Submitted on: 20.06.2022

Title of the Satellite Meeting: *New Horizons in Artificial Intelligence in Libraries*
Date*: July 21-22, 2022*
Location: *National University of Ireland, Galway, Ireland.*

Steps Towards Building Library AI Infrastructures: Research Data Repositories, Scholarly Research Ecosystems and AI Scaffolding

Raymond Uzwyshyn
Division of Library Collections and Digital Services,
Texas State University, San Marcos, USA.
E-mail address: ruzwyshyn@gmail.com

ABSTRACT:

Artificial Intelligence possibilities for Deep Learning, machine learning, neural nets and natural language processing present fascinating new AI library service areas. Most of these areas will be integrated into traditional academic library 'information' and 'digital' literacy programs and university research environment to enable research faculty, students and library staff. Most university faculty, graduate students and library staff working outside of Computer Science disciplines will require help infrastructures to enable their data and research towards new AI possibilities. This research overviews methodologies and infrastructures for building new AI services within the 'third interdisciplinary space' of the academic library. A library is a very suitable space to enable these new 'algorithmic literacy' services. The work utilizes the pragmatic steps taken by Texas State University Libraries to set up good foundations. Data-centred steps for setting up digital scholarly research ecosystem infrastructure are reviewed. Setting needed data-centred groundwork for library AI services enables research, data and media towards wider global online AI possibilities. Library AI external scholarly communications infrastructures and services are discussed as well as educational methodologies involving incremental steps for foundational AI scaffolding. Bootstrapping tools build on present systems and allow for the later enablement of future AI insights. Pathways from data collection to data cleaning, analytics and data visualization to AI applications are clarified. Preliminary focused steps needed are forwarded to move library staff, research faculty and graduate students towards these new AI possibilities. Data-centred ecosystems, retooling and building on present library staff expertise are reviewed. Data research repositories, algorithmic and programmatic literacy are recommended for later AI possibilities. Preliminary AI library working groups and R&D prototype methodologies for scaling up future library services and human resource infrastructures are considered. Recommended emergent pathways are prescribed to create library AI infrastructures to better prepare for a currently occurring global AI paradigm shift.

Keywords: Artificial Intelligence, Deep Learning, Data Research Repositories, Academic Libraries, Research Libraries

1 INTRODUCTION

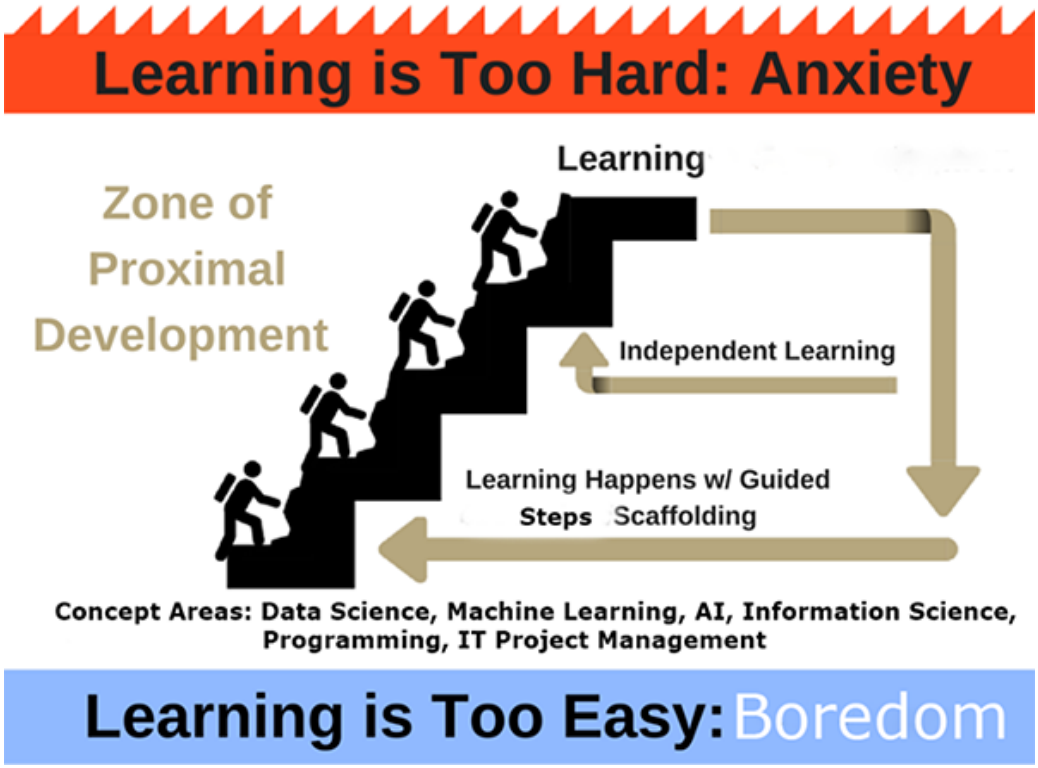
Deep learning, machine learning, neural nets and natural language processing are fascinating new areas of AI. Most university research faculty, graduate students and library staff work outside Computer Science AI disciplines. A majority of the university community don't know where to begin with enabling their research data with new AI paradigms. This research overviews pragmatic steps taken by Texas State University libraries to set up good foundations for AI possibilities. These steps include data research repository foundations, digital scholarly research ecosystem infrastructure,s and relevant tools and services to begin to set important groundwork for research, data and media towards new arising AI possibilities .

Texas State University is a US Carnegie Higher Research Activity University and Libraries are designated an Association of Research Libraries member. With approximately 40,000 students, the university is a comprehensive research and teaching university across undergraduate, graduate and doctoral levels. This research may be most productively utilized and adapted by any similar-sized medium or larger academic library or research institution thinking about beginning AI programs.

Texas State University Libraries, https://www.library.txstate.edu/

Developing AI-related library scholarly services for research faculty, graduate students and library staff begins with education and incremental steps to enable new insights and knowledge. Research data gathering and experimental data is plentiful at academic research institutions and in libraries. This research may now be corralled towards many new, productive and innovative AI pathways. This article outlines preliminary steps needed to enable library staff, research faculty and graduate students towards these new possibilities. Technological ecosystems, new hires and retooling possibilities for these new AI infrastructures for academic libraries will be discussed. How research centers may strategically move into these new areas will be reviewed. Methodologies, challenges, scholarly communications models and preliminary infrastructures begin with data research repositories, scholarly research ecosystems and algorithmic literacy programs . These all allow bootstrapping towards AI possibilities and setting strong foundations for larger successful new millennia AI library programs.

2 EDUCATIONAL STEPS AND SCAFFOLDING



Educational Scaffolding and Steps Towards Learning, Warren, 2021.[1]

To build any successful library AI program, educational steps and scaffolding are needed. Because the learning curve for AI is steep, staff education should be thought about in detail by library managers and administration. Artificial Intelligence combines concept areas of data collection, data science, programming, information science and IT project management. The larger goals of AI staff professional development and algorithmic literacy programs are also explicitly not to turn disciplinary research faculty, graduate students and library staff into AI experts. Education here develops a more sophisticated vocabulary towards AI programmatic literacy and larger conversations so that the university's larger learning community is conversant and enabled with the language of AI paradigms. Later they will be able to converse knowledgably on project possibilities and work with AI engineers and Ph.D.'s.

3 AI PARADIGMS AND ORIGINS

AI has many origins - each with unique algorithmic paradigms. Some paradigms are better suited than others to solve particular problem areas. It is best for any algorithmic literacy program to begin generally. Introduce university research faculty, graduate students and library staff to the wider field before delving more deeply into particular areas. There are many good introductory texts, documentaries, online courses and Youtube videos online to inspire before beginning the linear algebra and calculus of Deep Learning's Back Propogate

[1] See: https://goodsensorylearning.com/blogs/news/scaffolding-development

(see Reference bibliographies). For example, Pedro Domingos, *The Master Algorithm* provides an excellent categorization of the different AI schools, origins, algorithms and best solutions for various problem areas or tasks (Domingos, 2015). Carnegie Mellon's Tom Mitchell or Karoly Zsolnai's 'Two Minute Papers' Youtube Videos both provide excellent inspiring overviews of recent AI development and progress.[2]

AI Paradigm	Origin	Algorithm	Problem	Solution
Deep Learning Machine Learning	**Neuroscience (Neural Nets)**	**Back Propagation Neural Nets**	Complex Tasks, Hidden Patterns	Back propagation
Symbolic AI	Logic, Philosophy	Inverse Deduction	Knowledge Composition	Inverse Deduction
Bayesian Inference	Statistics, Probability Theory	Probabilistic Inference	Uncertainty	Probabilistic Inference
Evolutionary Computation	Evolutionary Biology (Complexity Theory	Genetic Algorithms	Structure Discovery	Genetic Programming
Reasoning by Analogy	Psychology	Kernel Machines (Support Vector Machines)	Similarity	Kernel Machines

Main AI Paradigms, Origins and Algorithms. Dr. Pedro Domingos, The Master Algorithm, 2015.

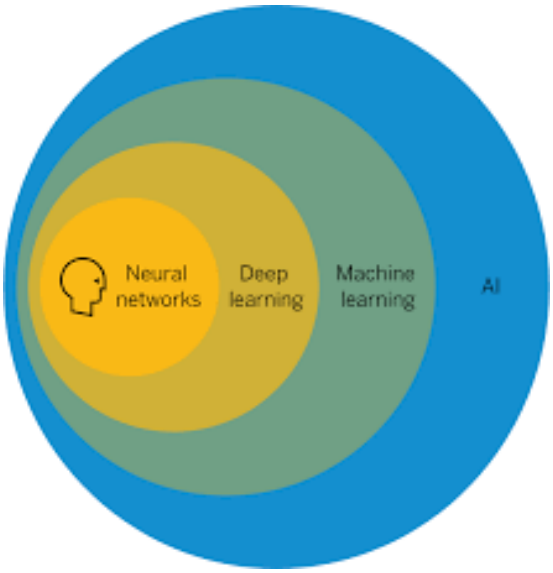
The larger idea is to begin building awareness in both a larger university community and, simultaneously, library staff so that there is a gradual building of awareness, inspiration and desire for further knowledge and to build skillsets. Present AI attention is largely focused on Deep Learning, Machine Learning and neural nets (See Carnes, 2019; Coldfusion, 2020; Lecun, 2022, Mitchell, 2022; Fridman, 2022 et al.). While there are other important areas, this is then an excellent place to focus on deeper beginnings to a wider program as there is both a lot of current attention here and many significant, inspiring gains.

[2] See Tom Mitchell: https://www.youtube.com/watch?v=ij9vqTb8Rjc and Karoly Zsolnai: https://www.youtube.com/c/K%C3%A1rolyZsolnai/videos

Levels of Learning

By focusing on a particular area after a group is inspired, there is some feeling of control and gaining of skills that may be achieved, enabling both library projects and faculty and graduate student research to new levels.

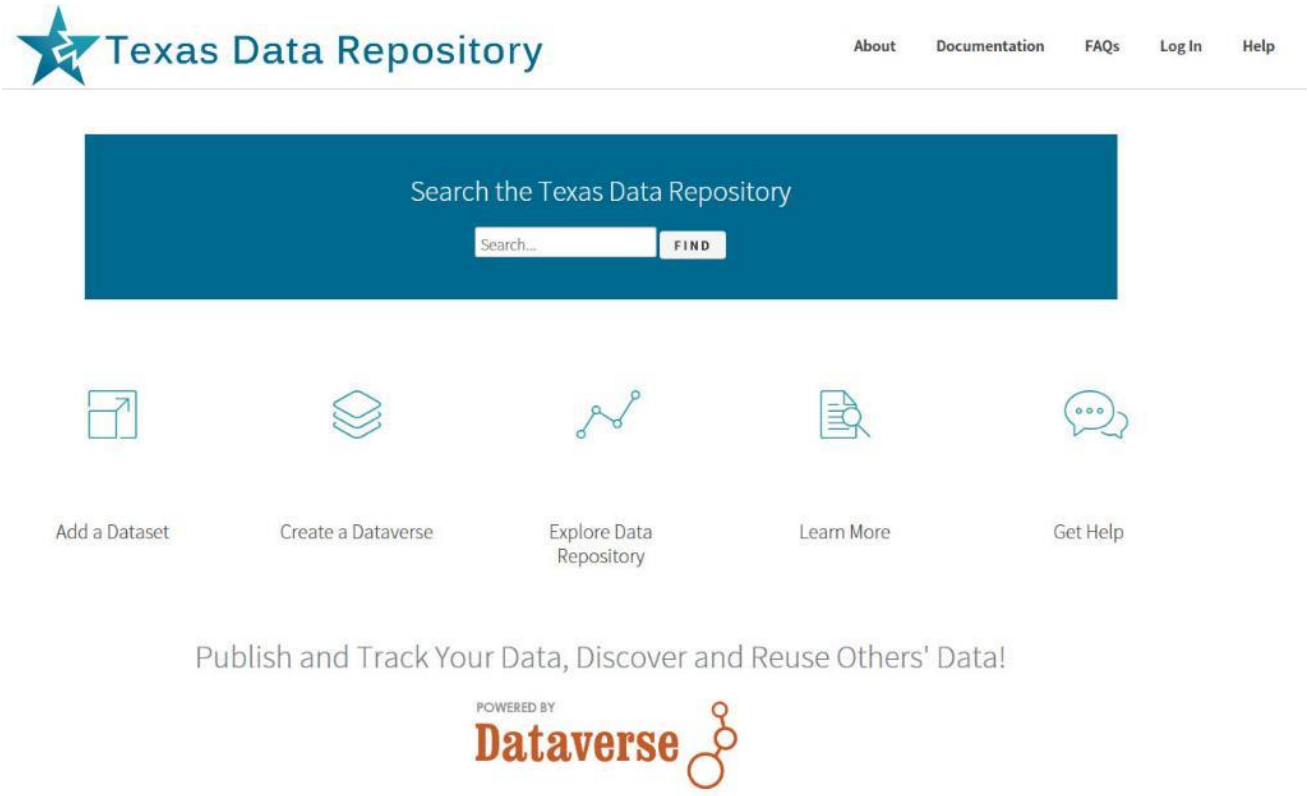
The last ten years of Deep Learning or Neural Net Algorithms have shown incredible progress with regards to results from natural language processing and conversational chatbots to cybersecurity to strategic reasoning (AlphaGo) to computer vision and object recognition (Mitchell, 2022). Here, it is best to both briefly overview the field but also, pragmatically set scalable limits so that progress may be made with both algorithmic paradigms and pragmatic application for both library staff and research faculty and graduate students so that projects may be achieved, and core research and data enabled.



Fields and Subfields of AI: Machine Learning, Deep Learning and Neural Networks

4 ONLINE DATA RESEARCH REPOSITORIES

With any Library AI program, it is also best to begin pragmatically. There is a clear trajectory in academic libraries from data and data collection to data science to data analytics, visualization and AI. This all begins with the data. Its organization and center is a good online data research repository. An academic online research data repository will allow both a university library to consolidate and share online faculty and graduate student research, manage university research data and provide important online data archiving, arching and publishing strategies for research data. It will also provide library staff and surrounding research faculty and graduate students important entry level skills needed towards AI. These foundational skills surround important tasks of data organization, data cleaning, creating structured data, data citation and creating metadata schemas, among other skills. These skills will all be important building blocks needed towards later AI's 'labeling' pathways.



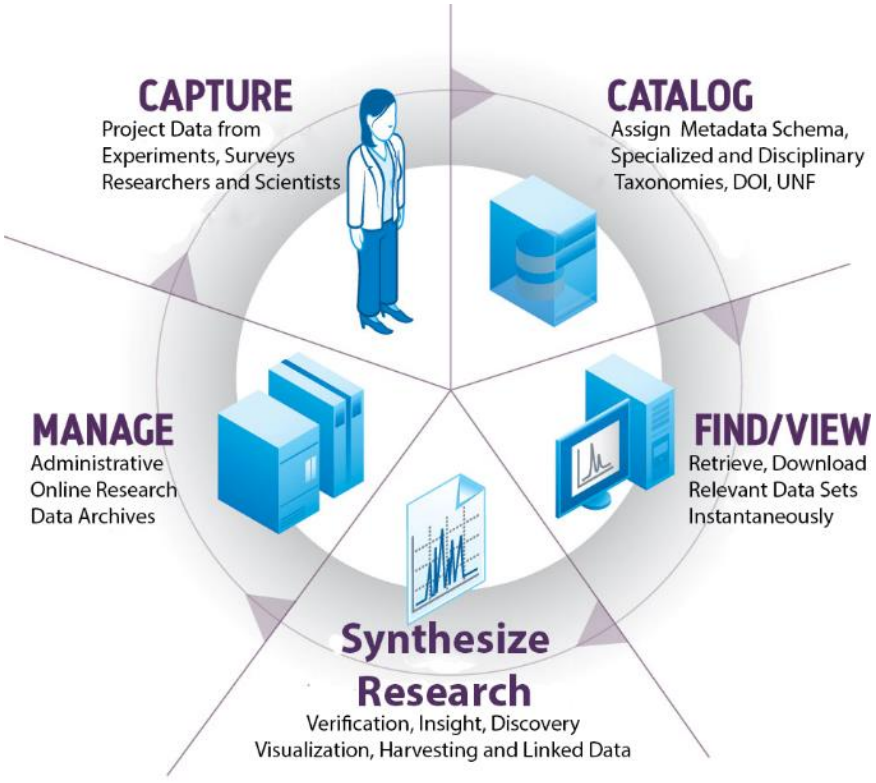
Texas Data Repository, http://data.tdl.org[3]

The Texas Data Repository is a good example of a preliminary foundation needed step to begin any AI program. The Texas Data Repository reconfigures Harvard's open source Dataverse as a consortial environmental aggregating research data from various Texas universities collaborating together. Setting up this type of open-source software on an individual institutional or consortial configuration will build many infrastructure skills. This will enable library staff in setting up this data-centered service for the larger university community, but also for the university researchers in beginning to build their 'data science' skills towards the later 'data' and algorithmic literacy needed for AI paradigms.

[3] Also See Uzwyshyn, Online Data Repositories (2016). https://www.researchgate.net/publication/304780954_Online_Research_Data_Repositories_the_What_When_Why_and_How

Open Refine https://openrefine.org/

Enabling a data repository for the institution will also encourage scholars and library staff to learn basic ‘data cleaning’ tools such as Open Refine. Open Refine is a powerful tool for working with messy data and transforming it so it will be in a suitable state to be taken from a repository and utilized by a later AI algorithm for later processing and training.



The Online Research Data Repository Lifecycle, Uzwyshyn, 2016.

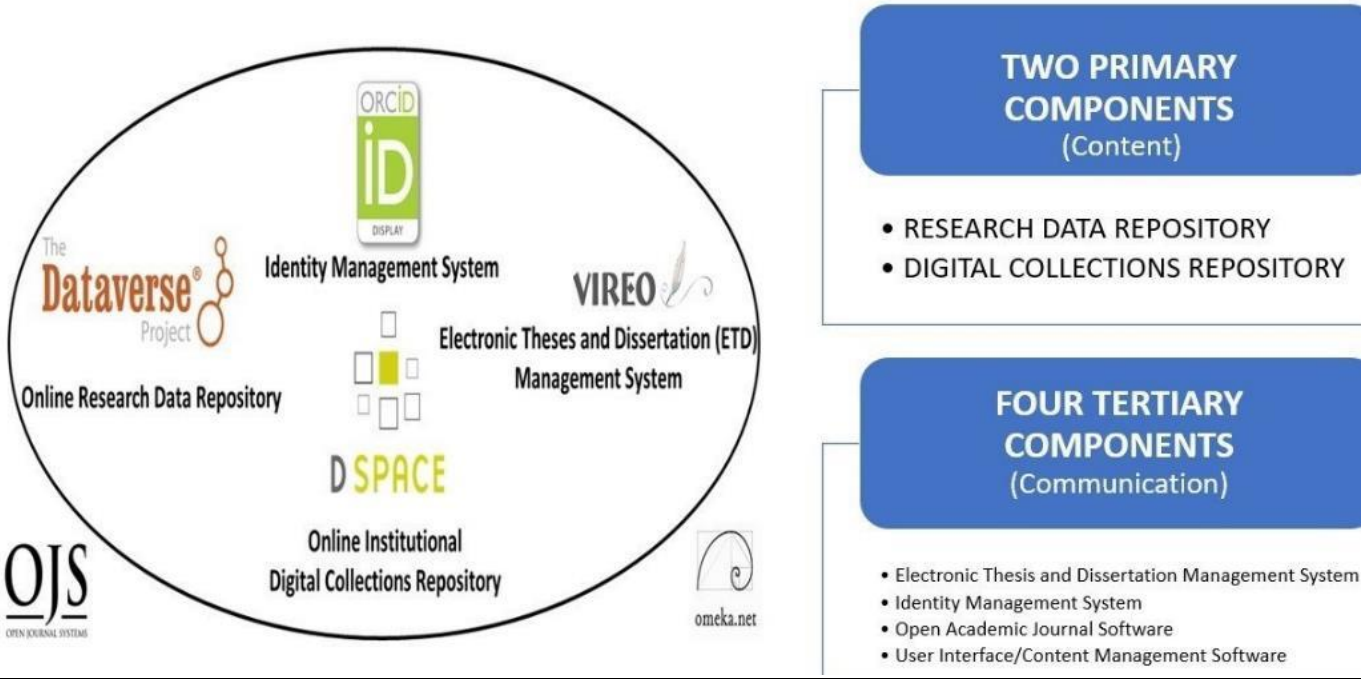
Gaining fundamental data literacy with the larger online research data repository lifecycle will serve both the university community and library staff well for excellent foundations on which to build the next levels of AI projects

To generalize, a university online data repository and larger research community competence with both the data research lifecycle and how the data repository enables this larger ecosystem will provide a great foundation for later AI pathways.

5 DIGITAL SCHOLARSHIP ECOSYSTEMS (DSE)

A digital scholarship ecosystem (DSE) should also be pursued in parallel or following quickly on the development to an online research data repository if a research library does not possess one already. While a Data Repository will always be central, an online institutional collections repository should also not be overlooked, especially for the ability to store and house both the metadata and core data for large textual files. These 'token'-based datasets are needed for natural language processing and text-based AI. A larger DSE generally consists of two primary online components for content and four tertiary online components for communication. A research data repository and digital collections repository will both make up together the primary content repositories (data, media and text).

As mentioned, Texas State University utilizes the consortial Texas Data Repository based on Harvard's Dataverse for the Data Research Repository. The well-known open-source software, DSpace is used for the university's digital collections repository. For universities and research institutions, the four tertiary components will enable better online global communication and networks. These are an online electronic theses and dissertation management system (ETD System, VIREO), identity management system (ORCID), open academic journal system software (OJS3) and user interface content management software (OMEKA). Together, these function as a unified digital scholarship ecosystem (DSE). This ecosystem allows great facility in later enabling larger AI pathways continuing to build on strong foundations.

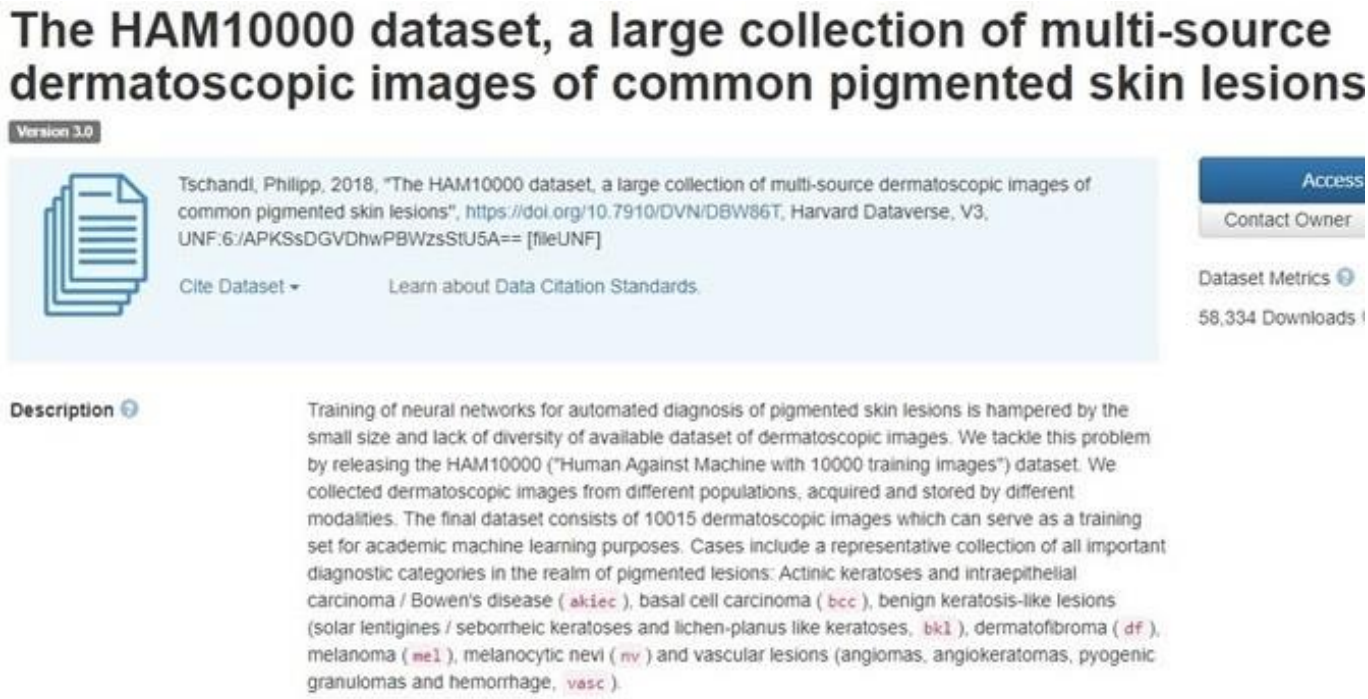


A Digital Scholarship Research Ecosystem, Six Components, Online Content and Communication[4]

[4] See Uzwyshyn, 2020. Available at: https://www.researchgate.net/publication/336923249_Developing_an_Open_Source_Digital_Scholarship_Ecosystem

6 OPEN SCIENCE, DATA, AI AND DIGITAL SCHOLARSHIP ECOSYSTEMS

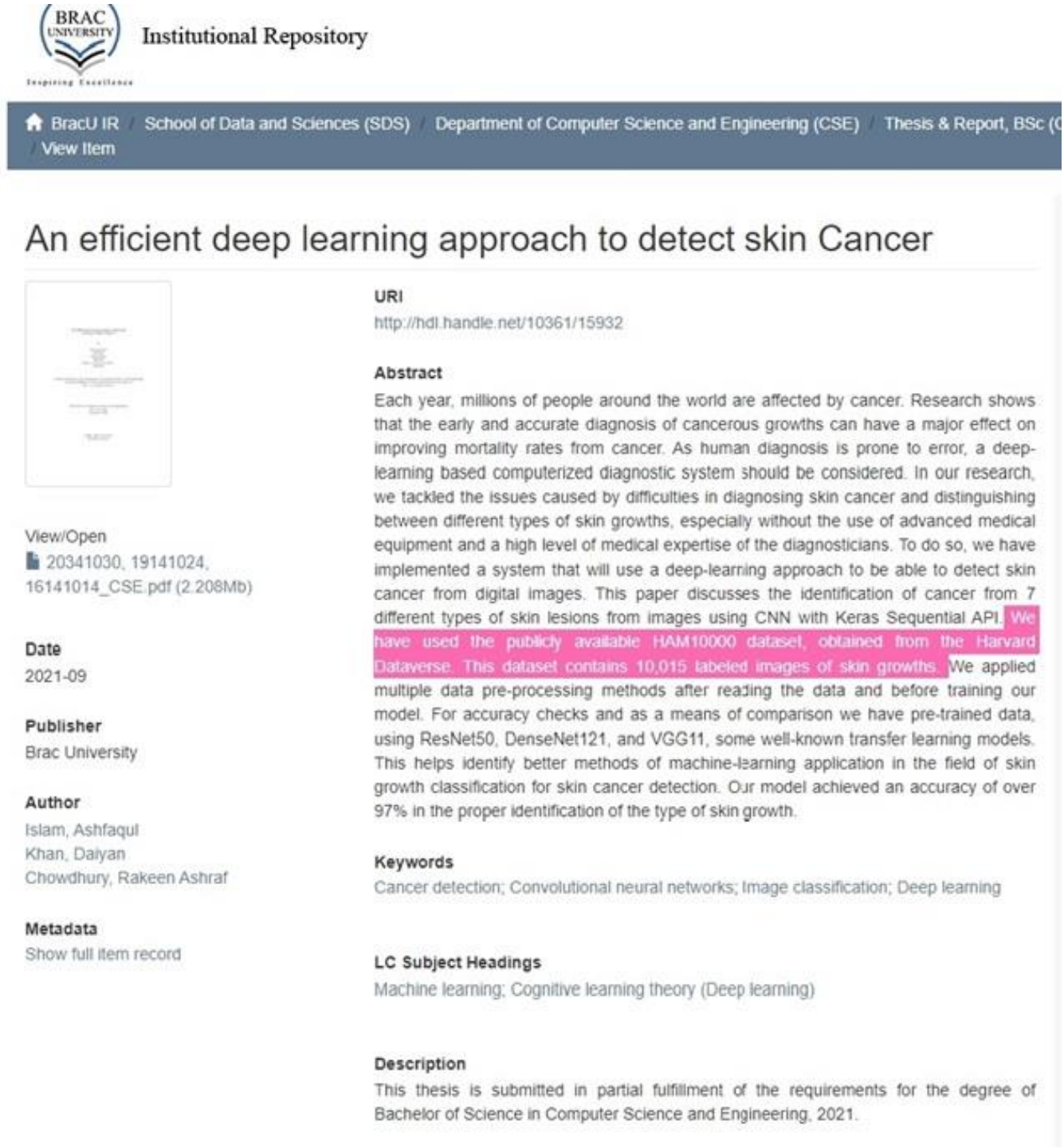
Innovative Open Science and AI possibilities are now enabled through affordances and combination of a digital scholarship ecosystem and data research repository. For example, the below HAM 10,000 image dataset is a large collection of multi-source dermatoscopic images of Cancerous skin lesions uploaded to Dataverse by Viennesse Dermatologist, Dr. Philip Tschandl, in 2018. Because Harvard's Dataverse allows for the uploading of datasets from other universities globally, appropriate research datasets may be uploaded for later sharing or use by anyone globally



HAM10000 Dataset, https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DBW86T

Below is a cover page from BRAC University from Dhaka Bangladesh that uses DSpace as an institutional repository to house theses and dissertations from the School of Data and Sciences, Dept. of Computer Science and Engineering. Here the students have downloaded and utilized the image data as training material to train a deep learning neural net algorithm to recognize cancer growths with efficiency greater than, or equal to, board certified dermatologists.

This is very good example of open science and AI possibilities operating on global levels through the enabling power of digital scholarship ecosystems and data repositories. Content and data that otherwise would be unavailable is brought together with new machine learning algorithmic techniques and new research and a very good thesis is produced. Globally dispersed content and knowledge from three different continents has been aggregated to advance the pursuit of knowledge and science. [5]



BRAC University Dspace Repository 2021 Deep Learning/AI Thesis http://dspace.bracu.ac.bd/xmlui/handle/10361/15932

[5] See also, the original inspirational discovery article for this thesis. Esteva, A, Thrun, S. et al. Dermatologist-level Classification of Skin Cancer with Deep Neural Networks. Nature, Volume 542 (February 2, 2017). pp. 115-119. doi:10.1038/nature21056

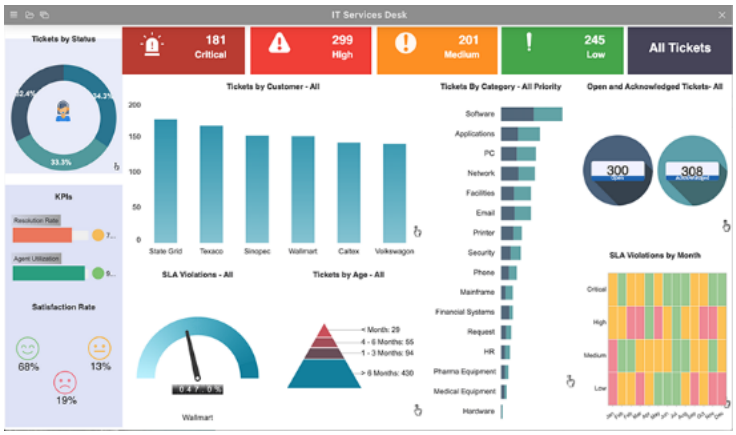
7 AI HUMAN RESOURCES

In creating an AI infrastructure for most libraries, hiring a whole new staff department will not be feasible. Many research and academic libraries, though, will have an online digital collections repository such as DSpace well in place and operational. This will also serve the library well. Many of the institutional repository content administrative skills gained with the Institutional Repository are transferrable to the Data Research Repository. This makes AI and data learning curves much easier to begin.

A staff member already in place for this repository position can initially take up a data-centered function with an upcoming data research repository as this begins to collect research faculty and graduate student research and a new 'data scientist' is hired. Other staff skills are similarly reconfigurable.

A cataloger, or cataloger focusing on metadata, can easily be reconfigured also part time to begin as the data repository metadata specialist. This will serve upcoming AI functions well, especially with regards to both machine learning. Much of the neural net training here is about what is called in computer science terms 'labelled' and 'unlabeled' data. Essentially, this is a simpler schema for metadata. Various other data cleaning and metadata skills will also come in useful here.

As a tiered gateway towards AI, hiring a Data Visualization specialist will also be very useful to gain support towards more complex AI initiatives. This specialist will initially provide library-wide support for data visualization and data analytics projects through dashboards and information visualizations to support data-driven decision making and finding insights from library data, but also from faculty and graduate student research data.



The Insights Possible from Data Driven Dashboards

Similarly, subject liaisons and research and information outreach librarians will come in very useful in communicating with both various departmental and school research faculty and conveying the libraries new data driven scholarly communication possibilities with information visualization and enabling research faculty with their possible data sets through both the repository and possible further information visualization help and resources. This will also serve to introduce both the new 'data repository', data/information/research visualization and bridge toward upcoming 'AI' possibilities and these pathways.

Library Data Driven Dashboards as Operational Gateways Towards AI

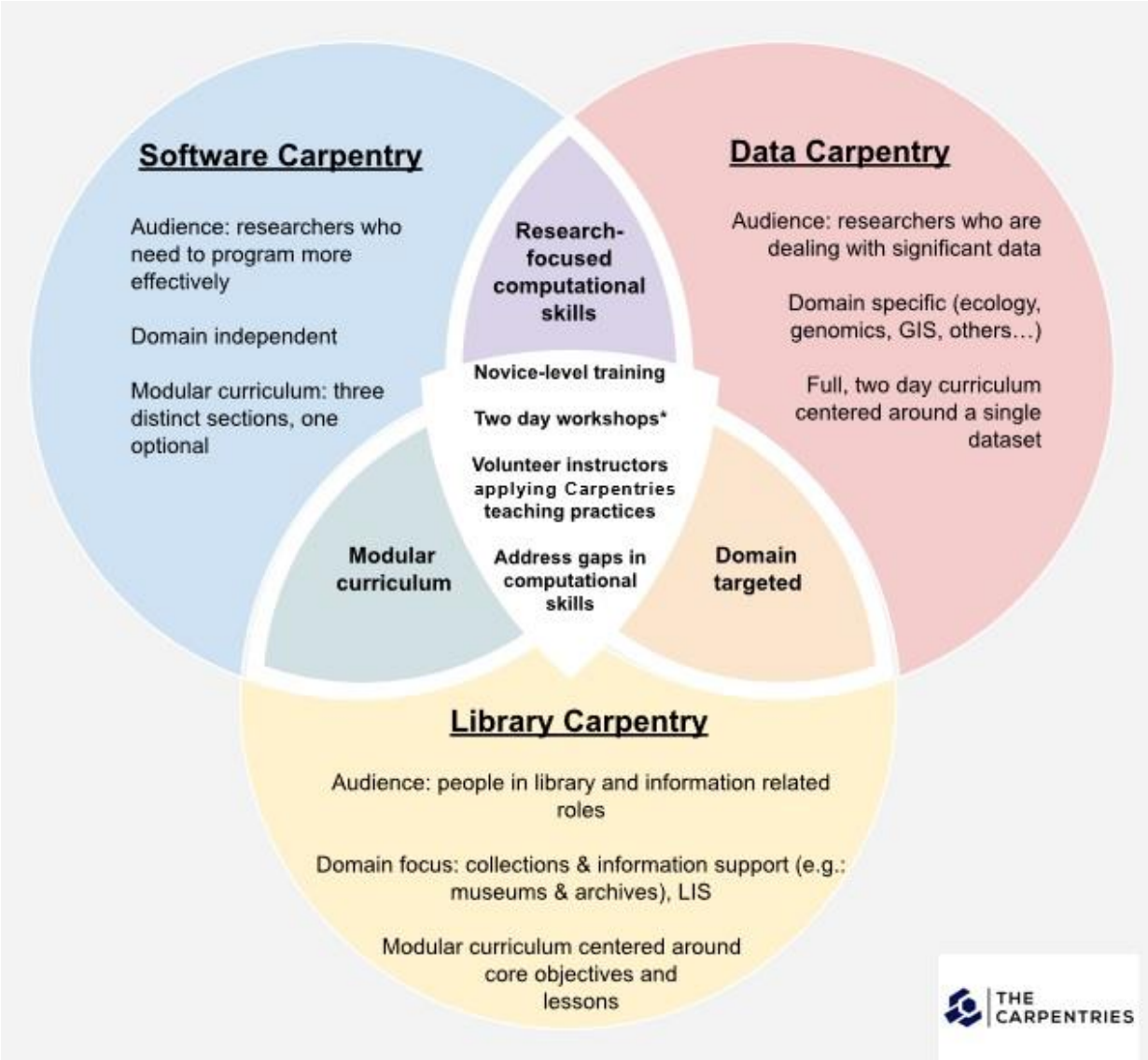
A Data Research and Information Visualization Specialist will also allow both higher university administration and research faculty and graduate students to see the usefulness of the programmatic possibilities with data towards greater insight. Education here on research faculty levels will begin to develop higher level strategies to clean and normalize data for future AI R&D projects and begin more complex programmatic analytics pathways with Python. The later Data Visualization Specialist or Librarian can then be the bridge for transferrable skills towards a more AI and full time Data repository role as needs increase.

8 AI LEARNING PATHS FROM DATA TO CARPENTRIES

Data Carpentries https://carpentries.org/

As the algorithmic literacy needs of both library staff and the surrounding university research community increase there will be needs arising from both researchers and library staff for more pragmatically oriented foundational coding and data science skills. With the

libraries' activities, research academics will begin to get a taste and understanding for enabling their research data towards insights on higher levels. Various areas of the library will also begin to realize the potential of these more algorithmic pathways. Here libraries worldwide can avail themselves of the Carpentries and host workshops in what is termed data carpentry, software carpentry or library carpentry.



Data, Software and Library Carpentries https://carpentries.org/

Carpentries workshops combine pragmatic programmatic knowledge needed for university researchers and graduate students with algorithmic literacy needs of library staff. They are also great scholarly communications bridges for dialogue and collaborative work between research faculty working on learning how to enable their research through data and programming and library staff who are also taking up these new methodologies towards larger library algorithmic literacy AI infrastructures and programs

9 LIBRARY AI CONFERENCES

As the learning curves and paths towards AI are steep, it will be important along these paths to keep staff both motivated and inspired - motivated with inspiring examples of benchmarks and milestones being achieved in our surrounding society (medicine, natural language processing, strategic games) but also inspired through new library AI conferences arising.

Stanford University Libraries Fantastic Futures Conference for Libraries Archives and Museums, https://library.stanford.edu/projects/fantastic-futures

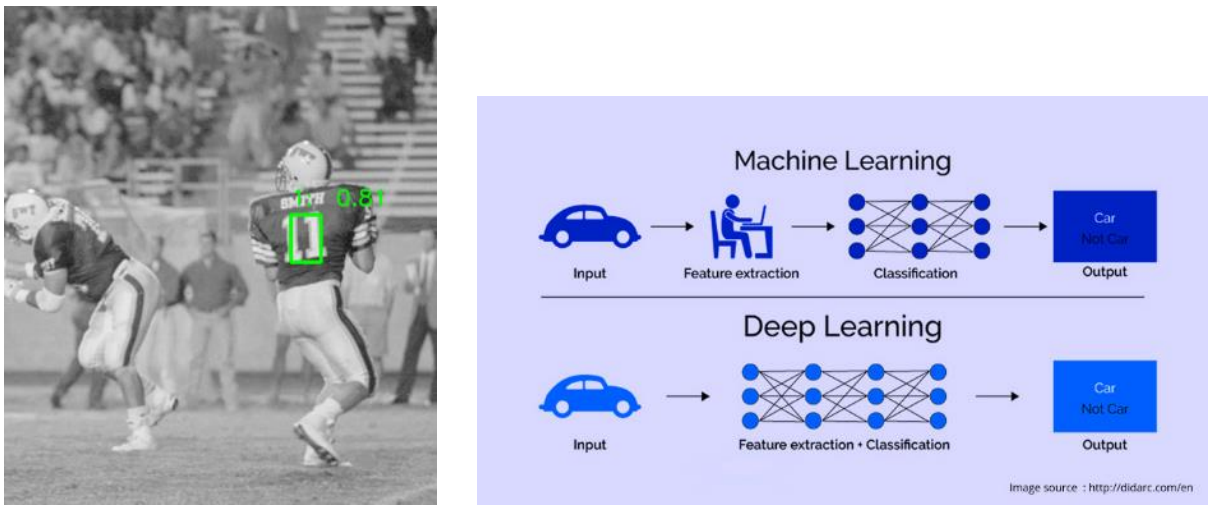
Two have become more well known: Carnegie Mellon's Artificial Intelligence for Data Discovery and Reuse and Stanford's 2nd International Conference on AI for Libraries, Archives and Museums. Stanford's Conference has travelled from the US to the Bibliothèque Nationale de France in Paris and then Norway. In subsequent years, Carnegie Mellon has fascinatingly combined with the CMU's Open Science Symposium.

Carnegie Mellon University Libraries AI for Data Discovery and Reuse, https://events.library.cmu.edu/aidr2020/

Both together can serve to inspire and motivate staff.

Artificial Intelligence presentations are also beginning to occur at more general library technology conference such as Computers in Libraries, Coalition of Networked Information Meetings, Texas Conference of Digital Libraries and IFLA's now historic first AI Satellite Conference New Horizons for Artificial Intelligence in Libraries, 2022.

10 FIRST LIBRARY AI R&D AND BETA PROTOTYPES



Texas State Convulational Neural Nets Football Jersey Number Feature Extraction Classification Prototype

It will be important at this time for motivated library staff to also be encouraged with pursuing AI beta projects. More than larger successes, these projects will be important for staff to gain initial understandings of models and the various pieces needed here for working on convolutional neural nets and deep learning model projects. Learning new areas of knowledge here are wide though and will range from AI processing power parameters needed (compute) and new video cards (NVIDIA GPU's) to the programming language of AI (Python) and vocabulary of pretrained and untrained models, classification, feature extraction, and image and natural language libraries, to name a few.

Texas State San Marcos Newspaper Neural Nets Object Recognition Metadata Classification Project, Peters 2022.

What is important in these first experiments is getting the staff accustomed to these larger models and what can and cannot be achieved. Understanding larger processes and possibilities is more important than large results at this stage and will begin to get library staff working on levels of R&D and iterating on results as classification R&D improves with model training.

11 LIBRARY AI WORKSHOPS, INSTITUTES AND FELLOWS PROGRAMS

Library specific AI Fellows institutes and workshops are also beginning to appear, and it is important to write recommendation letters and send motivated employees who apply for these at times funded workshops to attend. These institutes will be important both for motivating staff, but also for sharing curriculum and creating larger networks with other motivated parties.

INSTITUTE of Museum and Library SERVICES	IDEA INSTITUTE ON ARTIFICIAL INTELLIGENCE
• Weeklong Fellows Program at University of Texas Austin (20 Fellows) • Onboarding, • AI Institute • Library Centered AI • Programming Workshops • Final Project • AI Specialist Support • Networking with National Library AI Experts • Other Fellows	■ AI challenges and opportunities ■ Ethical considerations and guidelines ■ UX-Human/AI Interaction Lifecycle ■ Existing library, archive, and museum projects ■ AI project planning o Project Design o Data collection, classification, and transformation o Roles and implementation ■ Python Basics, Python for Machine Learning ■ APIs and bibliometrics ■ AI in search and discovery ■ Machine learning and coding ■ Harvesting, evaluating, and training data sets for use in AI ■ Conversational AI – Theoretical foundations ■ Conversational AI – applications ■ Linked open data Machine learning for text with topic modeling and clustering

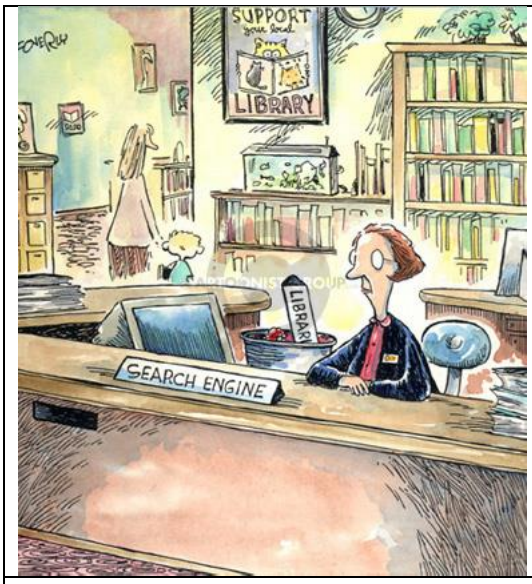
IMLS IDEA Institute on Artificial Intelligence. https://idea.infosci.utk.edu/

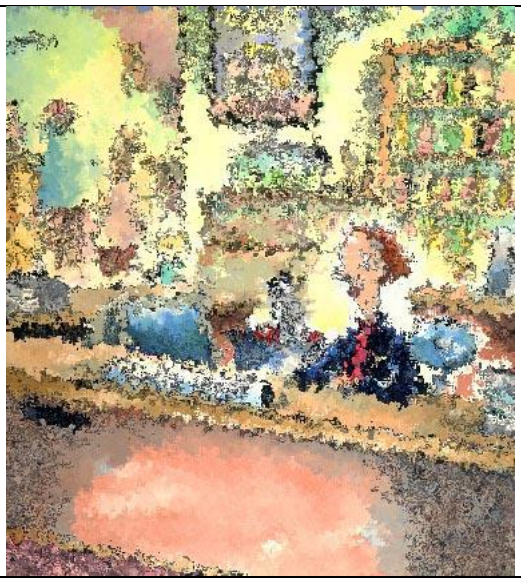
Following these types of institutes, it is important to open the door towards both projects that have begun through these opportunities but also gestate leadership opportunities for sharing curricula and beginning to think about further infrastructures that may be constructed locally to develop out of organic needs.

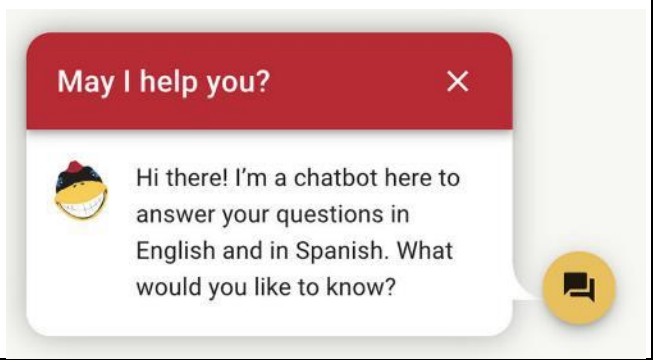
12 UNIVERSITY LIBRARY COLLABORATIONS

It is important to embrace change and it is paramount to understand that as the larger library begins to retool for the paradigm shift of AI so, too, will the larger parent research institution and associated IT infrastructures. As this occurs, unexpected opportunities for collaboration

and participation and partnership will begin to occur. For example, there is a current trend of dissolving and downsizing the traditional libraries research and information outreach services. The traditional reference and subject librarian services are transforming as both a cost saving measure, but also a less used, in person, help system to online and other modalities. Simultaneously, many universities are now adopting new AI Chatbot infrastructure for the students and faculty campuswide. This new chatbot infrastructure presents opportunities for the libraries to retool and utilize their skills in different ways. This allow the subject librarians to train the chatbot, retool and transform previous research skills and expertise towards new paradigm possibilities.

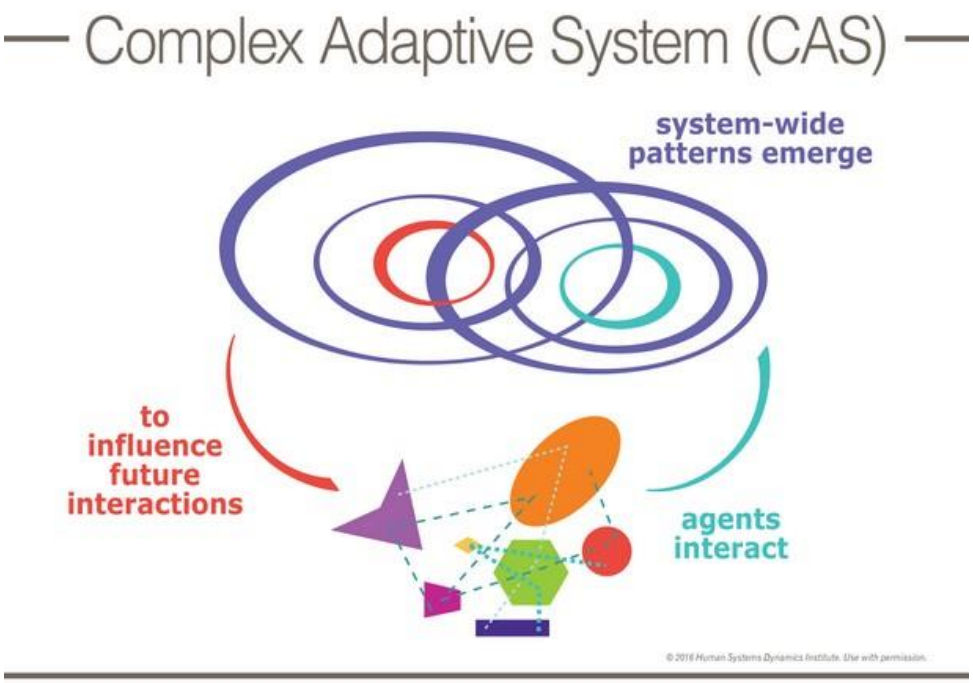






These unexpected AI changes should also be embraced as they also open doorways towards exploration toward research. Previous research skills can be now retooled towards the investigation and fine tuning of future AI natural language processing models ranging from Open AI's GPT3 and GPT4 to Google's DeepMind Gopher and other upcoming models. Help is not disappearing in the 21st century but the chatbot may not need some greater help in helping its human audiences. This help may now be accompanied by a background guide at the side, the previous research and instruction librarian, now Chatbot Administrator.

13 LIBRARIES AS COMPLEX DYNAMIC ADAPTIVE SYSTEMS



Libraries as Complex Adaptive Systems, Bryant, Dortmund, Lavoie, 2020.[6]

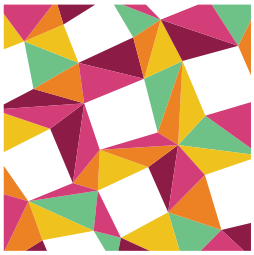
Libraries are complex interactive, adaptive, dynamic systems. As library staff, university faculty and graduate students interact from the bottom, system-wide patterns begin to emerge from the top. As this emergence begins to occur, it is important to begin to formalize some of this activity through a more formalized AI Working Group (AIWG). Ideally, the membership of this group should begin through informal partnerships, friendships and collaborations of interested library staff. As this activity ramps up, it should begin to be a little more formalized to continue conversations. This will provide direction, responsibility and accountability for artificial intelligence projects arising and important ethics related discussion, decisions, policy and later oversite.

Once this group has also become a little more formalized, it can be also opened further so both university research faculty and graduate students are invited to also help guide and offer suggestion and direction to possible paths, but also add to the innovation and possibility.

[6] See: https://www.oclc.org/content/dam/research/publications/2020/oclcresearch-social-interoperability-research-support-a4.pdf

14 NEXT STEPS, GRADUATE STUDENTS, POST-DOCS AND PERMANENT AI HIRES

At a certain point, library staff will realize they have come to the end of their AI learning path. This will signal the time to hire a first graduate student and eventually, a Ph.D. or Post Doc in Machine Learning or AI. Good graduate students can be found more easily in university engineering schools or computer science departments. Course listings will have names like *Machine Learning for Engineering Applications* or *Neural Nets and Deep Learning for Computer Science*. Associated research faculty will normally be only too glad to help assign suitable students and work with library staff to have their students nominally paid for part time AI research assistant projects or theses work. Relationships are key here and these will be win/wins for both the library, student, professor, and university in many ways.



CLIR Postdoctoral Fellowship Program https://postdoc.clir.org/ CLIR AI Research in Archives https://haira.clir.org/blog/

Following these successes, graduate students can be hired permanently. Ph.Ds. and Post-Docs can also be sought. There are even library specific Post-Doc programs through the Center for Library and Information Services on both US and CLIR global levels. These may be pursued to make great use of the recent Ph.D.'s specialized skills and bring new AI skillsets into libraries.

15 CONCLUSIONS – THE ROAD TO LIBRARY AI SUCCESSES

The new road to library AI successes is still largely untrammeled and open. There are many opportunities for most internal sections of the library. This ranges from Special Collections and Archives to possibilities for better understanding the black box of acquisitions' budgets to metadata possibilities with AI and natural language processing. Search and retrieval, library usage data, statistics and deriving insights from vast arrays of data now make up the 21st century academic research library landscape.

Beyond this, there are the incredible possibilities of connecting research faculty and graduate students and their research data collections with AI. Most of these researchers come from traditional academic disciplines. These are still widely outside of artificial intelligence, machine learning and computer science and these possibilities are completely unknown to them. The opportunities for these new infrastructures are incredible. This is both towards the advancement of the next levels of human knowledge and exploration for discovery and insight previously thought impossible. On these levels, it is important for libraries to begin thinking along these paths to enable new library artificial Intelligence infrastructures.

REFERENCES

(Select AI, Data Science and Learning Resources Bibliography)

Articles, Reports and Books

Bryant, R. Dortmund, A. and Lavoie, B. Social Interoperability in Research Support: Cross Campus Partnerships and the University Research Enterprise. *OCLC Research Reports.* 2020. https://www.oclc.org/content/dam/research/publications/2020/oclcresearch-social-interoperability-research-support-a4.pdf

Carnes, Beau. *Grokking Deep Learning in Motion*. Manning Publications (O'Reilly), 2019.

Cordell, R. *Machine Learning and Libraries: A Report on the State of the Field*. LC Labs: Library of Congress, Washington D.C. 2020.

Domingos, Pedro. *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. New York: Basic Books, 2015.

Esteva, A, Thrun, S. et al. Dermatologist-level Classification of Skin Cancer with Deep Neural Networks. Nature, Volume 542 (February 2, 2017). pp. 115-119. doi:10.1038/nature21056

Hervieux, S. and Wheatley, A. *The Rise of AI: Implications and Applications of Artificial Intelligence in Academic Libraries*. ACRL 2022.

Johnson, Steven. *Emergence: The Connected Lives of Ants, Brains, Cities and Software*. New York: Scribner, 2001.

Kleinveldt, Lynn. Smarter high education learning environments through AI: What this means for academic libraries. *Trends and Issues in Library Technology: Special Issue on AI*: June 2022. pp. 12-15. https://repository.ifla.org/handle/123456789/1940

Recht, Benjamin and Hart, Moritz. 2022. *Patterns, Predictions and Action: Foundations of Machine Learning* (Early History). Princeton University Press. 2022.

Sanchez, Nogales, Patrana Garcia, Alicia and Cerdan Medina, Jose Carlos. Digital Transformation, Data Reuse and Heritage Collections at the National Library of Spain. *Trends and Issues in Library Technology: Special Issue on AI*, June 2022. pp. 16-21. https://repository.ifla.org/handle/123456789/1940

Sartorius. *Understanding the Relationship Between Data Science, Artificial Intelligence and Machine Learning*. 2020 https://www.sartorius.com/en/knowledge/science-snippets/data-science-vs-artificial-intelligence-vs-machine-learning-602514

Uzwyshyn, R. 2020 Developing an Open-Source Digital Scholarship Ecosystem. *International Conference on Education and Information Technology*. St. Anne's Oxford, United Kingdom. February 2020. Available from: https://www.researchgate.net/publication/357286044_Research_Data_Repositories_and_Global_Scholarly_Ecosystem_Possibilities.

Uzwyshyn, R. 2016. Online Research Data Repositories: The What, When Why and How. *Computers in Libraries*. 36:3, April 2016. pp. 18-21. http://rayuzwyshyn.net/TXU2016/OnlineDataResearchRepositoriesUzwyshyn.pdf

Warren, Erica. Maximize Learning: Keeping Students in the Zone of proximal Development. *Good Sensory Learning* 2021. https://goodsensorylearning.com/blogs/news/scaffolding-development

General AI Video Overviews/Introductions

In the Age of AI. Frontline PBS. 2021. https://www.youtube.com/watch?v=5dZ_lvDgevk

Artificial Intelligence and Algorithms: Pros and Cons. Tilman Wolff und Ranga Yogeshwar. DW 2021. Available: https://www.youtube.com/watch?v=s0dMTAQM4cw Part II, https://www.youtube.com/watch?v=-ePZ7OdY-Dw

Artificial Intelligence. Machine Learning. Neural Networks. Future Technology. Bloomberg Businessweek Canada. 2022. https://www.youtube.com/watch?v=ypVHymY715M

ColdFusion (2018). *Why Deep Learning Now?* (Documentary Overview). https://www.youtube.com/watch?v=b3IyDNB_ciI

Domingos, P. *The Master Algorithm: Talks at Google*. Nov 27, 2015, Retrieved from: https://www.youtube.com/watch?v=B8J4uefCQMc

Domingo, P. *The Ultimate Software*. PSW Science Talk. September 13, 2021. https://www.youtube.com/watch?v=7K9X2WiBvu

Domingos, P. *Managing in the Age of AI*. Miami: Transform, 2022. https://vimeo.com/694708852/9bbb077e93

Hofstadter, Douglas. *Analogy as the Core of Cognition*. Stanford Presidential Talk. September 10, 2009. https://www.youtube.com/watch?v=n8m7lFQ3njk

Jordan, Michael I. *Machine Learning, Recommender Systems and the Future of AI*. February 20, 2020. https://www.youtube.com/watch?v=EYIKy_FM9x0

Larochelle, Hugo. *The Deep End of Deep Learning*. TedxBox, 2017. https://www.youtube.com/watch?v=dz_jeuWx3j0

Lecun, Y. *Heroes of Deep Learning: Chief AI Scientist at Facebook*, 2022. https://www.deeplearning.ai/hodl-yann-lecun/

Mitchell, Tom. 2022 *Where on Earth is AI Headed?* Carnegie Mellon. https://www.youtube.com/watch?v=ij9vqTb8Rjc

LeCun, Yann. *Is AI Just Statistics*. January 22nd, 2022. https://www.youtube.com/watch?v=iug8FAl0B1U

Peters, Todd and Long, Jason. *An Attempt at Metadata Enhancement through Machine Learning*. Texas Conference on Digital Libraries (Zoom Video and PPT). 2022.

Perl, Judea. (2020). *Causal Reasoning, Counterfactuals and the Path to Artificial General Intelligence*. (Interviewed by Lex Fridman)
https://www.youtube.com/watch?v=pEBI0vF45ic

Zsolnai, Karoly. *Two Minute Papers*. Recent AI R&D Papers.
https://www.youtube.com/c/K%C3%A1rolyZsolnai/videos 2022.

Python Online Courses

Lateef, Zulaikha. *A Comprehensive Guide to Artificial Intelligence with Python*. March 29, 2022. Retrieved from: https://www.edureka.co/blog/artificial-intelligence-with-python

Severance Charles (Professor at U Michigan School of Information). *Introduction to Python for Everybody*. Codecamp (Free). https://www.freecodecamp.org/learn/scientific-computing-with-python/python-for-everybody/

Machine Learning Courses

Machine Learning Crash Course (with TensorFlow Api's. Google. (Free).
https://developers.google.com/machine-learning/crash-course

Carin, Lawrence. *Introduction to Machine Learning*. Coursera (Duke University). Free.
https://www.coursera.org/learn/machine-learning-duke

Fridman, Lev. *MIT Deep Learning and Artificial Intelligence Lectures*.
https://deeplearning.mit.edu/ 2022.

Gugger, Sylvain and Jeremy Howard. *Practical Deep Learning for Coders with Fastai and Pytorch: AI Applications without a PhD*. https://course.fast.ai/

Ng, Andrew. *Deep Learning*. Coursera. https://www.coursera.org/specializations/deep-learning

Zhang, A, Lipton, Z., Li Mu, Smola Al. *Dive into Deep Learning*. https://d2l.ai/index.html

Hinton, Geoff. 2016. *Neural Networks for Machine Learning*. (Coursera Course).
https://www.youtube.com/watch?v=cbeTc-Urqak&list=PLoRl3Ht4JOcdU872GhiYWf6jwrk_SNhz9

Anna Dabrowski Digital Carpentry Materials. *GitHub - ajdabrowski/carpentries-instruction-materials: Materials for teaching Carpentries workshops*

AI Library Related Conferences and Organizations.

Artificial Intelligence for Data Discovery and Reuse (2019-2022). Carnegie Mellon. https://events.library.cmu.edu/aidr2019/

The Carpentries. Foundational Coding and Data Science Skills for Researchers and Libraries Worldwide https://carpentries.org/ 2022.

CLIR Global. *Council on Library and Information Resources*. https://www.clir.org/global/ Postdoctoral Library AI Program: https://postdoc.clir.org/

Fantastic Futures (Stanford Based). Norway (2022). https://www.nb.no/artikler/fantastic-futures/. FY21 (France). USA Stanford https://library.stanford.edu/projects/fantastic-futures

Leonard, Peter (Digital Humanities Library, Yale University). *Neural Networks: Machine Vision for the Visual Archive*. Coalition for Networked Information (CNI). 2018. https://www.cni.org/wp-content/uploads/2018/04/cni_neurel_leonard.pdf