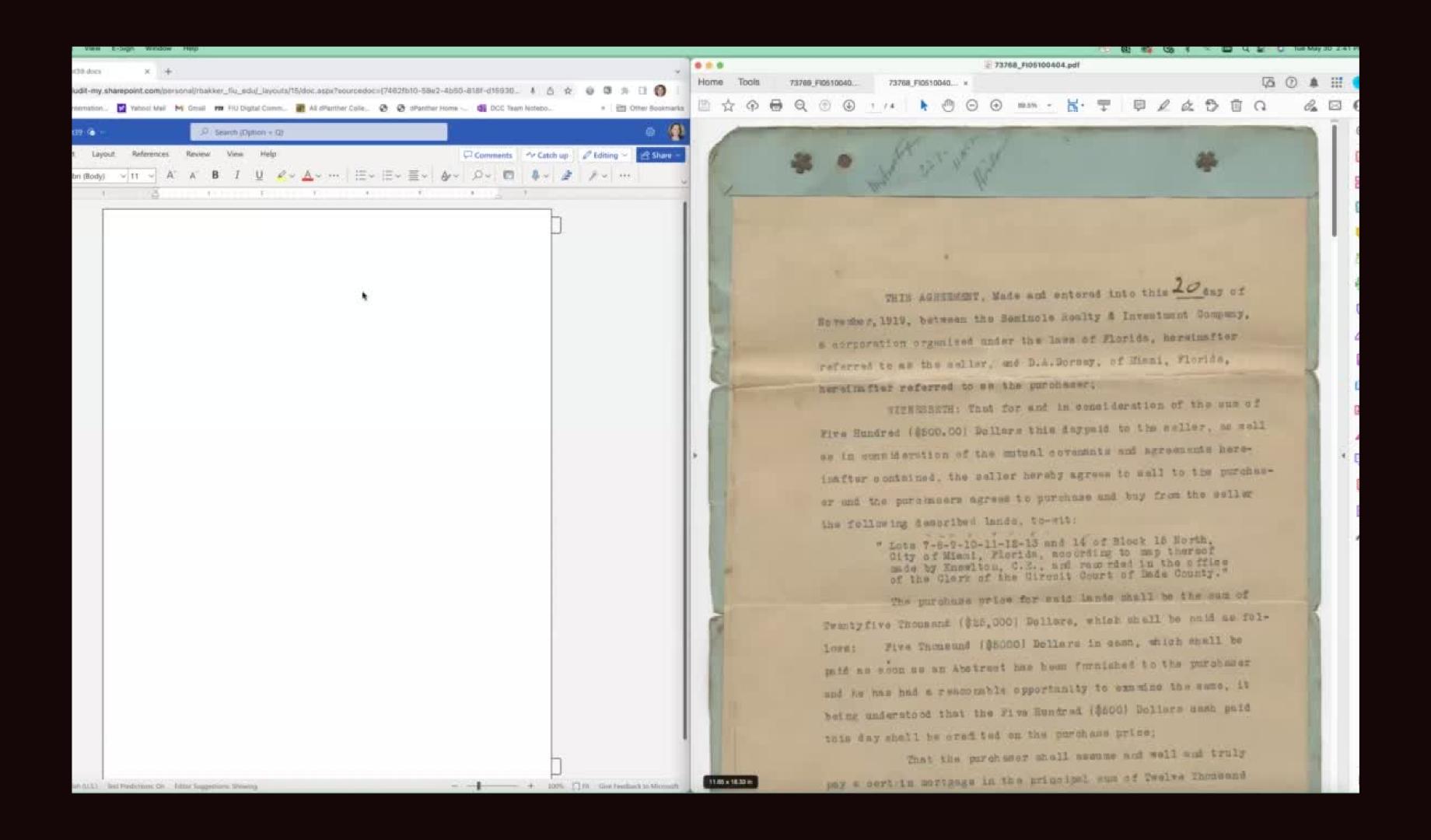
Dana Dorsey Project



Reasons for Using Dictation

- Replicability
- Format for Text Analysis &
 Discoverability
- Simultaneous Data Extraction

Juan Lopez, project intern, in the Digital Scholar Studio



https://crowd.loc.gov/get-started/how-to-transcribe/

- Transcription
 - Hyphen complete the word on the same line then start the next word on the next line.
 - Signature
 - Marginalia -
 - * * [handwritten]*February 2*[/handwritten]
 - Stamp/Notarized -
 - [stamped]State of Florida[/stamped]
 - [seal] County [/seal]
 - Using brackets in transcription
 - [unclear][/unclear]
 - [deletion][/deletion]
 - Use is something is crossed out completely
 - [insertion][/insertion]
 - It looks like "insertion" isn't needed for these documents this
 does not help with full text search or use of the data
 - [correction] [/correction]
 - [handwritten][/handwritten]
 - "handwritten" is not necessary unless it written in as an addendum
 - not as part of filling out the document
 - o Always close out the tags e.g. [tag][/tag]
- Data Extraction
 - o Add People names
 - make sure to include FI#
 - Make sure the name is not on the list already
 - Separate initials with a space (e.g. J. L. Rogers)
 - Add Entity names
 - make sure to include FI#
 - If entity name is very similar but not exactly the same add it
- If you come across something unusual add to the "dorsey_metadata" file
 - Add to the Notes Quality Control and Issues
- For tables write

[see table in original document]

Transcription format [Arial 12pt]

Link to Item: [hyperlink]

FI#: FI########

Title: Title

Typing Transcription Tips



BY THE PEOPLE

About Campaigns ✓ Resources Discuss How-To ✓

Login Register

Home / Get Started / Transcription: Basic Rules

Transcription: Basic Rules

Instructions

Get Started

Transcription: Basic Rules

Transcription: Things to Avoid

Transcription: Printed Text & Images

Transcription: Unusual Text

How to Review

How to Tag

Instrucciones en español

Text order

Transcribe text in the order it appears on the page. If you're unsure, transcribe the text how it would make sense to read aloud.

Preserve original spelling, grammar, and punctuation

Transcribe original text spelling, grammar, and punctuation. Type what you see. Do not paraphrase.

Dashes and other punctuation can be a little unusual before the mid-20th century. Make your best guess on whether something is best represented by an en dash, em dash, equals sign, period, etc.

If a misspelling will impact search of the document, you can tag the page with the correct spelling. **Example:**"Abe Linkin" - Transcribe as spelled and tag with "Abraham Lincoln"

Line breaks

Preserve line breaks. Line breaks make it easier for someone to review your transcription. Sometimes text in the transcription box spills over two lines in the transcription box. If you don't hit enter, no line break is recorded.

Dictation Tips

1. Speak Naturally

- Normal Pace: Speak at a clear and normal pace for best results.
- Microphone Use: Use a microphone if you speak quietly or are in a loud environment for better accuracy.

2. Real-Time Corrections

- Keyboard Handy: Keep a hand on the keyboard to fix small mistakes on the go.
- Cursor Control: Ensure your cursor returns to the right spot after corrections to continue dictation smoothly.

3. Commands and Navigation

- Next Line: Use the command "next line" or press "enter" to move to the next line.
- Delete: Use the command "delete" to erase the most recent word.
- Punctuation: Speak punctuation commands like "open parentheses" and "closed parentheses" for accuracy.

Dictation Tips

- 4. Handling Interruptions
 - Stay on Tab: Ensure you remain on the text-to-speech tab to avoid interruptions.
 - Chunk Reading: Read in chunks and use zoom features or sticky notes to keep track.
- 5. Patience and Adaptation
 - Patience: Be patient as the software learns your speech patterns.
 - Repeat if Necessary: If a word isn't picked up, repeat it clearly to help the software learn.

Patience is key, text to speech is learning from you as well

Clean, Uniform Text Files

	LEASE HEFTY'S FORM R. E. 29 THE HEET PHILSS, INC., MIAMI
	This Indenture, Made this lat day obecember, 35, between, of the one part
	witnesseth: That the said D. A. Dorsey Lessor do by
	these presents lease unto the said H. E. 3. Reeves, Lessee One Store Room known as 1112 N.W.3rd Ave.
	Lessee to furnish water and light.
	TO HAVE AND TO HOLD the premises aforesaid unto the said lesses from the day of December , 19 35 , for the term of Two Years (24 months)
	them sext ensuing, the said lessee yielding and paying for the same unto the said lessor, His heirs and assigns, the rent or sum of\$600.00Payable at \$25.00
-	per month, on the first day of each and every month until the total amount shall have been paid.
The state of the s	
The state of the s	IT IS AGREED between the parties hereto that should this rent at any time remain unpaid for five days after the same shall be due and payable, the said lessor may, at His option, then consider said lessee as tenant at will, and re-enter upon and repossess All of the said premises. And should the said lessee at any time permit the said rent to be in
	arrears and unpaid for five days after the same shall be due and payable, then the said lessee_ covenant_ and agree_ that in case of such default the entire rent for the term for which said premises are leased shall at once become due and payable, and may be recovered forthwith by distress or otherwise; and in all proceedings under this lease for recovery of rent in arrears, whether by distress or other action at law, the said lessee_ for His_ heirs, executors and administrators, hereby waive_ the benefit of all the exemption laws, any law to the contrary
-	notwithstanding. AND IT IS FURTHER AGREED that should the property of the lessee be removed from the aforesaid premises, the lessor hereby authorized, at any time within ninety days thereafter, to enter upon the same wherever found, and to seize and to sell

Link to Item: http://dpanther.fiu.edu/dpService/dpPurlService/purl/FI05082910/00001

FI#: FI05082910

Title: Lease between Dana A. Dorsey and H. E. S. Reeves

Lease Hefty's Form R. E. 29 The Hefty Press, Inc. Miami This indenture, Made this 1st day of December, 1935

between D. A. Dorsey, of the one part

and H. E. S. Reeves, of the other part.

Witnesseth: That the said D. A. Dorsey, Lessor do _ by

these presents lease unto the said H. E. S. Reeves, Lessee

One store room known as 1112 N.W. 3rd Ave.

Lessee to furnish water and light.

To have and to hold the premises aforesaid unto the sad leasee from the 1st

day of December, 1935, for the term of two years (24 months)

the next ensuing, the said lessee yielding and paying for the same unto the said lessor

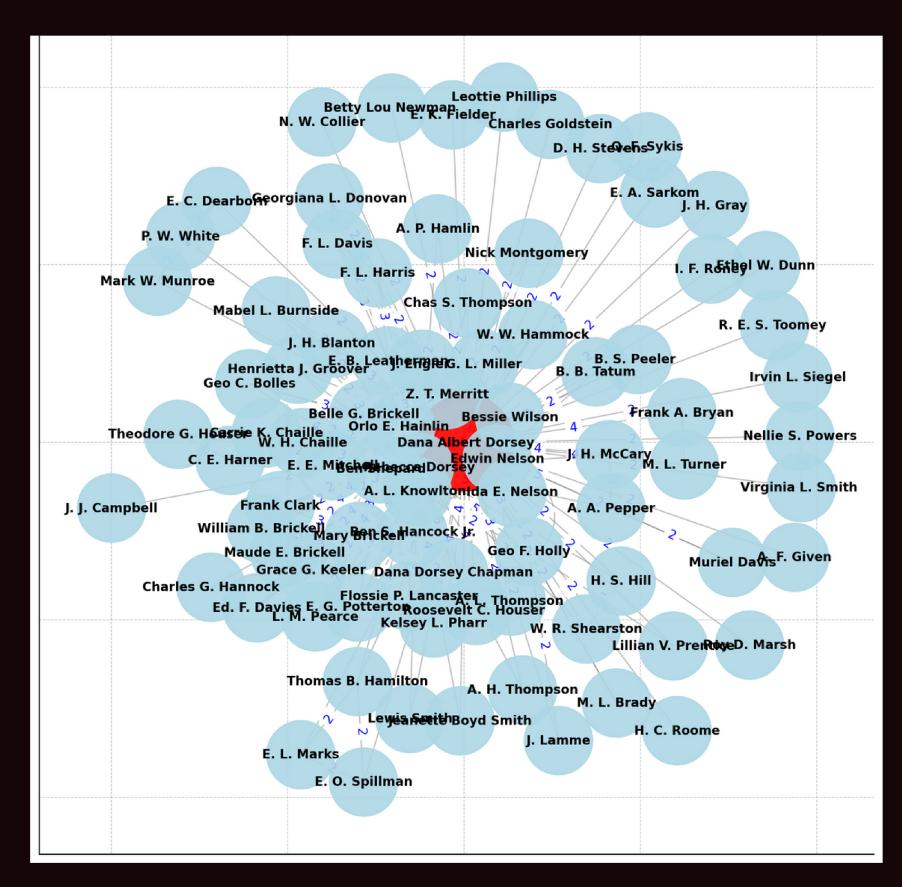
his heirs and assigns, the rent or sum of \$600- - Payable at \$25.00

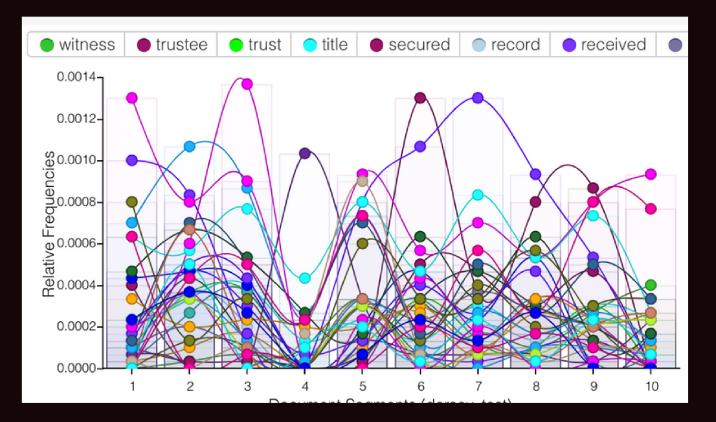
per month, on the first day of each and every month until the

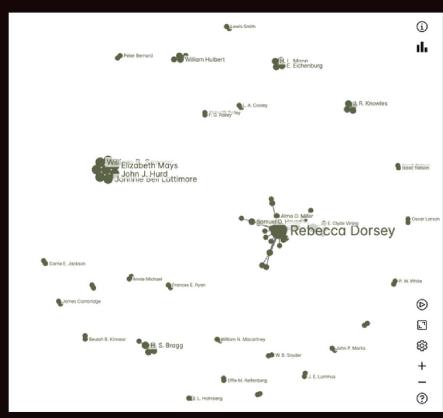
total amount shall have been paid.

It is agreed between the parties hereto that should this rent at any time remain unpaid for five days after the same shall be due and payable, the said lessor may; at his option then consider said lessee as tenant at wil, and re-enter upon and repossess all of the said premises. And should the said lessee at anytime permit the set rent to be in arrears and unpaid for five days after the same shall be due and payable, then the said lessee covenant and agree that in case of such default the entire rent for the term for which said premises are leased shall at once become due and payable, and may be recovered forthwith by distress or otherwise; and in all proceedings under this lease for recovery of rent in arrears, whether by distress or other action at law, said lessee for his heirs, executors and administrators, hereby waive the benefit of all the exemption laws, any law to the contrary notwithstanding. And it's further agreed it should the property of the lessee be removed from the aforesaid premises, the lessor hereby authorized, at any time within ninety days thereafter, to enter upon the same wherever found, and to seize and to sell so much thereof as will fully satisfy the said lessor, for all arrears or rents then due or to become due, as above provided, and the cost thereon; and then the sales of said property lessor is to have the rights and privileges granted under this lease, especially that waiving the benefits of the exemption laws, so much so as if no removal of the lessee's property had been made. It is also expressly understood that the said premises shall not be underlet or rented without the lessor's written consent to the transfer of this lease or such underletting. On the expiration of this lease, the property is to be given up in as good order in all respects as it now is,

Data, text, and relational analysis







Thank you for coming!

Rebecca Bakker rbakker@fiu.edu Digital Collections Center FIU Libraries



References

- Library of Congress. (n.d.). How to transcribe. By the People. Retrieved May 22, 2024, from https://crowd.loc.gov/get-started/how-to-transcribe/
- López, Gustavo, Luis Quesada, and Luis A Guerrero. "Alexa vs. Siri vs. Cortana vs. Google Assistant: A Comparison of Speech-Based Natural User Interfaces." Advances in Human Factors and Systems Interaction. Cham: Springer International Publishing. 241–250. Web.
- Sheftel, Anna, and Stacey Zembrzycki. "Slowing Down to Listen in the Digital Age: How New Technology Is Changing Oral History Practice." The Oral history review 44.1 (2017): 94–112. https://www.tandfonline.com/doi/full/10.1093/ohr/ohx016
- Shtern, J. (2020, February 11). A brief history of ASR (Automatic Speech Recognition). Medium. Retrieved from https://medium.com/descript/a-brief-history-of-asr-automatic-speech-recognition-b8f338d4c0e5
- Sydell, L. (2016, August 24). Voice recognition software finally beats humans at typing, study finds. NPR. https://www.npr.org/sections/alltechconsidered/2016/08/24/491156218/voice-recognition-software-finally-beats-humans-at-typing-study-finds
- Trower, Shelley. Sound Writing: Voices, Authors, and Readers of Oral History. 1st ed. United Kingdom: Oxford University Press, 2023. Web.
- Yu, Dong. Automatic Speech Recognition: A Deep Learning Approach. 2015th ed. London: Springer Nature, 2014. Web.