

Steps Towards Building Library AI Infrastructures and Programs

(Research Data Repositories, Scholarly Research Ecosystems and AI Scaffolding)

Presented for New Horizons in AI for Libraries
IFLA Satellite Conference, Galway, Ireland
National University of Ireland, July 21, 2022

Ray Uzwyshyn, Ph.D. MBA MLIS
Director, Collections and Digital Services
Texas State University Libraries, USA
July 2022, ruzwyshyn@txstate.edu



Laddered Processes Towards Building Library AI Awareness & Competencies

Awareness, Building Skills, Knowledge, Mastery



Multi-Year Process 2014-2022

Texas State University Libraries



Large Academic Library system,
ARL Library
Main campus Library and other
offsite libraries (Health
Professions, Austin/Roundrock)



Texas State University,
Undergraduate, Graduate
and Doctoral Institution
40,000 Students

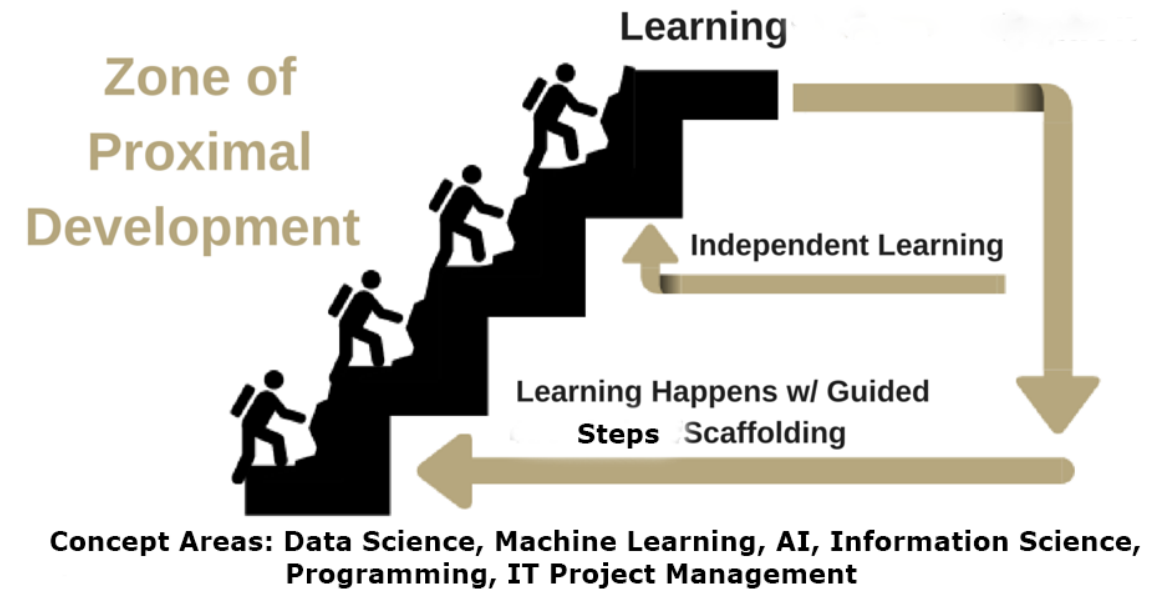


Texas State Libraries,
Academic Research Library
ARL Library

Steps and Scaffolding Towards Building Library AI Infrastructures



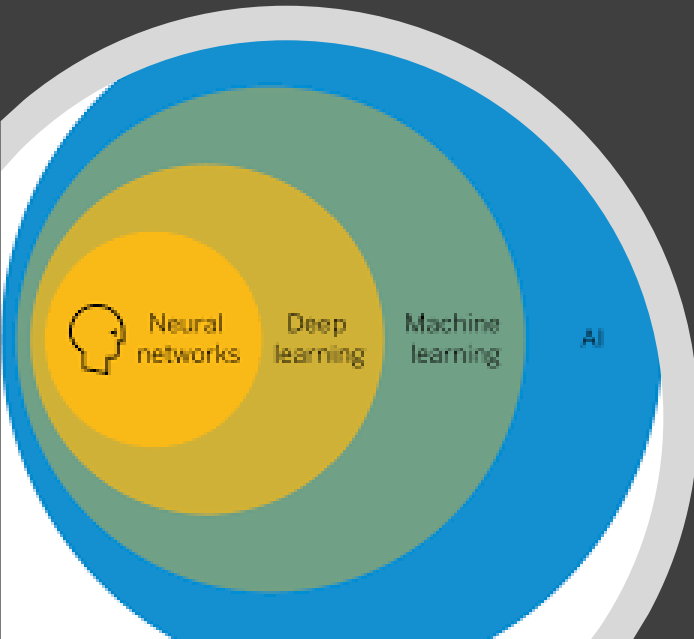
Learning is Too Hard: Anxiety



Learning is Too Easy: Boredom

Last Ten Years Has Shown Incredible Progress of AI

AI (Machine Learning (Deep Learning)) =
Better Algorithms + Greater Computing Power + Large Data Sets

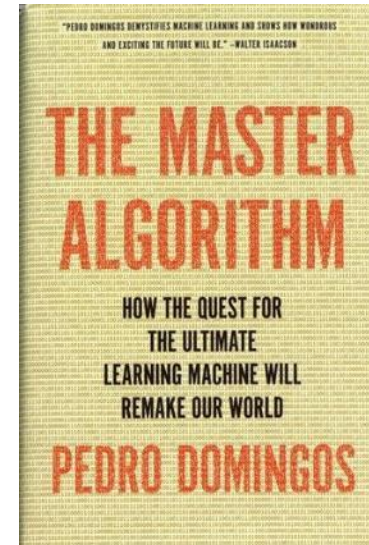


- Natural Language Processing (Speech to Text, Translation)
- Fraud Detection & Cybersecurity
- Conversational Chatbots & Robotic Agents
- Strategic Reasoning (AlphaGo)
- Computer Vision (Facial + Object Recognition Cancer Cell Detection)

AI Has Many Paradigms and Origins

Algorithms, Suitable Problem and Solution Methods,
Dr. Pedro Domingos, University of Washington

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



Recommendation #1: Begin with an Academic Data Research Repository



[About](#) [Documentation](#) [FAQs](#) [Log In](#) [Help](#)

Search the Texas Data Repository

FIND



Add a Dataset



Create a Dataverse



Explore Data
Repository



Learn More



Get Help

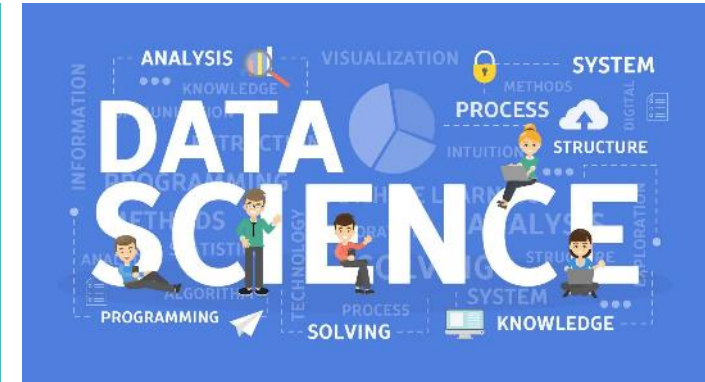
Publish and Track Your Data, Discover and Reuse Others' Data!

POWERED BY

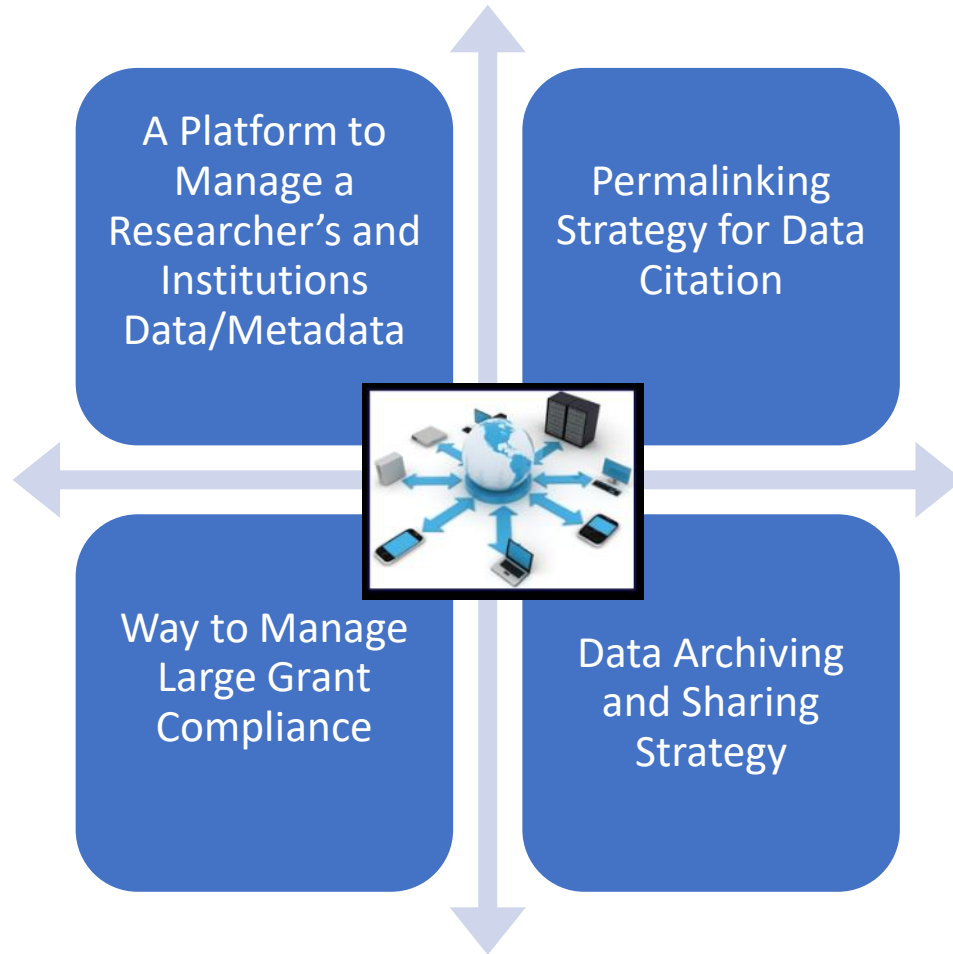


<http://data.tdl.org>

Clear Trajectory
in Libraries from
Data Collection
To Data Science ->
Data Repositories ->
Data Analytics ->
Data Visualization >
AI



What is an Online Data Research Repository?

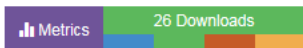


TEXAS RESEARCH DATA REPOSITORY



Texas Digital Library Test Dataverse

A statewide collaboration of higher education institutions in Texas



Share, publish, and archive your data. Find and cite data across all research fields.

Welcome to the Texas Digital Library Test Dataverse!

IMPORTANT: This Dataverse server does NOT include the [TwoRavens add-on](#).

Because of this, you may receive errors when ingesting certain datasets and the "explore" button will not work.

A screenshot of the Texas Digital Library Test Dataverse website. The interface features a header with the logo and a description. Below the header is a navigation bar with four tabs: Trinity University Dataverse, UT Medical Branch Dataverse, University of Texas Dataverse, and Texas State University Dataverse. The Texas State University Dataverse tab is highlighted with a red oval. Below the tabs is a search bar with a 'Find' button and an 'Advanced Search' link. A red oval highlights the search bar area. To the right of the search bar is an 'Add Data' button.



Texas Digital Library Consortium of 22 universities across Texas leveraging technological cooperation among academic libraries

Data Repositories Allow Building Skills For AI

Data Organization, Data Cleaning, Structured Data Citation, Sensitive Data and Metadata Schemas



OpenRefine is a powerful tool for working with data: (cleaning it)

Harvard Dataverse Network

Search icon, Information icon, Comment icon, Create Account link

REPLICATION DATA FOR: A MULTIVARIATE MODEL OF STRATEGIC ASSET ALLOCATION

hdl:1902.1/QBXRSFLBQJUNF:3:ZnYhHkZe2veTJAWaBDpPKA==

Version: 2 – Released: Thu Oct 03 16:46:32 EDT 2013

CATALOGING INFORMATION

Data & Analysis

Comments (0)

Versions

i If you use these data, please add the following citation to your scholarly references. [Why cite?](#)

John Y. Campbell; Yeung L. Chan; and Luis Viceira, 2007, "Replication data for: A Multivariate Model of Strategic Asset Allocation", <http://hdl.handle.net/1902.1/QBXRSFLBQJUNF:3:ZnYhHkZe2veTJAWaBDpPKA==> The Harvard Dataverse Network [Distributor] V2 [Version]

Citation Format

i Results found in this publication can be replicated using these data.

Original Publication

Campbell, John Y.; Chan, Yeung Lewis; and Viceira, Luis M., 2003, "A multivariate model of strategic asset allocation," Journal of Financial Economics, Elsevier, vol. 67(1), pages 41-80: [article available here](#)

Publications

John Y. Campbell & Yeung Lewis Chan & Luis M. Viceira, 2001. "A Multivariate Model of Strategic Asset Allocation," NBER Working Paper, National Bureau of Economic Research, Inc. [article available here](#)

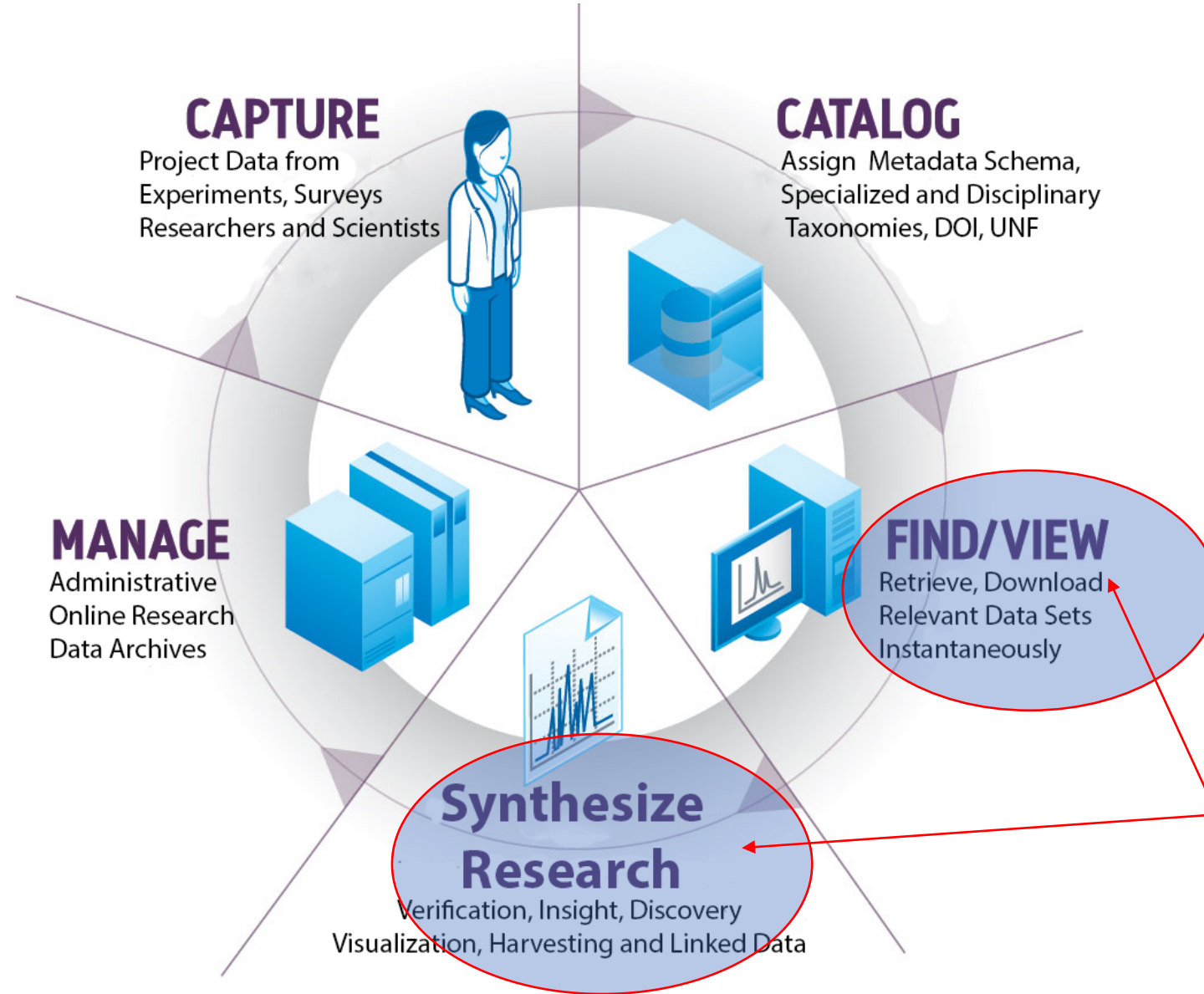
Campbell, John Y & Chan, Yeung Lewis & Viceira, Luis M, 2001. "A Multivariate Model of Strategic Asset Allocation," CEPR Discussion Paper 3070, C.E.P.R. Discussion Papers. [article available here](#)

Data Citation Details

Title	Replication data for: A Multivariate Model of Strategic Asset Allocation
Study Global ID	hdl:1902.1/QBXRSFLBQJ
Authors	John Y. Campbell (Harvard University); Yeung L. Chan; and Luis Viceira
Producer	John Y. Campbell
Production Date	2003
Funding Agency	National Science Foundation; Hong Kong RGC Competitive Earmarked Research Grant (HKUST 6965/01H); Division of Research of the Business School

The Research Data Repository Lifecycle

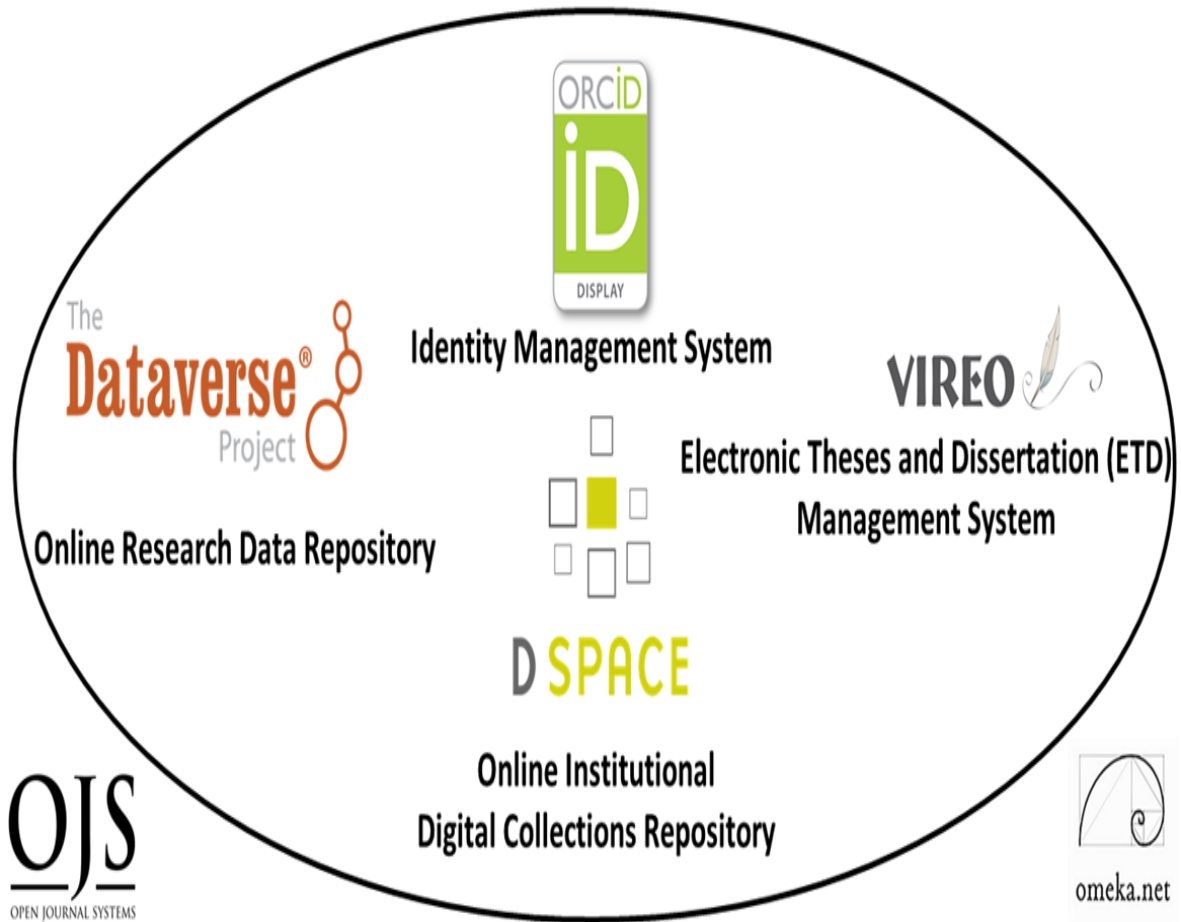
Setting Better Foundations & Organization for AI Infrastructures



Data Repository
provides
Basic AI, Machine
Learning, Open Science
and Research Needs.

Recommendation #2: Digital Scholarship Ecosystems, Foundations for AI

Six Open Source Software Components



TWO PRIMARY COMPONENTS (Content)

- RESEARCH DATA REPOSITORY
- DIGITAL COLLECTIONS REPOSITORY

FOUR TERTIARY COMPONENTS (Communication)

- Electronic Thesis and Dissertation Management System
- Identity Management System
- Open Academic Journal Software
- User Interface/Content Management Software

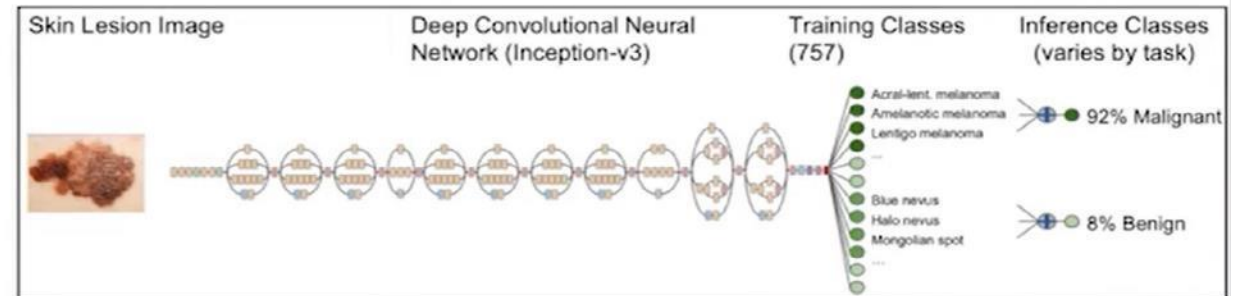
Dermatologist-level Classification of Skin Cancer with Deep Neural Networks,

Nature 2017, Andre Esteva, Brett Kupress, Sebastian Thrun et al.

AI Models, Deep Learning, Convolutional Neural Nets, Labeled Medical Data from Image Data Archives

Skin Cancer Diagnosis:

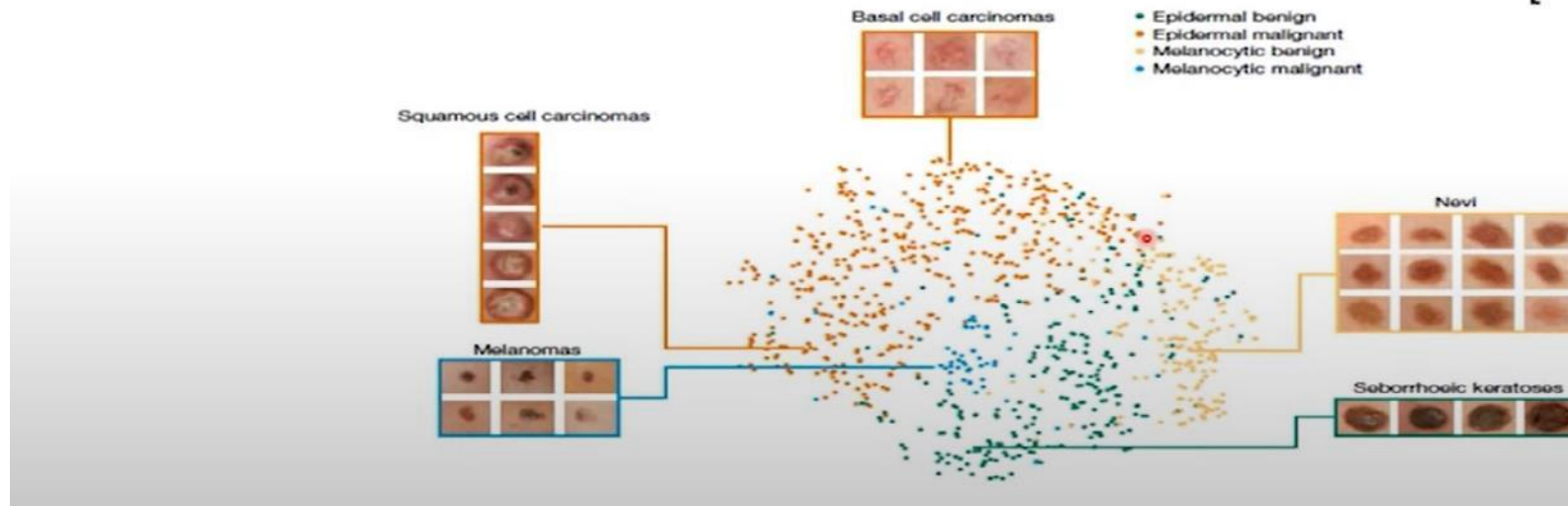
Trained on 1.4 M standard photographs
Retrained on 129,450 skin images
Deep net Inception v3 architecture
Outperforms doctors



[Esteva et al., *Nature* 2017]

[Video](#)

[Stanford
Overview](#)




Open Science, Data Research Repositories, Discovery and AI

Dataverse Data Research Repository

Dermatology Image Dataset,
Dr. Philip Tschandl, Vienneesse
Dermatologist

- Great Example of Open Science
- <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DBW86T>

 HARVARD
Dataverse


Add Data ▾ Search ▾ About User Guide Support Sign Up Log In

ViDIR Dataverse
(Medical University of Vienna)

Harvard Dataverse > ViDIR Dataverse >

The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions

Version 3.0



Tschandl, Philipp, 2018, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions", <https://doi.org/10.7910/DVN/DBW86T>, Harvard Dataverse, V3, UNF:6:APKSsDGVdhwPBWzsStU5A== [fileUNF]

Cite Dataset ▾ Learn about Data Citation Standards.

Access Dataset ▾

Contact Owner Share

Dataset Metrics ⓘ
58,334 Downloads ⓘ

Description ⓘ

Training of neural networks for automated diagnosis of pigmented skin lesions is hampered by the small size and lack of diversity of available dataset of dermatoscopic images. We tackle this problem by releasing the HAM10000 ("Human Against Machine with 10000 training images") dataset. We collected dermatoscopic images from different populations, acquired and stored by different modalities. The final dataset consists of 10015 dermatoscopic images which can serve as a training set for academic machine learning purposes. Cases include a representative collection of all important diagnostic categories in the realm of pigmented lesions: Actinic keratoses and intraepithelial carcinoma / Bowen's disease (**akiec**), basal cell carcinoma (**bcc**), benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus like keratoses, **bk1**), dermatofibroma (**df**), melanoma (**mel**), melanocytic nevi (**nv**) and vascular lesions (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage, **vasc**).

Files Metadata Terms Versions

Search this dataset...



Filter by

File Type: All ▾

Access: All ▾

☐ 1 to 6 of 6 Files



[HAM10000_images_part_1.zip](#)

ZIP Archive - 1.3 GB

Published Jun 4, 2018

15,433 Downloads

MD5: 463...e46



[HAM10000_images_part_2.zip](#)

ZIP Archive - 1.3 GB

Published Jun 4, 2018

11,809 Downloads

MD5: da4...84b



[HAM10000_metadata.tab](#)

Tabular Data - 810.9 KB

Published Jan 29, 2021

5,938 Downloads

8 Variables, 10015 Observations UNF:6:WcXi...myQ==

Metadata and Image Data for Download From Data Repository

BRAC University Dhaka, Bangladesh Institutional Repository

Digital Collections
Repository

Dspace
<http://dspace.bracu.ac.bd/>



Institutional Repository

BracU IR / School of Data and Sciences (SDS) / Department of Computer Science and Engineering (CSE) / Thesis & Report, BSc (Computer Science and Engineering) / View Item

An efficient deep learning approach to detect skin Cancer



View/Open

20341030, 19141024,
16141014_CSE.pdf (2.208Mb)

Date

2021-09

Publisher

Brac University

Author

Islam, Ashfaqui
Khan, Daiyan
Chowdhury, Rakeen Ashraf

Metadata

Show full item record

URI

<http://hdl.handle.net/10361/15932>

Abstract

Each year, millions of people around the world are affected by cancer. Research shows that the early and accurate diagnosis of cancerous growths can have a major effect on improving mortality rates from cancer. As human diagnosis is prone to error, a deep-learning based computerized diagnostic system should be considered. In our research, we tackled the issues caused by difficulties in diagnosing skin cancer and distinguishing between different types of skin growths, especially without the use of advanced medical equipment and a high level of medical expertise of the diagnosticians. To do so, we have implemented a system that will use a deep-learning approach to be able to detect skin cancer from digital images. This paper discusses the identification of cancer from 7 different types of skin lesions from images using CNN with Keras Sequential API. We have used the publicly available HAM10000 dataset, obtained from the Harvard Dataverse. This dataset contains 10,015 labeled images of skin growths. We applied multiple data pre-processing methods after reading the data and before training our model. For accuracy checks and as a means of comparison we have pre-trained data, using ResNet50, DenseNet121, and VGG11, some well-known transfer learning models. This helps identify better methods of machine-learning application in the field of skin growth classification for skin cancer detection. Our model achieved an accuracy of over 97% in the proper identification of the type of skin growth.

Keywords

Cancer detection; Convolutional neural networks; Image classification; Deep learning

LC Subject Headings

Machine learning; Cognitive learning theory (Deep learning)

Description

This thesis is submitted in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science and Engineering, 2021.

POLICY GUIDELINES

- BracU Policy
- Publisher Policy

Search



☒ Search BracU IR

☐ This Collection

BROWSE

All of BracU Institutional Repository

Communities & Collections

By Issue Date

Authors

Titles

Subjects

This Collection

By Issue Date

Authors

Titles

Subjects

MY ACCOUNT

Login

Register

- Table of Contents
- List of Figures
- List of Tables
- Nomenclature
- Introduction
- Related Work
- Different Types of Skin Cancer
- Dataset Description
- Dataset Pre-processing
- Model Training
- Model Building and Evaluation by CNN Model using Keras Sequential API
- Model Building and Evaluation using RESNET50
- Model Building and Evaluation using DENSENET121
- Model Building and Evaluation using VGG11
- Conclusion
- Bibliography

An Efficient Deep Learning Approach to Detect Skin Cancer

by

Ashfaul Islam
20341030

Daiyan Khan
19141024

Rakeen Ashraf Chowdhury
16141014

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
September 2021

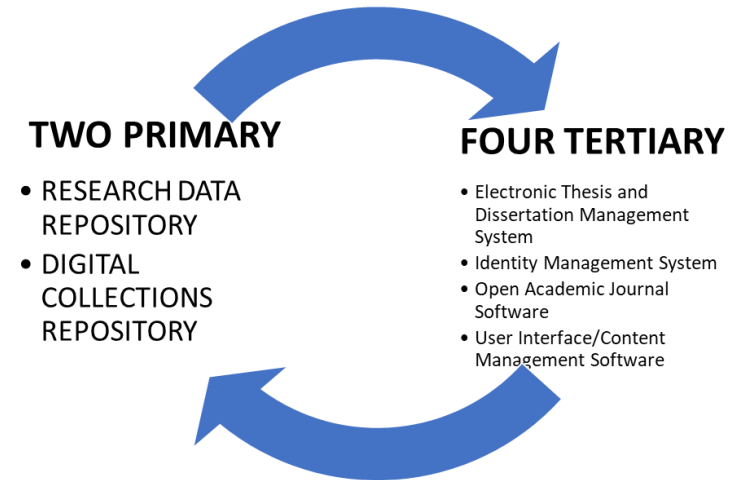
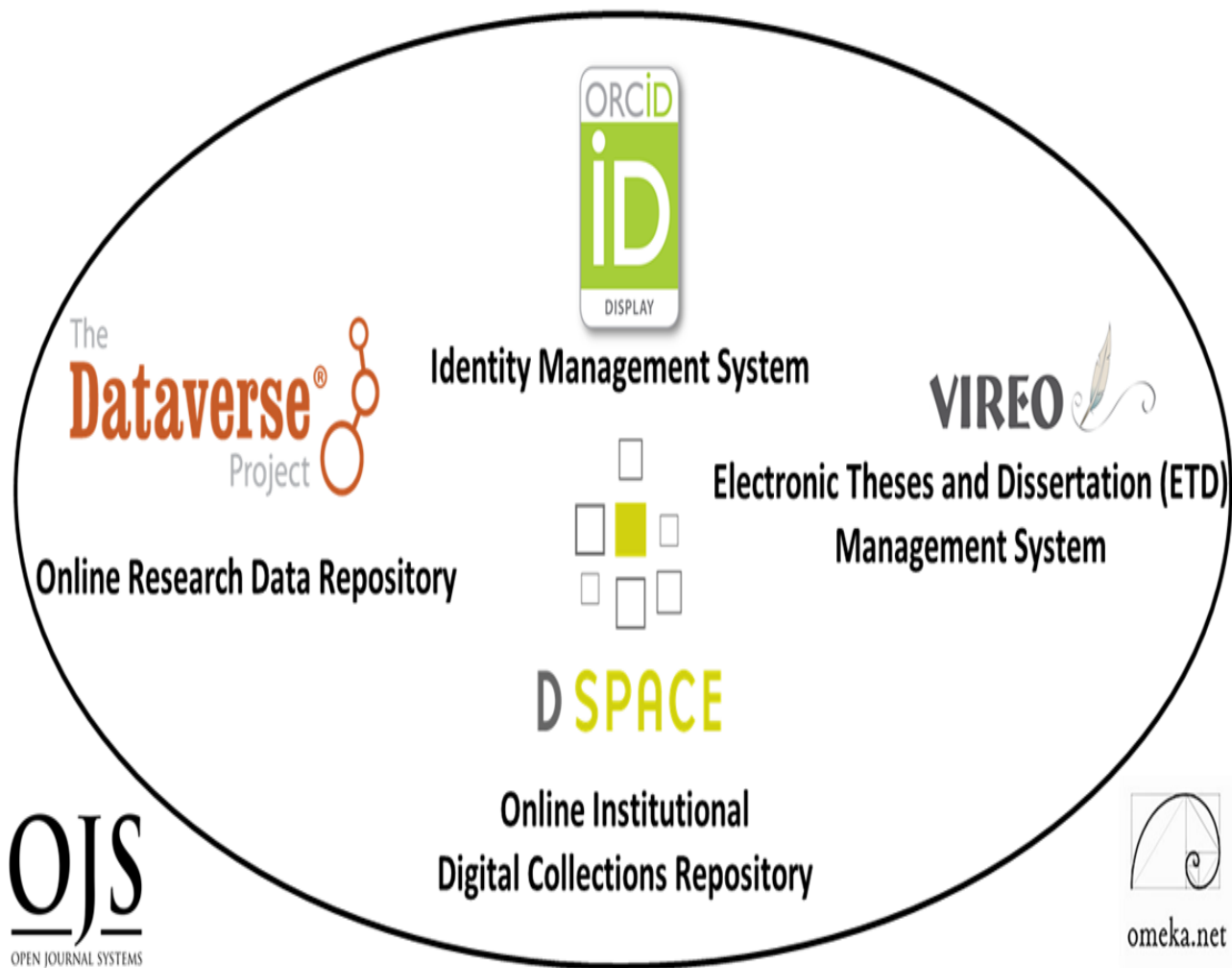
The Progress of Knowledge

2017 Stanford Nature Deep Learning Cancer ID Article

2018 Harvard Dataverse Datarepository Upload
Open Source Vienneesse Dermatological Image Library

November 2021
Dspace Repository
Undergraduate Thesis
BRAC University, Dhaka
Bangladesh, Dept. of
Computer Science and
Engineering
Downloaded July 2022

Digital Scholarship Ecosystem Centered on Research Data Repository and Collections Repository



Many Useful Data Science Skills for AI Will Be Useful Here

Metadata Schemas

Data Organization

Data Cleaning

Data Classification

Creating Dataset Benchmarks

Standardization of Data

Part II: Human Resource Infrastructures (Working Teams)



Future Hires

Machine Learning/Deep Learning/AI Specialist/ Data Scientist and/or AI Librarian (working with the data)

Data Visualization and Analytics Specialist
(Tableau, Bayesia, Power BI)

Committee for Data Repository Workflows & Policies

Onsite Staff Skills

Metadata Specialist/Cataloger
Data Repository Faculty/Student Liaison
Subject Liaisons (Outreach)

Current Staff

Digital Collections Librarian

(Texas State Data Repository Librarian

Dataverse/Publications Repository: D-Space)



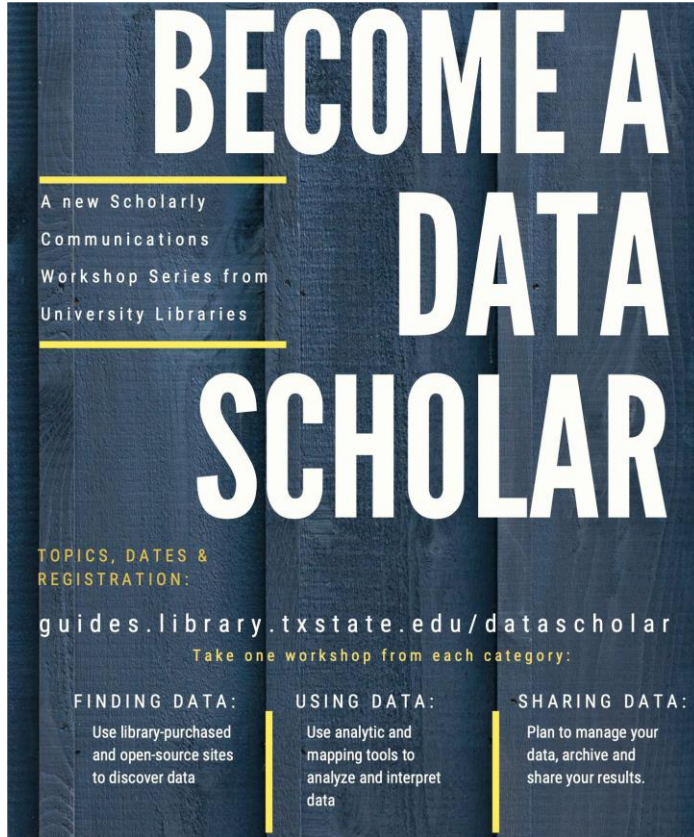
RESPONSIBILITIES: Develop and maintain a data visualization and analytics strategy. Develop strategies to clean and normalize data for use in further analysis. Utilize data visualization strategies to report and present analytics and answer questions related to data analytics and data visualization. Pursue professional development activities to improve knowledge, skills, and abilities and perform special projects and other duties as needed.

- **Required:** Ability to read, analyze, and understand data in a variety of formats; strong written, oral, and interpersonal skills, including ability to work effectively in a team; knowledge of data visualization applications such as PowerBI, Tableau or others; analytical skills; proficiency with Microsoft Excel; ability to utilize analytics/visualization tools in new, creative, and effective ways.

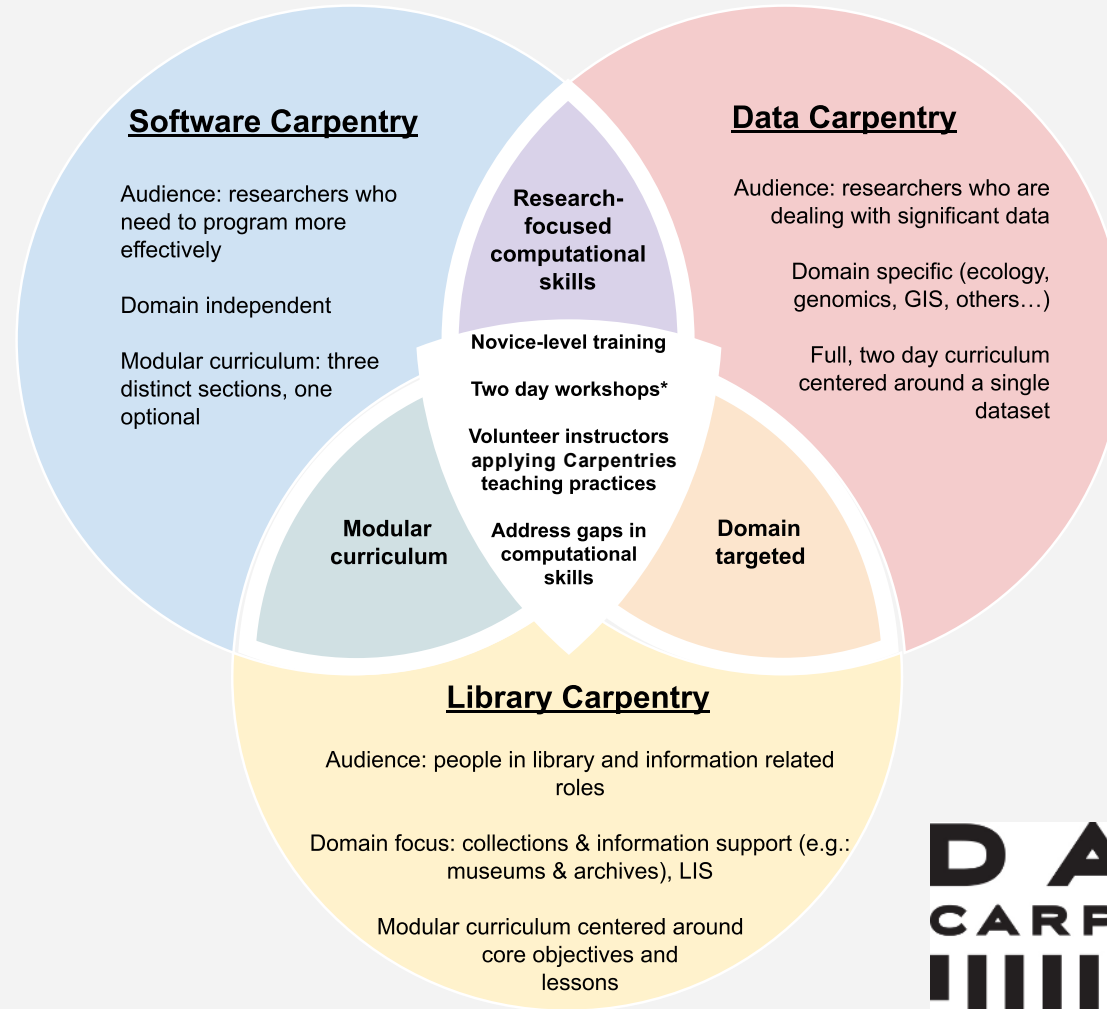
- **Preferred:** Degree in information science, applied statistics, business analytics, computer science or another quantitative or data visualization field; experience with SQL or other query language; experience with R, Python, statistical analysis languages, predictive analytics, and/or AI software.

Further Learning Paths: Data to Carpentries

Foundational Coding and Data Science Skills for researchers Worldwide



Libraries Can
Host Carpentry
Workshops



<https://carpentries.org/>



Conferences and Learning

Library IT and Digital Services May Be Getting Interested in AI



Fantastic Futures

2nd International Conference on AI
for Libraries, Archives and Museums
Stanford Libraries (2019)

Artificial Intelligence

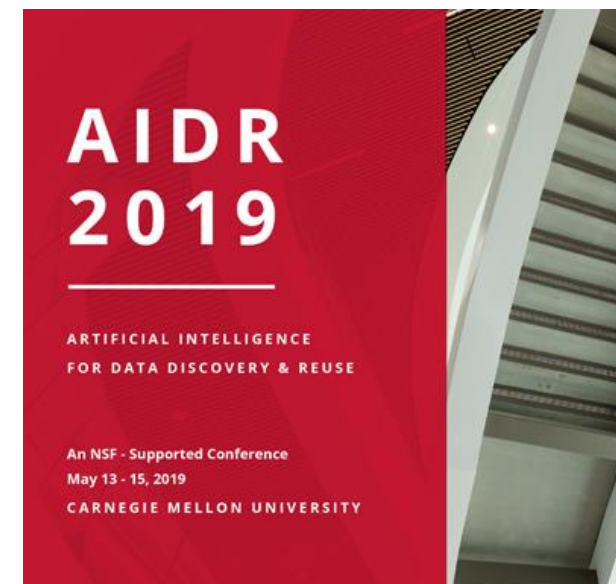
for Data Discovery
& ReUse & Open Science
Symposium (2020), Carnegie Mellon



Texas Conference on Digital Libraries,

Patrice Andre Prud'homme (TCDL)
Oklahoma State,

Computers in Libraries, Yale Art
History, Pixplot (Image Categorization, CNI, C



R&D & Learning, Area 1: Digital and Web Services

Deep Learning Models and Convolutional Neural Nets

- **University Archives**

San Marcos Public
Newspaper Image Negatives
90 years of digitization 800, 000 images

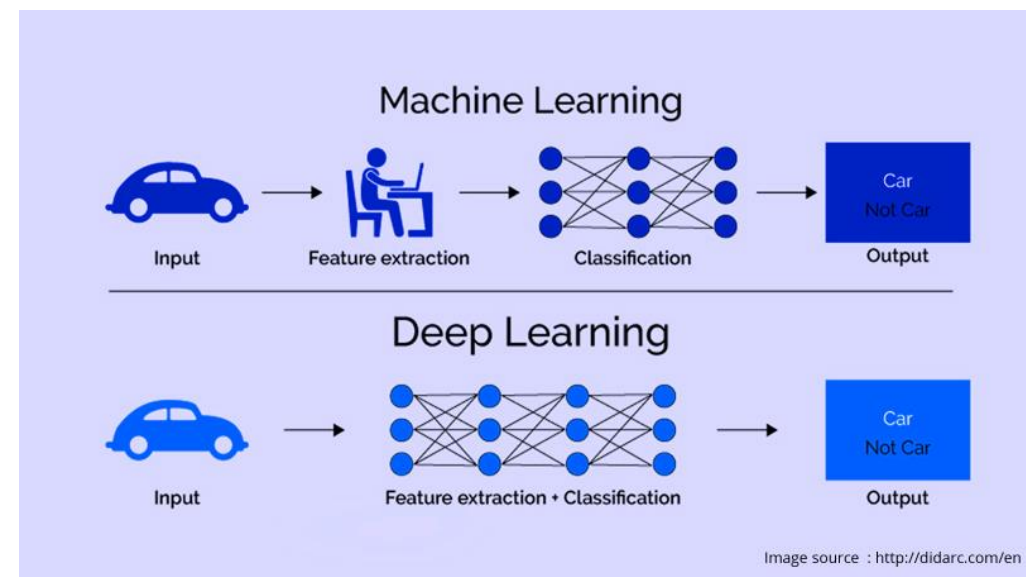
- **Processing Power**
(Compute)

- **Python**

- **Video Cards**
(NVIDIA GPU's)

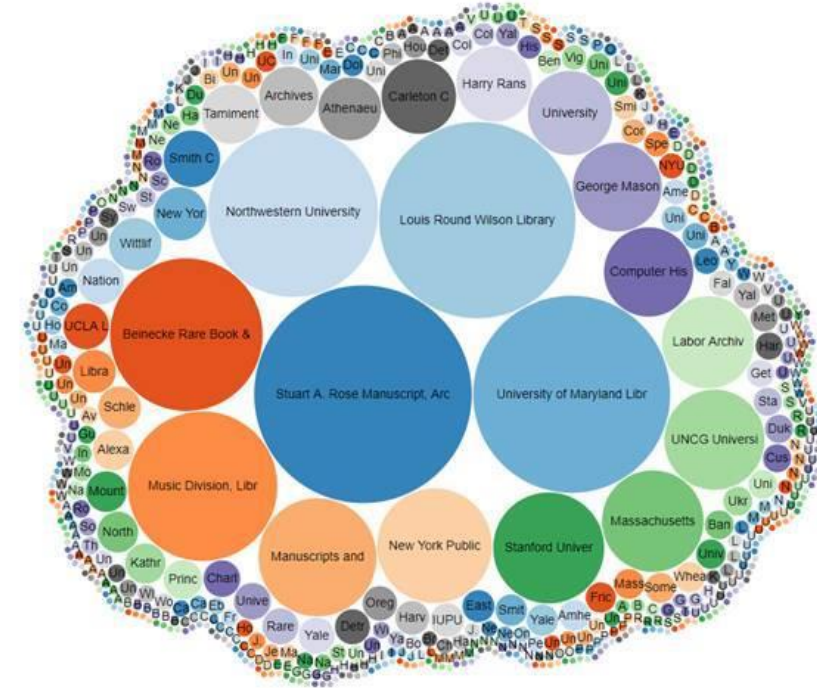
- **Pretrained Models**

- **ResNet, YOLO, COCO**
(200k labeled images, 80 categories)



Area II: Metadata, Wikidata, Semantic Web, AI

- Metadata Services Cataloger
- Crosswalking between Systems
- Successful Linked Data Project
Wikidata Semantic Web Project: Faculty, Oral History Collections, Wittliff Archives, OJS Journals
(Data that is Machine Readable, ie. Google etc)
Moving from MARC Silos to Online (Wikidata)
- Learning Many Data Science Skills, Data Models, Data Batch & Cleaning Tools:
OpenRefine, Quickstatements, Python (7 staff)



Main page
Community portal
Project chat
Create a new item
Recent changes
Random item
Query Service
Nearby
Help
Donate

Lexicographical data
Create a new Lexeme
Recent changes
Random Lexeme

Tools
What links here
Related changes
Special pages
Permanent link
Page information
In Wikipedia

Project page Discussion

Read Edit View

Wikidata:WikiProject PCC Wikidata Pilot/Texas State University Libraries

< Wikidata:WikiProject PCC Wikidata Pilot

A WikiProject for work performed by Texas State University Libraries Staff under the PCC Wikidata Pilot.

Contents [hide]

- 1 Aim and Scope
- 2 Projects
- 3 Contributors
- 4 Models for Current Faculty
 - 4.1 Basic statements
 - 4.2 Extended statements
- 5 Models for Wittliff Archival Collections
 - 5.1 Basic statements
 - 5.2 Extended statements
- 6 Models for University Archives Oral History Collections
 - 6.1 Basic statements
 - 6.2 Extended statements
- 7 Models for Digital Collections Journal Project
 - 7.1 Basic statements
 - 7.2 Extended statements
- 8 Queries
- 9 Useful Resources

Workshops and Training IDEA Institute on Artificial Intelligence

(Recommendation Letter, July, 2022)

- Week Long Fellows Program at University of Texas Austin (20 Fellows)
- Onboarding, Institute, Library Centered AI, Final Project
- Networking with National Library AI Experts and Other Fellows



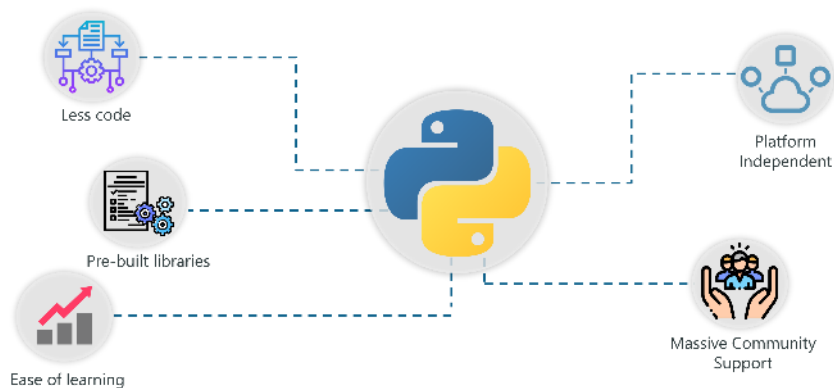
- AI challenges and opportunities, Ethical considerations and guidelines
- UX-Human/AI Interaction Lifecycle
- Existing library, archive, and museum projects
- AI project planning
- Project Design
- Data collection, classification, and transformation
- Roles and implementation
- Python Basics, Python 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



Steps Towards AI: Learning Python, Spring 2022

Hi Ray,

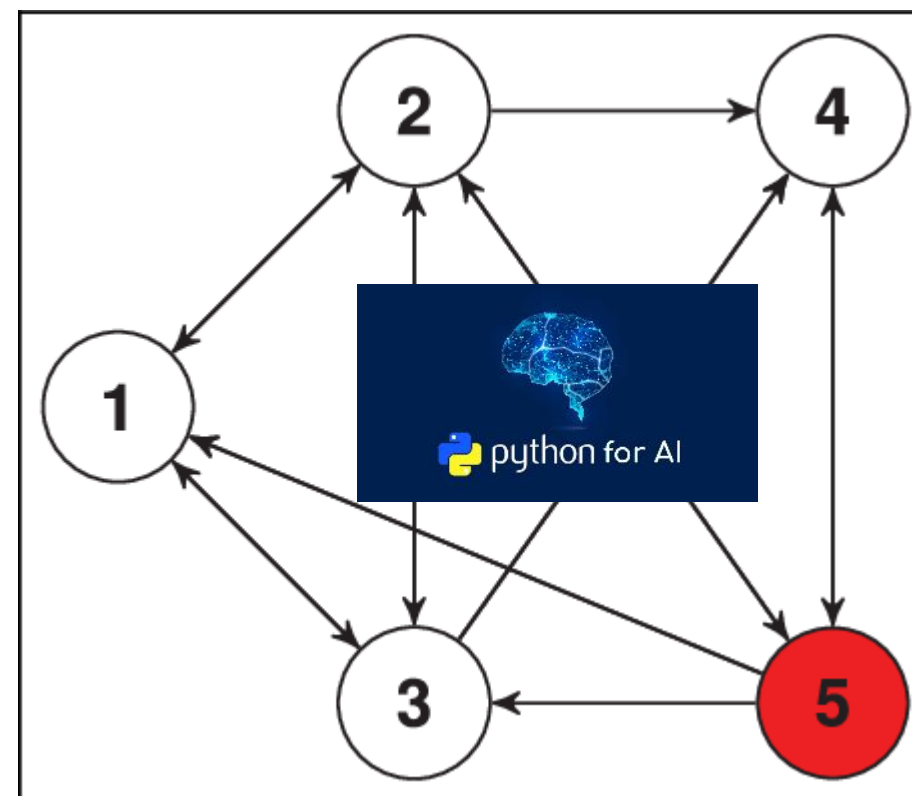
I wanted to let you know that we've started a Teams Group for myself, Carol, Alex and Amanda named "Python learners" so we could share tips and help each other on our various learning paths in an encouraging, safe space. 😊 ---Mary



Courses: Getting up to Speed with Python, Python and Machine Learning
Why Python for AI? – Artificial Intelligence with Python

Carol, Library Management
System (LMS) Usage Data Insights

Mary,
Metadata
AI Extraction



Alexandria
Collections Analytics, Data
Visualization

Amanda,
Collections
Budget,
Insight and
Analytics

Todd/Jason
Image
Recognition
Neural Nets
Part II?

Area III: Ocelot Chatbot Administrator



Transformation of Research and
Information Outreach Services
(Reference & Subject Librarians)



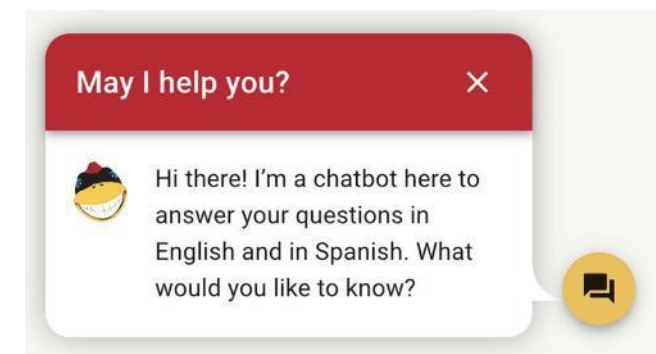
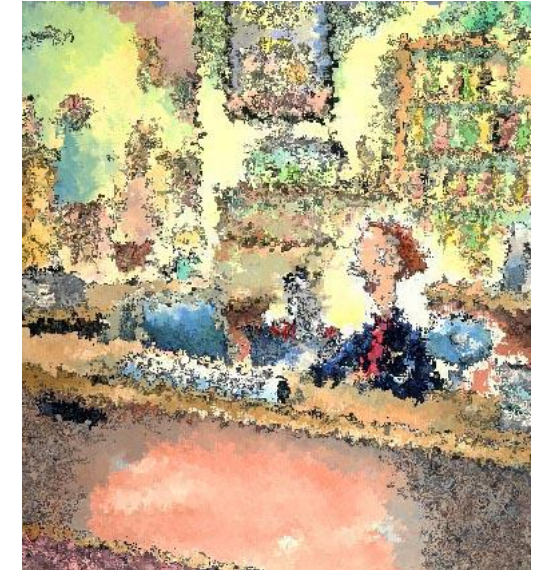
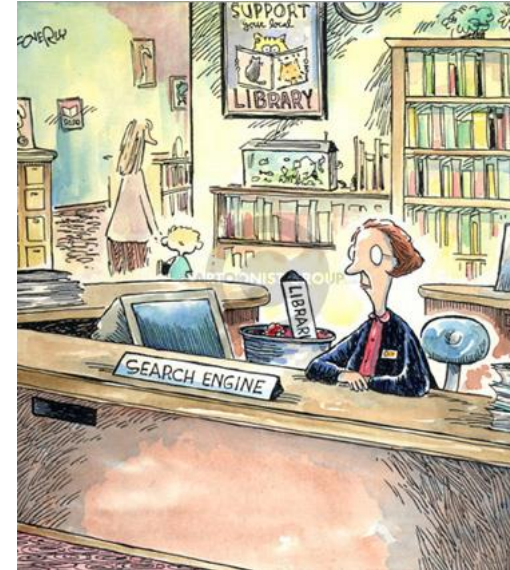
University IT Adopts New Ocelot
Chatbot Infrastructure



Digital and Collection Services
receives New Libraries
Chatbot Administrator



Future Natural Language
Processing R&D
(GPT3-4, DeepMind Gopher)

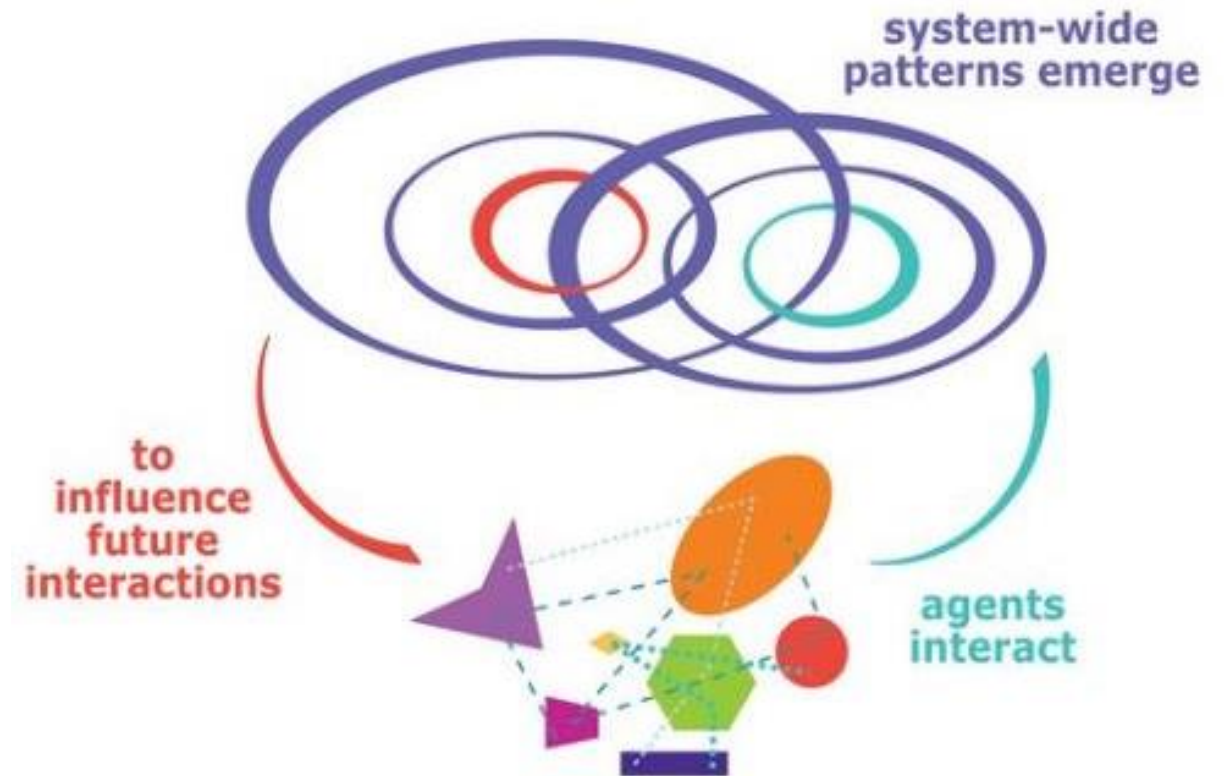


Libraries are Complex Adaptive Non-Linear Dynamic Systems

Emergence, Chaos Theory Complexity, Genetic Algorithms



- Complex Adaptive System (CAS)



TEXAS STATE AI WORKING GROUP

(AIWG)



Purpose: An AI Working Group, information sharing, direction, responsibility and accountability for:

- 1) Artificial Intelligence Projects, Policy Ethics related discussions



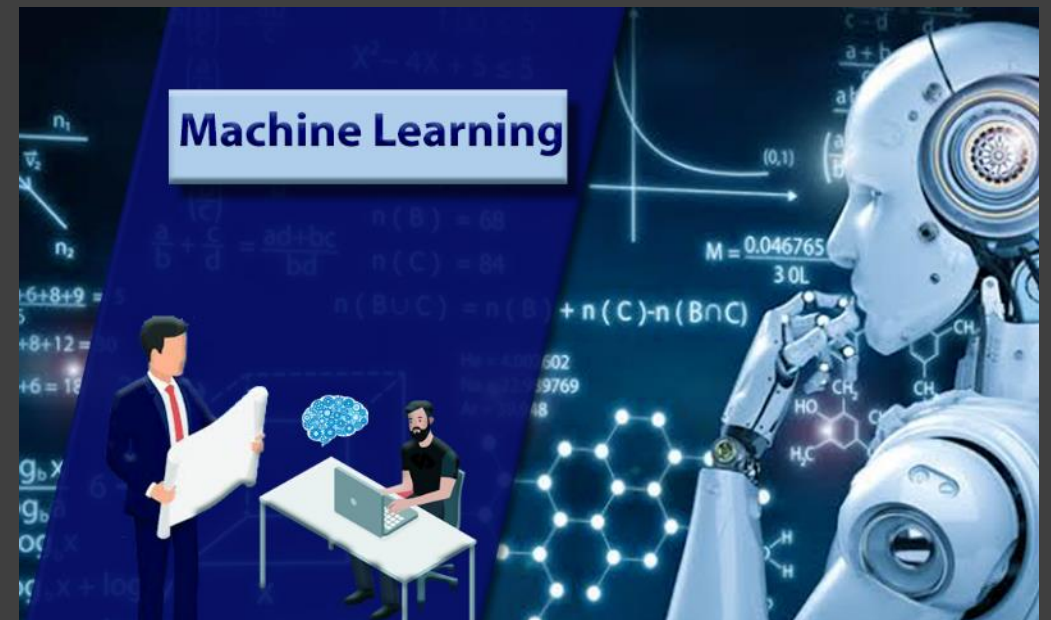
Membership: Metadata, Acquisitions, Digital Services, Special Collections, New Technologies, and Research Services. (9-10 Staff).



Next Steps: A Graduate Student or Two

- Any University Engineering School or Computer Science Department Will Have Graduate Students and Courses Like This:
- E5331 – [Machine Learning for Engineering Applications](#) (2019-2021) [Listed Fall – 2022]
- EE4331 – [Intro to ML for Engineering Applications](#) (2020-2021) [Listed Fall – 2022]

Form a Relationship with the Electrical Engineering Professor and Hire His Graduate Student to Help with your team as part of their final project, graduate theses or part time Research Assistant to provide computational assistance and resources. These can be Masters Candidates or good advanced undergraduate students.



Future Steps: AI Postdoctoral Fellows and Permanent Library AI Positions

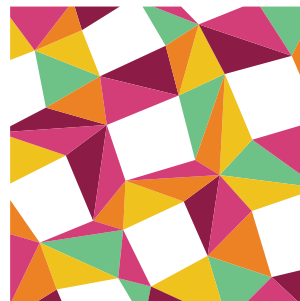
Postdoctoral Fellowship Program offers recent AI/Machine Learning related Ph.D. graduates the chance to develop research tools, resources, and services while exploring new career opportunities and opening Library possibilities.



<https://www.clir.org/>

<https://www.clir.org/global/>

Postdoctoral AI Fellows work with library staff, faculty and graduate students on library related projects that forge and strengthen connections among library collections, archives, special collections digital technologies, and their current AI research and skills.



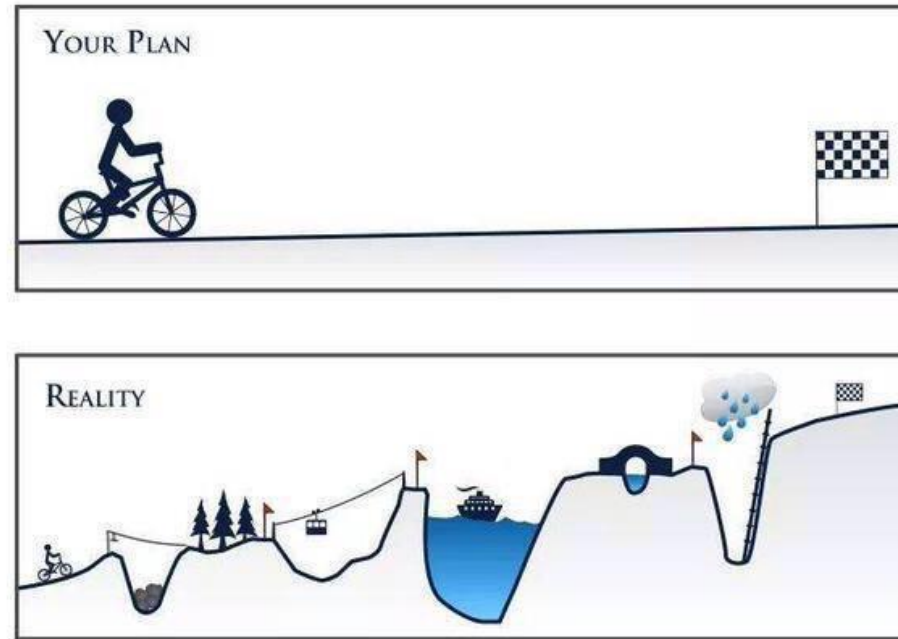
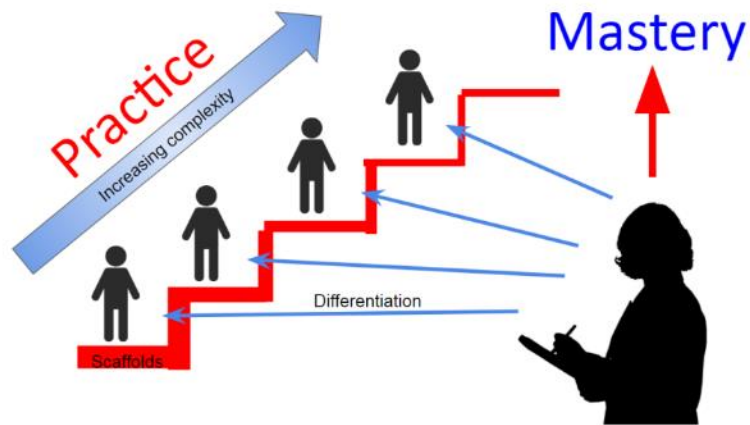
CLIR
POSTDOCTORAL
FELLOWSHIP
PROGRAM

<https://postdoc.clir.org/>

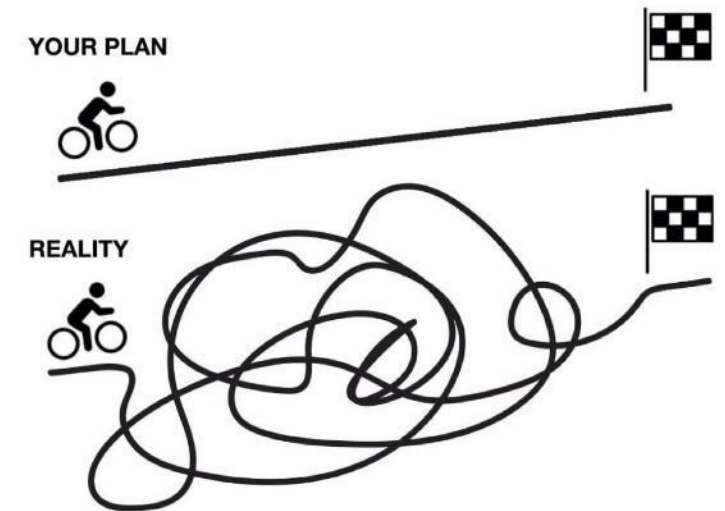


<https://haira.clir.org/blog/>

Steps and Ideas For Scaffolding Towards Library AI Projects and Foundational Infrastructure Success



What reaching success looks like:



Questions/Comments



Ray Uzwyshyn, Ph.D. MBA MLIS
Director, Collections and Digital Services
Texas State University Libraries, USA
ruzwyshyn@txstate.edu , <http://rayuzwyshyn.net>
July 2022

References and Further Resources

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>

Uzwysyn, R. 2022. Online Research Data Repositories and Digital Scholarly Ecosystems: From Research Data and Datasets to Artificial Intelligence and Discovery. IFLA WLIC 2022. Dublin, Ireland. DOI: 10.13140/RG.2.2.12354.86728

---. 2022. **Steps Towards Building Library AI Infrastructures: Research Data Repositories, Scholarly Research Ecosystems and AI Scaffolding**. New Horizons in Artificial Intelligence in Libraries (IFLA Satellite Conference), National University of Ireland, Galway, IR. DOI: 10.13140/RG.2.2.21120.30728

---. 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.

---. 2016. Online Research Data Repositories: The What, When Why and How. *Computers in Libraries*. 36:3, April 2016. pp. 18-21.
<http://rayuzwysyn.net/TXU2016/OnlineDataResearchRepositoriesUzwysyn.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=b31yDNB_cil

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*. TedxBBox, 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=iug8FAI0B1U>

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

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. 2022. *Two Minute Papers*. Recent AI R&D Papers. <https://www.youtube.com/c/K%C3%A1rlyZsolnai/videos>

Machine Learning, Data Science and Python Online Courses

Dabrowski, Anna. Digital Carpentry Materials. [GitHub - ajdabrowski/carpentries-instruction-materials: Materials for teaching Carpentries workshops](https://github.com/ajdabrowski/carpentries-instruction-materials)

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/>

Hinton, Geoff. 2016. *Neural Networks for Machine Learning*. (Coursera Course). https://www.youtube.com/watch?v=cbeTc-Urqak&list=PLoRI3Ht4JOcdU872GhiYWf6jwrk_SNh9

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

Machine Learning Crash Course with TensorFlow API. Google. (Free). <https://developers.google.com/machine-learning/crash-course>

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

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/>

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

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. *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

: