

Research and Applications

ACR's Connect and AI-LAB technical framework

Laura Brink ¹, Laura P. Coombs¹, Deepak Kattil Veetil¹, Kashyap Kuchipudi¹,
Sailaja Marella¹, Kendall Schmidt¹, Sujith Surendran Nair ¹, Michael Tilkin¹,
Christopher Trembl¹, Ken Chang², and Jayashree Kalpathy-Cramer^{2,3}

¹Department of Information Technology, American College of Radiology, Reston, Virginia, USA, ²Department of Radiology, Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Boston, Massachusetts, USA and ³Department of Ophthalmology, University of Colorado School of Medicine, Aurora, Colorado, USA

Corresponding Author: Laura Brink, MSc, Department of Information Technology, American College of Radiology, 1891 Preston White Dr., Reston, VA 20191, USA; lbrink@acr.org

Received 8 August 2022; Revised 11 October 2022; Editorial Decision 16 October 2022; Accepted 31 October 2022

ABSTRACT

Objective: To develop a free, vendor-neutral software suite, the American College of Radiology (ACR) Connect, which serves as a platform for democratizing artificial intelligence (AI) for all individuals and institutions.

Materials and Methods: Among its core capabilities, ACR Connect provides educational resources; tools for dataset annotation; model building and evaluation; and an interface for collaboration and federated learning across institutions without the need to move data off hospital premises.

Results: The AI-LAB application within ACR Connect allows users to investigate AI models using their own local data while maintaining data security. The software enables non-technical users to participate in the evaluation and training of AI models as part of a larger, collaborative network.

Discussion: Advancements in AI have transformed automated quantitative analysis for medical imaging. Despite the significant progress in research, AI is currently underutilized in current clinical workflows. The success of AI model development depends critically on the synergy between physicians who can drive clinical direction, data scientists who can design effective algorithms, and the availability of high-quality datasets. ACR Connect and AI-LAB provide a way to perform external validation as well as collaborative, distributed training.

Conclusion: In order to create a collaborative AI ecosystem across clinical and technical domains, the ACR developed a platform that enables non-technical users to participate in education and model development.

Key words: artificial intelligence, machine learning, software design, radiology, data science

LAY SUMMARY

The American College of Radiology (ACR) has developed a free, vendor-neutral software suite called Connect that provides medical professionals the tools for local data analysis and artificial intelligence (AI) development. With AI-LAB, an application on Connect, users can import medical data, prepare the data, download public pre-trained models, develop new algorithms, and perform statistical evaluation, all while keeping their data safe behind their firewall. Connect and AI-LAB provide a powerful code-free, end-to-end pipeline to democratize AI for all.

INTRODUCTION

The development of artificial intelligence (AI) in radiology and other medical domains has been increasing at a rapid pace in recent years. Deep learning models have shown state-of-the-art performance for various tasks, performing certain interpretation and diagnostic tasks at the level of clinical experts.¹ Key to the success of these algorithms are 3 components: (1) clinical professionals who can drive direction, validation, and translation, (2) data scientists who can design, train, and deploy such algorithms, and (3) the availability of large quantities of diverse, well-annotated patient data. Because such synergy is only accessible within certain academic institutions, it is unsurprising that the barrier to entry to participate in AI development remains high for many radiologists. According to the 2020 American College of Radiology (ACR) Data Science Institute (DSI) Artificial Intelligence Survey, about 30% of responding radiologists use AI as part of their clinical practice, and of those radiologists, most are from larger institutions with collaborative access to scientists.² While some no-code web-based applications have been developed to lower this barrier, none provide a solution that is both end-to-end and code-free.^{3–5} Conversely, data scientists who are not working closely with clinicians may build algorithms that do not make a meaningful impact on patient care.

Access to high-quality patient datasets is also a key hurdle.⁶ Because medical institutions are reluctant to share their data for security and privacy reasons, only a few hospitals participate in the testing and development of AI tools. Furthermore, most AI developers train and test exclusively on their own or a single partner institution's data, omitting some patient groups from training or testing datasets.⁷ As such, AI tools designed using small patient cohorts are often not representative of the larger population. Studies show that AI developed on single-institutional data may not perform well on data from other institutions due to the lack of diversity in the original training dataset.^{8,9} Downstream, algorithms underperforming in certain patient subgroups could reinforce care biases.¹⁰ Thus, there exists a need to enable multi-institutional, collaborative approaches for the purpose of creating robust models, while simultaneously ensuring patient privacy.^{11,12} In order to create equitable and widely applicable AI tools, more radiology professionals and data scientists need to be involved in every step of AI development and have access to diverse patient data.^{13,14}

Launched in 2019, AI-LAB has evolved to meet the needs of radiology professionals and AI developers for a generalizable testing and training framework that maintains data privacy. ACR has developed a free, vendor-neutral software suite called ACR Connect (Connect), to help address some of these difficulties. Connect is a multi-user, multi-service software platform that links clinicians, data scientists, and organizations into the greater ACR network of services. Connect lives on-premise and allows users to keep their sensitive data in-house. AI-LAB is an application within Connect that gives non-technical users tools to train, test, and run AI algorithms without moving the data off-premises. In this article, we discuss the current technical framework behind Connect and AI-LAB and how the software will evolve in the future.

MATERIALS AND METHODS

ACR infrastructure and Connect

ACR has a history of managing important imaging, clinical, and other data for registries and clinical trials. Because institutions have become increasingly concerned about sharing data due to patient

privacy concerns, ACR has developed tools, like Connect, to allow for local processing and distributed testing. [Figure 1](#) shows how Connect fits into ACR's existing data workflow.

Connect is a component-based software platform that supports advanced data processing on-premises allowing sites to join a network of organizations participating in AI development and testing among other cross-organization research initiatives. Given increasing concerns about data security, Connect is designed to give sites complete control over their data. It has a variety of utilities to connect to local systems and manage data, as well as applications that support different project-specific needs. [Figure 2](#) shows the technical design of Connect on-premise and how it interacts with the ACR cloud.

System requirements

Connect can be installed on a stand-alone server, an on-premise virtual machine (VM), or a VM from a cloud provider. Connect is optimized for Linux machines but a pared down version, Connect Windows, supports Windows machines. See [Supplementary Tables S1 and S2](#) for more detailed requirements. When Connect is installed, an administrator account is automatically created. The administrator can create new users on Connect and limit user access to applications or services through the User Management utility.

Importing and managing data

Another utility on Connect is the Digital Imaging and Communications in Medicine (DICOM) service that interfaces with local imaging systems (ie, PACS) to allow users access to their local data. Users can import data into Connect by connecting to their local PACS through either DIMSE (TCP/IP) or DICOMWeb (HTTP). Alternatively, users can import data by moving files to a local drive on the bind mount of the server. Data from electronic health records (EHR) can be ingested into Connect via HL7, FHIR, or as an extract loaded from a drive. Connect can handle most file formats (eg, CSV, JSON, XML).

All imported data can be managed via another utility called Data Manager. Within Data Manager, users can create, edit, and add to datasets. The datasets all have associated objects, or artifacts, which may be accessed or generated by the applications within Connect. An example dataset is shown in [Supplementary Figure S1](#). Once a dataset is created in Data Manager, Connect applications can access it to allow for further processing and use.

Data Manager also has an anonymization pipeline that utilizes anonymization profiles. Users can choose to use the default anonymization profile or create a new profile. An anonymization profile is a set of rules that dictate how the data should be de-identified by calling an API interface for each DICOM tag. For example, users can specify if they want the new Study Date to be empty, incremented by some value, or encrypted. Connect's Image Pixel Cleaner is a simple interface that lets users quickly browse through hundreds of images, identify those with PHI burned into image pixels, draw the area, and clean the images individually or in bulk with a few clicks.

AI-LAB

AI-LAB is an AI toolkit designed for users with little to no experience in programming or machine learning. It is built on cloud-native technology and a public instance, AI-LAB Central, is hosted on ACR's servers.¹⁵ Users come to AI-LAB Central to learn about AI, demo some of AI-LAB's algorithm building and testing, and participate in AI-LAB community activities like AI Challenges. AI-LAB

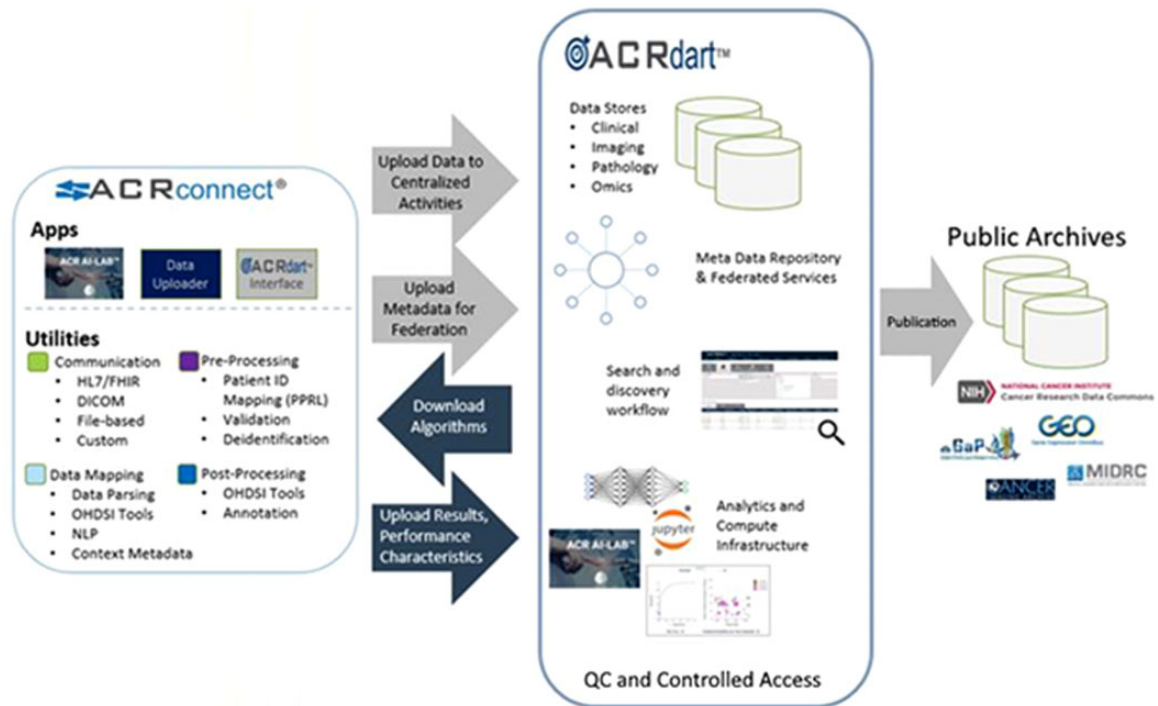


Figure 1. The ACR’s AI research data ecosystem. Connect resides on-premise, providing connectivity to local clinical systems and tools for advanced local processing. A containerized “app model” allows for workflow apps that perform tasks such as preparation of data for upload (eg, de-identification) and advanced local processing (eg, AI-LAB). The local Connect node communicates in a secure manner with ACR cloud services to facilitate the exchange of data, meta data, AI models and the results of AI experiments as appropriate. Central systems, including those powered by ACR’s Data Analysis and Research Toolkit (DART), interact with local nodes and process information to enable registries, clinical trials, and distributed AI activities such as validation and federated learning. Among the many post-processing activities is the preparation of data for safe publication in public archives.

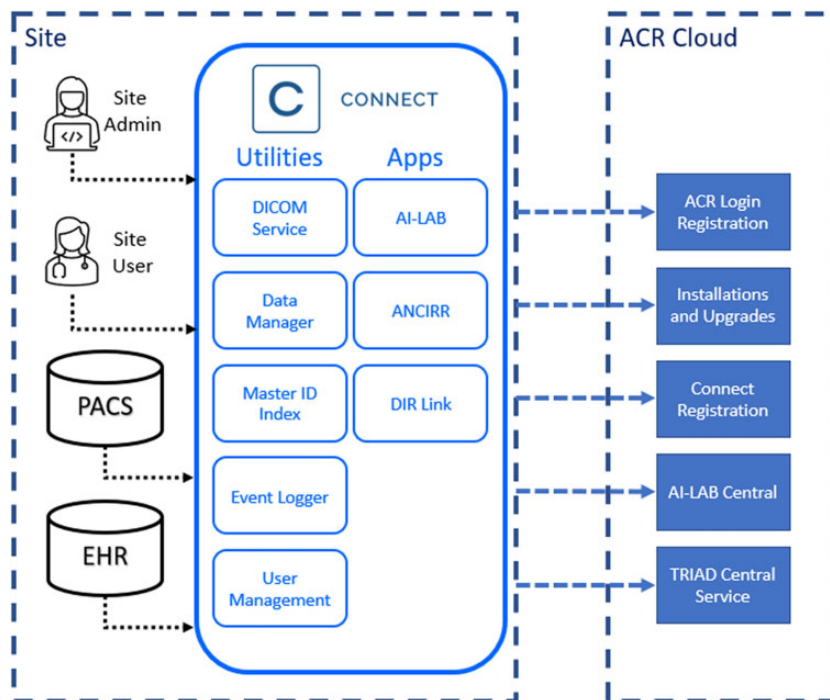


Figure 2. Technical framework of ACR Connect on-premise. Connect talks to various on-premise health IT systems like Picture Archiving and Communication System (PACS), Electronic Health Record (EHR), etc. via standard protocols like Digital Imaging and Communications in Medicine (DICOM), and HL7 Fast Healthcare Interoperability Resources (FHIR). The framework provides utilities to manage users and data. Applications run on the Connect platform like Dose Index Registry (DIR), ACR National Clinical Imaging Research Registry (ANCIRR), and AI-LAB make use of these utilities to allow users/facilities to participate in AI testing and training, clinical registries, and federated learning.

Central also hosts important repositories for the AI-LAB community like the Data Science Institute (DSI) Use Case documents and AI-LAB models. Connect hosts AI-LAB On-prem; a local instance that resides behind the institution's firewall on-premise and is available only to users with the application privileges. Each institution with Connect has their own instance of AI-LAB On-prem that enables users to download and test algorithms on their own data, create algorithms using their own data, and participate in AI-LAB community activities like federated learning (FL). AI-LAB's distributed infrastructure allows data to stay protected behind institutional firewalls while models can be shared among the AI-LAB community.

Learn module

To fully engage with AI, it is important to conceptually understand the fundamentals of how deep learning models are trained. For this purpose, a series of short videos were created and made available in the Learn module. These videos provide a high-level overview of AI and the neural network architectures that power it. Covered topics include a description of the core principles behind training, proper choice of cost function, metrics of performance, and more. To make the verbiage of machine learning less intimidating to beginners, these videos are presented as a conversation between a radiologist and a data scientist, providing continuous translation between the technical and clinical domains.

AI use cases

To ensure that all AI algorithms developed or tested using AI-LAB are clinically relevant, ACR's DSI maintains and creates AI use case documents.¹⁶ Each use case is created by radiologists and allied professionals to ensure that AI developers understand the clinical value, daily workflow, and the potential impact on patient care. Additionally, each use case specifies dataset considerations including relevant procedures, views, and any other important distinguishing factors.

For AI developers, the most important aspect of the use case document is the technical specifications detailing the inputs and outputs that a model must adhere to be clinically useful. AI models that address the same AI use case can be tested on the same dataset and directly compared. Possible inputs include DICOM studies, electronic health records, and pathology reports. The outputs defined in each use case consist of primary (required) and secondary (optional) data elements. Currently, there are 4 supported data element types: classification, continuous, bounding box, and segmentation. Most data elements are linked to a RadElement, defined by ACR and the Radiological Society of North America (RSNA).¹⁷

AI models

AI-LAB is designed to allow AI models to plug and play, enabling users to seamlessly test models developed by different vendors. ACR has created a set of standards for AI models to adhere to in order to be compatible with AI-LAB.¹⁸ ACR's model standards are in line with the recent work by Integrating the Healthcare Enterprise (IHE) Radiology to standardize AI models to increase interoperability.¹⁹ While IHE Radiology outlines multiple different pathways for AI models to integrate with clinical systems, AI-LAB currently supports a specific implementation using Docker containers.²⁰ AI models are expected to be packaged into a Docker container that uses Docker volumes to pass input and output data to and from AI-LAB. ACR's model standards specify how the Docker containers should be packaged, recommended logging practices, input and output formats, as well as mandatory run-time arguments like available GPUs.

An AI model must be packaged in its own Docker container and support at least one use case. AI models can support multiple use cases if the model outputs satisfy the primary outputs for each of the use cases. The supported use cases are specified in the model manifest, a ReadMe.txt file, which must accompany each AI model. The model manifest is in line with the Open Application Model Specifications.²¹ Before adding a new AI model into AI-LAB, model developers must share the Dockerfile and the associated ReadMe.txt with ACR for security checks. Once a model is approved for distribution, it will be added to AI-LAB On-prem "Models" page and users within the AI-LAB community can download it for testing. Optionally, model developers can create a training mode for the Docker container to allow users to train new models with their code on the Create page.

RESULTS

Workflow within AI-LAB

Figure 3 details a user's workflow within AI-LAB to train and test algorithms with their own local data. Assuming a dataset has already been created within the Data Manager utility on the local instance of Connect, users go to the "Prepare" page within AI-LAB first. Users prepare the dataset for a specific AI use case. AI-LAB Central provides the definition of the AI use case, including the structure of the annotations. Dataset artifacts like PNGs or NumPy objects may be generated as part of the data preparation step.

After dataset preparation, users navigate to the "Annotate" page to annotate studies according to the AI use case, which may involve labeling classes, drawing bounding boxes, providing pixel-level segmentations, or counting pathologies (like nodules). Users can review annotations already created or imported through the DICOM Service utility on Connect. AI-LAB currently utilizes the Cornerstone viewer.²²

The annotated dataset is available for model training and testing but a specific model must first be selected. Users can download a model by selecting from the list of available AI models located on AI-LAB On-prem "Models" page.

To train a new model, users navigate to the "Create" page where they can select the initial base model, choose various training hyperparameters, and set an annotated dataset as the training data. Supplementary Figure S2 shows the hyperparameter options for training a Breast Density model.²³ Once model training is complete, users can inspect the accuracy and loss curves of both the training and validation datasets and the performance of the new model on the hold out test dataset. Figure 4 shows an example of the training outputs after a Breast Density training session. Users can choose to share their new local model with the larger AI-LAB community by uploading it to AI-LAB On-prem "Models" page.

To test an AI model on local data, users navigate to the "Evaluate" page. When users initiate an evaluation, the model generates predictions for the local dataset according to ACR's model standards. AI-LAB then compares the model's predictions to the annotations and calculates relevant validation metrics (like a confusion matrix and receiver operating characteristic curve). The results from an evaluation can be uploaded to AI-LAB Central, facilitating distributed validation of models on real-world data.

Users can also try an AI model on a specific study on the "Run" page. Users select a study from a local dataset and can view the model's prediction and associated report text for that particular study. Supplementary Figure S3 shows the prediction of a Breast Density

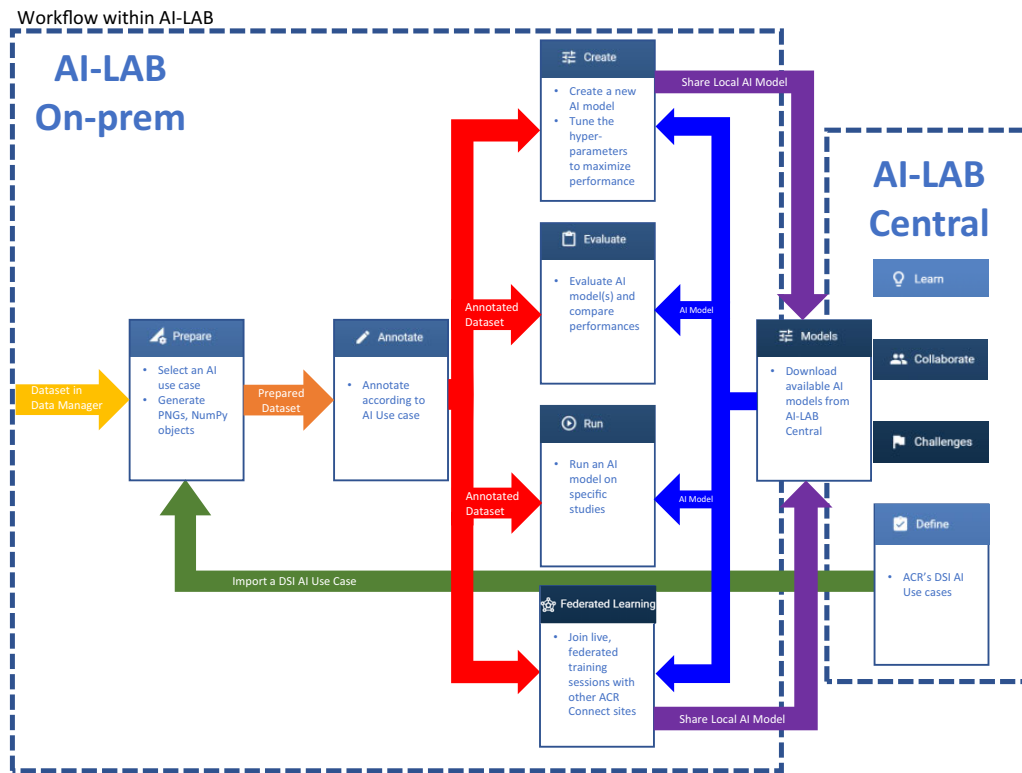


Figure 3. The workflow within AI-LAB for a user to train, test, and run AI models using their local data.

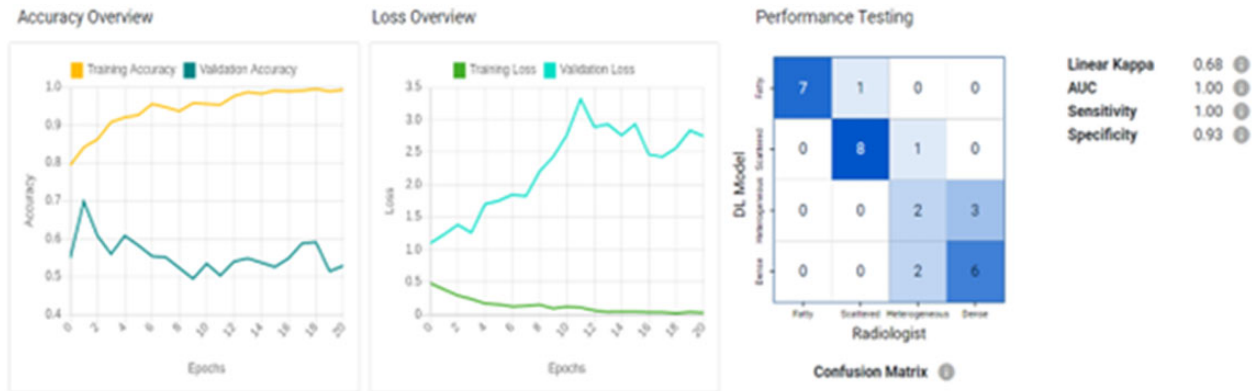


Figure 4. An example of a completed training of a breast density model on the “Create” page. The user can examine the accuracy and loss on the training and validation datasets at each epoch. In this particular training session, the model is most likely overfitting, or memorizing, the training data. The user can see evidence of the overfitting by noting that the training accuracy approaches 1 while the validation accuracy floats around 0.6. A similar trend is seen in the loss overview where the training loss approaches 0 while the validation loss actually increases. The performance testing shows the results of the new model on the hold out test dataset. Studies from the Digital Mammographic Imaging Screening Trial are used for this example.²⁴

model on the selected study. All of these pages on AI-LAB enable users to participate in algorithm testing and training without the need for AI expertise.

Evaluating AI models

AI-LAB utilizes a set of standard validation metrics allowing users to compare the performances of different AI models directly. The code to calculate the metrics is publicly available in order to increase transparency, maintain research integrity, and encourage public contribution.²⁵ These metrics are in line with the work by other groups

like Medical Imaging and Data Resource Center (MIDRC). Within AI-LAB, these metrics are used to calculate all evaluations on the Evaluate page as well as the performance testing after training is completed on the Create page.

The metrics are separated into the 4 data element types: classification, continuous, bounding box, and segmentation. The repository has the most up to date information on the validation metrics by data element type.¹⁸ Another important feature of the metrics is the reporting of unknowns and failures by the model. If the model does not return a prediction or it is not in the correct format, the

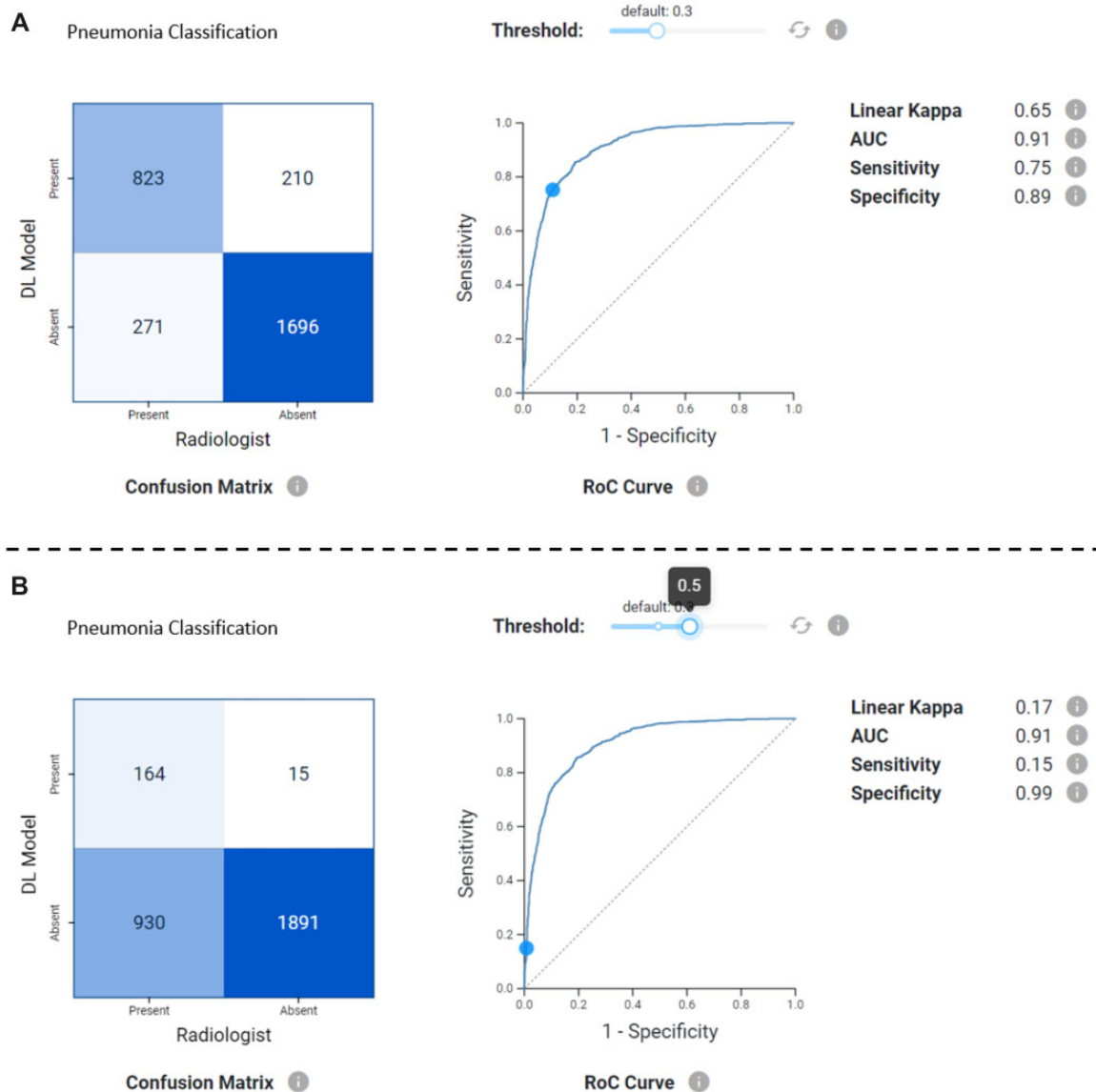


Figure 5. Pneumonia Classification data element’s validation metrics on the “Evaluate” page. For binary classification elements, like pneumonia classification, users can adjust the binary threshold separating the positive and negative classes by using the Threshold slider at the top. A shows the metrics for a binary threshold at 0.3. B shows the metrics for a binary threshold at 0.5. On the left is the interactive confusion matrix. The user can click on a cell in the confusion matrix to view those specific studies. In the middle is the receiver operating characteristic curve. On the right are the classification metrics calculated for the entire dataset. For this particular model, a binary threshold at 0.3 is preferable to 0.5 depending on the target sensitivity/specificity. Studies from 2018 RSNA Pneumonia Detection Challenge are used for this example.²⁷

study is flagged as a “failure”. If the model is unsure of what to predict for a specific study, the model can return a Null value and this is deemed an “unknown”. Users are shown the total number of studies successfully processed, the total number of failures, and the total number of unknowns. The validation metrics shown are for the successfully processed studies.

For the Pneumonia Detection use case defined in [Supplementary Figure S4](#), there is one classification data element and one bounding box element.²⁶ [Figure 5](#) shows the classification metrics for the Pneumonia Classification data element and [Figure 6](#) shows the bounding box metrics for the Pneumonia Bounding Box data element. By separating the metrics into the data element types, the validation metrics are flexible and can be applied to any of the AI use cases.

The “Evaluate” page has several interactive features to enhance user’s engagement. For example, users can click on any of the cells in the confusion matrix to open a viewer and inspect those studies more carefully. This is especially useful when trying to determine if there is a trend or common feature in the studies the model is misclassifying. For binary classification data elements, users can also adjust the binary threshold by sliding the Threshold bar as shown in [Figure 5](#). All the classification metrics update when the binary threshold is adjusted so users can explore the trade-off between sensitivity and specificity with the interactive receiver operating characteristic curve and monitor the changes in the metrics and confusion matrix.

AI-LAB empowers users to question and investigate the performance of models on many different levels. For example, users can monitor an algorithm’s performance over time using the predicted

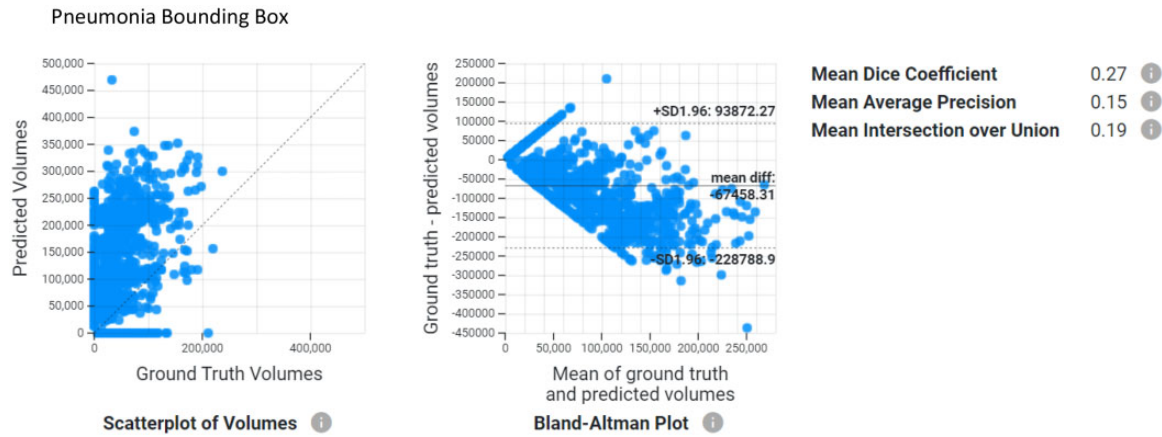


Figure 6. Pneumonia Bounding Box data element’s validation metrics on the “Evaluate” page. On the left is the scatterplot of volumes. This particular model tends to have larger predicted volumes than the ground truth volumes. In the middle is the Bland-Altman plot where the same trend of the larger predicted volumes can be seen. On the right are the bounding box metrics calculated for the entire dataset. Studies from 2018 RSNA Pneumonia Detection Challenge are used for this example.²⁷

probability over time scatterplot. [Supplementary Figure S5](#) shows an example of this scatterplot for the pneumonia classification data element. The plot distinguishes between the true and false positives as well as the true and false negatives and users can adjust the binary threshold to find the optimal separation. By plotting the predictions against the study date, users can monitor for “concept drift”, or changes in the algorithm’s performance over time.²⁸ Users can inspect how the model’s performance changes over time, addressing the dynamic nature of AI algorithms as populations and diseases change.

Federated learning

While the AI-LAB “Create” page provides a way for users to fine-tune available models to their local data, as eluded to earlier, models trained on single-institutional data generally do not perform well at other institutions.^{8,9} FL provides a solution for developing AI models that generalize well across institutions while allowing data to remain at sites throughout model training.

By sharing model weights instead of the data, FL allows model training across a distributed network of sites, each with local datasets which remain secure and private throughout. Models produced through local training are then sent to a central location for aggregation into a single final model. FL has been shown to produce final models which approach the performance levels of centrally trained models.^{29,30} One real-world implementation of FL produced a breast-density classification model that was more accurate on a test dataset from each of the study’s sites than the models produced with each of the site’s local data alone.³¹

With the ability to leverage the Connect network, AI-LAB is ideally suited to host FL sessions. [Supplementary Figure S6](#) diagrams the AI-LAB FL architecture with 3 participating sites. FL through AI-LAB uses a central ACR server to coordinate training. To participate, sites must have Connect On-prem or access to a secure instance in the cloud. To start FL, the central coordinating server distributes a base AI model to each site’s instance of AI-LAB for local training. Each participating site uses their local data to train a new model that is biased for their specific population. Upon completion of training, each site passes their new model back to the central server for aggregation. The server combines the models that are biased for each specific site to generate a single new model that will perform well at all the sites. The server then distributes the new

aggregated model back to each site’s local instance. This process repeats many times to produce a final aggregate model which generalizes well to data from all sites as well as a local model at each site fine-tuned on that site’s data.

AI-LAB in action

At the time of writing this article, there have been over 11.6k views of the videos in AI-LAB and Connect and AI-LAB have been installed at 13 US-based hospitals. In 2019, AI-LAB hosted a crowdsourcing breast density annotation challenge to investigate reader variability and evaluate an AI model.³² Over 3000 annotations were submitted by the crowd (ranging from students to experienced breast radiologists) and showed that the crowd and the AI model had higher agreement than the crowd and the original interpreting radiologists.

There are multiple ongoing experiments within AI-LAB including distributed validation and fine-tuning of a breast density classification algorithm and the inaugural FL experiment to train a generalizable COVID-19 classification model for chest X-rays. The plug and play design of AI-LAB leaves the possibility for future experiments wide open.

DISCUSSION

Here, we present a framework to democratize AI for all individuals at all institutions, fostering an ecosystem for collaboration across clinical and technical domains. To do this, we combined education, annotation, model development/evaluation, data/model sharing, and distributed training on a singular platform in a code-free environment. Importantly, the release of our platform is supported by growing evidence that code-free software can improve the access of AI to both clinicians and researchers.^{33,34} In this article, we reviewed the technical design and framework behind Connect and AI-LAB and how this software suite empowers radiologists to participate in AI development in numerous ways. Users of AI-LAB Central can learn about AI through videos and tutorials, define AI use cases, demo the training and testing of algorithms, and join the AI community through challenges. Connect on-premise enables institutions to easily and securely curate datasets locally behind their own firewall. Through ACR’s model standards and the DSI use cases, new

models can be incorporated easily and shared with the AI-LAB community of sites. Because all of the models are evaluated using the same set of public validation metrics, users can test and investigate the performances of multiple models using their local datasets. With model monitoring features and FL experiments, users can help develop robust models and reduce model biases.

Among the many utilities of this platform, several stand out. The first is education. Very few clinicians have experience with programming languages, let alone machine learning. Conversely, very few data scientists have the expertise to identify and define pertinent clinical problems. This platform translates technical jargon into understandable concepts. In tandem, clinical problems can be transcribed into computer algorithm inputs and outputs. In this coupling, data scientists can be educated in clinical medicine and clinicians can be educated in data science, with the goal of minimizing loss of key information at the interface. Another advantage is the preservation of patient privacy. The rapid growth of big data approaches have made patient privacy and data ownership a major priority.¹³ Indeed, this is one of the reasons why most published machine learning models have been trained on patient data from a single institution.⁶ The platform provides a way to perform external validation as well as collaborative distributed training without patient data leaving the local sites. Additionally, as a hub for AI models, users are able to adapt models trained by other users to their own local data via transfer learning. For example, a model developed for chest radiographs at one institution might be a good initial model to fine-tune for abdominal radiographs at another institution. Finally, while large academic hospitals may have access to large datasets, clinicians who have dedicated research time, and data scientists to develop algorithms, individuals and other hospitals likely will not have all these components. The platform allows anyone with interest to fully engage in AI development.

In the future, Connect will be adding more control over specific user access within the platform. Administrators will be able to add restrictions to datasets and models. AI-LAB will allow multiple annotations by different users to help mitigate inter-rater variability. AI-LAB will add filtering functionality to the validation metric calculations. Users will investigate biases of AI models by selecting subgroups from their datasets. For example, users might investigate whether a model performs better or worse on male vs. female patients. Users can filter the dataset by sex and examine how the metrics change depending on the sub-group. Additionally, users can investigate many combinations of sub-groups, like African-American females over the age of 65 years, ensuring that all potential biases are examined. By incorporating sub-group filtering, AI-LAB will continue to enable users to monitor health disparities and inequities within AI models.

Connect and AI-LAB will empower individuals to develop ideas, datasets, and models that will provide benefit beyond their institution. In this way, AI can move beyond large academic medical centers, to everyone in the clinical and research community. Physicians at community hospitals and private practices would be able to tap into this collective resource and implement their own ideas to meet their specific clinical needs. The democratization of AI represents an investment in the community, which in turn will improve the care provided to patients.

CONCLUSION

In this technical framework, we described the data ecosystem within Connect and AI-LAB. While the data remains safely behind the insti-

tution's firewalls, AI-LAB enables model sharing among institutions to facilitate distributed validation and FL. We also discussed the importance of standards for AI use cases, AI models, and AI validation metrics. We hope to continue to enhance the features in Connect and AI-LAB to further empower our users to participate in AI development and testing.

AUTHOR CONTRIBUTIONS

All authors contributed to the design and implementation of Connect and AI-LAB.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *JAMIA Open* online.

ACKNOWLEDGMENTS

We would like to thank Medical Imaging and Data Resource Center (MIDRC) for their support in developing the predicted probability over time scatterplot and the thumbnail view of the confusion matrix within AI-LAB.

CONFLICT OF INTEREST STATEMENT

JK-C has received research funding to Massachusetts General Hospital from GE and Bayer.

DATA AVAILABILITY

The data underlying this article are available in the article and in its online [supplementary material](#).

REFERENCES

1. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019; 25 (1): 44–56.
2. Allen B, Agarwal S, Coombs L, Wald C, Dreyer K. 2020 ACR Data Science Institute artificial intelligence survey. *J Am Coll Radiol* 2021; 18 (8): 1153–9.
3. Lösel PD, van de Kamp T, Jayme A, et al. Introducing Biomedisa as an open-source online platform for biomedical image segmentation. *Nat Commun* 2020; 11 (1): 5577.
4. Burkhardt J, Sharma A, Tan J, et al. N-Tools-Browser: web-based visualization of electrocorticography data for epilepsy surgery. *Front Bioinform* 2022; 2: 857577.
5. Egger J, Wild D, Weber M, et al. Studierfenster: an open science cloud-based medical imaging analysis platform. *J Digit Imaging* 2022; 35 (2): 340–55.
6. Langlotz CP, Allen B, Erickson BJ, et al. A roadmap for foundational research on artificial intelligence in medical imaging: from the 2018 NIH/RSNA/ACR/The Academy Workshop. *Radiology* 2019; 291 (3): 781–91.
7. Kim DW, Jang HY, Kim KW, Shin Y, Park SH. Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: results from recently published papers. *Korean J Radiol* 2019; 20 (3): 405–10.
8. AlBadawy EA, Saha A, Mazurowski MA. Deep learning for segmentation of brain tumors: impact of cross-institutional training and testing. *Med Phys* 2018; 45 (3): 1150–8.
9. Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. *PLoS Med* 2018; 15 (11): e1002683.

10. Seyyed-Kalantari L, Zhang H, McDermott MBA, Chen IY, Ghassemi M. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. *Nat Med* 2021; 27 (12): 2176–82.
11. Chang K, Balachandar N, Lam C, *et al.* Distributed deep learning networks among institutions for medical imaging. *J Am Med Inform Assoc* 2018; 25 (8): 945–54.
12. Rieke N, Hancox J, Li W, *et al.* The future of digital health with federated learning. *NPJ Digit Med* 2020; 3 (119): 119.
13. Allen B, Agarwal S, Kalpathy-Cramer J, Dreyer K. Democratizing AI. *J Am Coll Radiol* 2019; 16 (7): 961–3.
14. Recht MP, Dewey M, Dreyer K, *et al.* Integrating artificial intelligence into the clinical practice of radiology: challenges and recommendations. *Eur Radiol* 2020; 30 (6): 3576–84.
15. AI-LAB Central. <https://ailab.acr.org>. Accessed March, 2022.
16. Define AI Use Cases. <https://www.acrdsi.org/DSI-Services/Define-AI>. Accessed March, 2022.
17. RadElements. <https://www.radelement.org/>. Accessed March, 2022.
18. AI-LAB Model Standards. https://github.com/ACRCode/AI-LAB_documentation/wiki/AI-LAB-Inference-Model-Standards. Accessed March 2022.
19. Genereaux B, Bialecki B, Diedrich K, *et al.* *AI Interoperability in Imaging. Integrating the Healthcare Enterprise Radiology* [White Paper]; 2021. https://www.ihe.net/uploadedFiles/Documents/Radiology/IHE_RAD_White_Paper_AI_Interoperability_in_Imaging_Rev1-0_PC_2021-03-12.pdf.
20. Docker. <https://www.docker.com/>. Accessed March, 2022)
21. OAM. <https://oam.dev/>. Accessed March, 2022.
22. Cornerstone. <https://cornerstonejs.org/>. Accessed March, 2022.
23. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition; 2016: 770–8; Las Vegas, NV.
24. Pisano ED, Gatsonis C, Hendrick E, *et al.*; Digital Mammographic Imaging Screening Trial (DMIST) Investigators Group. Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 2005; 353 (17): 1773–83.
25. AI-LAB Evaluation Metrics Code. <https://github.com/ACRCode/acrconnect-ai-lab-evaluation-metrics>. Accessed March, 2022.
26. Pneumonia Use Case. <https://www.acrdsi.org/DSI-Services/Define-AI/Use-Cases/Pneumonia>. Accessed March, 2022.
27. Shih G, Wu C, Halabi S, *et al.* Augmenting the National Institutes of Health Chest Radiograph Dataset with expert annotations of possible pneumonia. *Radiol Artif Intell* 2019; 1 (1): e180041.
28. Pinykh OS, Langs G, Dewey M, *et al.* Continuous learning AI in radiology: implementation principles and early applications. *Radiology* 2020; 297 (1): 6–14.
29. Sheller MJ, Edwards B, Reina GA, *et al.* Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Sci Rep* 2020; 10 (1): 12598.
30. Sadilek A, Liu L, Nguyen D, *et al.* Privacy-first health research with federated learning. *NPJ Digit Med* 2021; 4 (1): 132.
31. Holger R, Chang K, Singh P, *et al.* Federated learning for breast density classification: a real-world implementation. *ArXiv*. 2020 doi: 10.1007/978-3-030-60548-3_18.
32. Chang K, Beers AL, Brink L, *et al.* Multi-institutional assessment and crowdsourcing evaluation of deep learning for automated classification of breast density. *J Am Coll Radiol* 2020; 17 (12): 1653–62.
33. Korot E, Guan Z, Ferraz D, *et al.* Code-free deep learning for multi-modality medical image classification. *Nat Mach Intell* 2021; 3 (4): 288–98.
34. Faes L, Wagner SK, Fu DJ, *et al.* Automated deep learning design for medical image classification by health-care professionals with no coding experience: a feasibility study. *Lancet Digit Health* 2019; 1 (5): e232–42.