My Desktop
Prepare & Submit Proposals
Prepare Proposals (Limited proposal types)
Demo Site: Prepare Proposals
Prepare Proposals in FastLane
Proposal Status
Awards & Reporting
Notifications & Requests
Project Reports
Award Functions
Manage Financials
Program Income Reporting
Grantee Cash Management Section Contacts
Administration
Lookup NSF ID

Preview of Award 1726188 - Annual Project Report

Cover |
Accomplishments |
Products |
Participants/Organizations |
Impacts |
Changes/Problems
| Special Requirements

Cover

Federal Agency and Organization Element to Which Report is Submitted:	4900
Federal Grant or Other Identifying Number Assigned by Agency:	1726188
Project Title:	MRI: High Performance Digital Pathology Using Big Data and Machine Learning
PD/PI Name:	Joseph Picone, Principal Investigator Tunde Farkas, Co-Principal Investigator Iyad Obeid, Co-Principal Investigator Yuri Persidsky, Co-Principal Investigator
Recipient Organization:	Temple University
Project/Grant Period:	01/01/2018 - 12/31/2021
Reporting Period:	01/01/2020 - 12/31/2020
Submitting Official (if other than PD\PI):	N/A
Submission Date:	N/A
Signature of Submitting Official (signature shall be submitted in accordance with agency specific instructions)	N/A

Accomplishments

*** What are the major goals of the project?**

As shown in Figure 001 (attached), there are three major goals of this project:

- Phase 1: Hardware Acquisition
- Phase II: Data Development
- Phase III: Algorithm Development

This first phase consists of hardware procurement, development and installation. This required managing a complex multi-organizational relationship between the vendor (Aperio/Leica Biosystems), the Temple Hospital Information Technology group

(TUHS-IT) and Temple University Main Campus Information Technology (TUMC-IT). Subgoals for this phase of the project included:

- Procurement of Scanner
- Installation and Verification
- User Training
- Procurement of Network Storage
- Installation and Verification
- Final Hardware Certification

The second phase of the project consisted of data development. This involved working closely with the vendor to certify and integrate the slide scanner hardware and software. Subgoals for this phase of the project included:

- Preliminary Archival Scanning
- User Acceptance Testing
- Production Archival Scanning
- Workflow Integration
- User Acceptance Testing
- Production Scanning
- IRB Application
- Preliminary Database Release
- User Acceptance Testing
- Production Database Release

The third phase of the project consisted of algorithm development. This involved the development of a deep learning-based system to automatically classify data. Subgoals for this phase of the project included:

- Pilot Experiments
- System Tuning
- System Performance Analysis
- Physician Feedback
- Final System Performance Evaluation
- Physician Acceptance Testing

*** What was accomplished under these goals and objectives (you must provide information for at least one of the 4 categories below)?**

Major Activities: In the third year of this project, we originally planned to focus on software development and annotation of data, while we kept data collection running as a mature task. Unfortunately, due to COVID-19, data collection was more or less suspended in March 2020 because it was no longer possible for our undergraduates to work at Temple University Hospital. At the time we were on track to reach 100,000 slides by the end of 2020.

We currently have digitized over 60,000 slides. These can be categorized as follows:

1. Breast Tumor (6%)
2. Urinary Prostate (36%)
3. Gastrointestinal (18%)
4. Lymph Nodes (8%)
5. Other (32%)

We have completed annotation of the Breast Tumor subset (3,640 slides) and will release these in early 2021. We are in the final stages of deidentifying the medical reports, which will be released with the data. This subset totals 1.23 Terabytes of data. The average image size is 372 Mbytes, which confirms the fact that these are extremely high-resolution images.

Part of our preparations to release this data included an extensive reconfiguration of our web server to support the dissemination of these large datasets. While this activity is a detail beyond the scope of the report, it was an important step to allow people to efficiently download such a large amount of data. We increased the amount of web space on the web server and reconfigured the anonymous rsync capability so that

multiple downloads could be managed efficiently using advanced disk caching strategies.

We also began annotating the next largest subset of the data - urinary prostate slides. To do this, we had to learn how to classify these slides. We created a collaboration with the Pathology Department at Fox Chase Cancer Center (FCCC), a center of excellence for cancer treatment in the U.S. Dr. Yulan Gong met with our undergraduates and trained them to annotate these slides. This training was completed at the end of 2020. We are now in the early stages of annotating this data.

In the process of creating this collaboration, we reached an agreement with FCCC to digitize pathology slides in their tumor bank, collaborating with Dr. Denise Connolly. FCCC maintains one of the largest repositories of this type of data in the country. It is extremely well-curated. We spent 9 months working through IRB issues with them (this is always a very slow process), but should begin transferring slides in Spring 2021 and continue digitization throughout 2021. The agreement for this data sharing has been approved by Temple University and is awaiting approval by FCCC. This is an enormous opportunity since these slides are very well documented and will serve as important ground truth data for machine learning. It will also allow us to expand the types of stains seen in our database, which will make the machine learning problem more difficult and yet more clinically relevant.

We are in the process of collaborating with FCCC on some future research opportunities in the pathology area and are very excited about the opportunity to bring them into this research area. Their pathology expertise will be extremely valuable.

We are developing an extensive annotation guidelines document that includes examples of each type of annotation performed. The first version of this document will be released with the breast tumor subset. It will be publicly available and open for community review. A preliminary version of this document is available at this URL: https://www.isip.piconepress.com/publications/reports/2021/tuh_dpath/annotations/. However, this document is not yet ready to be widely disseminated. It is interesting to note this document was written by the undergraduate annotation team, which is an extremely competent group of students.

We have also made significant progress on the development of software to automatically classify images. We have been developing deep learning-based whole slide imaging (WSI) approaches that operate on these high-resolution images directly. Common analysis methods that are routinely performed by pathologists are the determination of histological grade and the hormone receptor status by immunohistochemistry (IHC). These analyses can be tedious and error-prone. Since the estimation of the number of cells that are positive for a particular antigen and the degree of positivity needs analysis of immunohistochemically stained slides based on their staining intensity, quantitative automatic image analysis methods can be used highly more effectively in comparison to conventional glass slides observations.

The microscopic examination of a tissue specimen to find the signs of disease is called histopathology. To do so, a surgical specimen or biopsy is put onto glass slides. Since the raw biopsy does not have distinctive colors, staining is used by pathologists to make each component of tissue in different colors and easily distinguishable. A common method of staining is done with hematoxylin and eosin (H&E). Cell nuclei get blue with hematoxylin and cytoplasm and connective tissue get pink by eosin. Breast tumor excisions or biopsies are fixed in formalin and embedded in paraffin. Then sections with a thickness of 3-5 μm are cut from paraffin blocks with microtome; which is a high precision cutting instrument. These sections are mounted on glass slides. The most interesting parts of the sample for pathologists are nuclei and cytoplasm. These parts cannot be visible on glass slides, so staining will be used to highlight them. Hematoxylin and eosin (H&E) are the most common materials which are used to make nuclei

blue/purple and eosin to make other structures such as cytoplasm, stroma, etc. pink. Other advanced staining techniques are not very common, such as IHC which highlights estrogen and progesterone receptors (ER and PR).

There are two types of slides that have frequently been used as input images in research and literature; the fine needle tissue microarrays (TMA) and whole slide images (WSI). TMAs were very common in literature since they are much smaller than WSI images. A typical image in TMA is about 2K in 2K pixels, while WSIs are 50K in 50K pixels. Smaller images demand less computational and storage capacity but are useless for clinical studies. Our focus is on WSIs, which are what is used in the vast majority of clinical settings.

Because of the size of the images, typically 50K pixels in one dimension and at least 25K pixels in another, we use several data reduction techniques to make machine learning more efficient. The images are preprocessed and segmented to identify regions of interest. This is done at full resolution to preserve important details. We use morphological techniques known as dilation and erosion to define areas of interest.

Images are then converted into features using a variety of well-known feature extraction techniques such as edge lengths, roundness, homogeneity and other shape-related parameters. An overview of the deep learning process used is summarized in Figure 002 attached below. In Figure 003, we summarize the training process used to train our classifiers. In Figure 004, we present the evaluation process used to classify images, postprocess the classification system's raw output and evaluate the results.

The accuracy on the development set was measured to be 97% while the accuracy on an open set test, referred to as the evaluation set, was 93%. These results are very promising given the difficulty of this data and are in the range of what would be considered clinically acceptable. We are very excited about these results.

All code for this system is written in Python and uses standard libraries such as PyTorch.

Specific Objectives: The specific objectives for the third year of this project were:

1. Digitize 100,000 images by December 2020.
2. Annotate and release several subsets of data.
3. Release software to automatically classify images.
4. Acquire user feedback on the value of the software.

Significant Results: We have digitized 60,000 slides and were on track for 100,000 before COVID-19. We have requested, and received, an extension to the project to allow us to continue digitizing images.

We have also created a data pipeline with Fox Chase Cancer Center that should pay long-term dividends. Slide scanning should commence in Spring or Summer 2021 depending on COVID-19.

We are on the verge of releasing our first large subset of data - the Breast Tumor Corpus - which will contain 3,640 annotated images.

We have developed software that can classify breast tumor data with high accuracy. We are working to release this software in early 2021.

Key outcomes or Other achievements: The breast tumor data is a substantial contribution because nothing like it exists in the open source. We expect it will quickly become quite popular.

The creation of a strategic partnership with FCCC should pay dividends in the long run as it gives us access to another data source.

*** What opportunities for training and professional development has the project provided?**

This project has employed four undergraduates in three capacities: web developer, application programmer and software engineer. The student doing web development had the opportunity to learn state of the art methods in full stack web development including using the package Boost to develop and maintain web pages. The application engineer position had the opportunity to learn a broad range of skills including real-time DSP programming, streaming interfaces, Unix shell programming and, of course, implementation of signal processing and deep learning in Python. The software engineering position involved learning how to design extensible software systems than facilitate the integration and adaptation of research software. Two graduate students were employed on the project. They were tasked with the development of the software tools, which involved how to apply and optimize deep learning software in Python. All students in the project also learned how to organize and manage data sets. All students have also been trained on how to present their work at a professional conference, and how to document their work. We use software management systems and a weekly reporting structure that emphasize good technical communication.

*** Have the results been disseminated to communities of interest? If so, please provide details.**

All data, resources, and software are available from the project web site as described in the original proposal. The project web site is located at https://www.isip.piconepress.com/projects/nsf_dpath/. The data is also crossreferenced on our generate data and resources site: https://www.isip.piconepress.com/projects/nedc/. We provide direct downloads and also support an anonymous rsync server (nedc@www.piconepress.com).

We also maintain a listserv we use to communicate with people. Anyone can sign up at this URL:

https://www.isip.piconepress.com/projects/nsf_ccri_dpath/html/request_access.shtml

*** What do you plan to do during the next reporting period to accomplish the goals?**

The main challenge right now is determining when we will be able to return to the hospital. It seems like it is safe to say we will be back there by the summer, which means we should reach our goal of digitizing 100,000 slides - hopefully more.

We expect to officially release these things:

- The Breast Tumor Corpus (3,640 slides)
- The Urinary Tract Corpus (10,000 slides)
- an annotations guidelines document for both corpora
- our software to automatically classify images

We also expect to annotate and release Fox Chase data. The nature of this depends on a number of factors and will be decided once the IRB is in place.

Supporting Files

Filename	Description	Uploaded By	Uploaded On
figure_001.pdf	The project timeline from the original proposal.	Joseph Picone	01/26/2021
figure_002.pdf	An overview of the layered deep learning approach based on CNNs that we use to classify images.	Joseph Picone	01/26/2021
figure_003.pdf	An overview of the training process used for our WSI image classification system.	Joseph Picone	01/26/2021
figure_004.pdf	An overview of the evaluation process used to classify images and evaluate them efficiently.	Joseph Picone	01/26/2021

Products

Books

Book Chapters

Nabila Shawki, M. Golam Shadin, Tarek Elseify, Luke Jakielaszek, Tunde Farkas, Yuri Persidsky, Nirag Jhala, Iyad Obeid, and Joseph Picone (2020). The Temple University Hospital Digital Pathology Corpus. *Machine Learning Applications in Medicine and Biology 1st.* 1. Obeid, Iyad Selesnick, Ivan Picone, Joseph. Springer. New York City, New York, USA. 1. Status = PUBLISHED; Acknowledgement of Federal Support = Yes ; Peer Reviewed = Yes ; DOI: 10.1007/978-3-030-36844-9.

Inventions

Journals or Juried Conference Papers

View all journal publications currently available in the NSF Public Access Repository for this award.

The results in the NSF Public Access Repository will include a comprehensive listing of all journal publications recorded to date that are associated with this award.

Campbell, C. and Mecca, N. and Duong, T. and Obeid, I. and Picone, J.. (2018). Expanding an HPC Cluster to Support the Computational Demands of Digital Pathology. *IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. 1 (1) 01 to 03. Status = Deposited in NSF-PAR doi:https://doi.org/10.1109/SPMB.2018.8615614 ; Federal Government's License = Acknowledged. (Completed by Picone, Joseph on 01/26/2021) Full text Citation details

Hunt, I. and Husain, S. and Simon, J. and Obeid, I. and Picone, J.. (2019). Recent Advances in the Temple University Digital Pathology Corpus. *IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. 1 (1) 1 to 4. Status = Deposited in NSF-PAR doi:https://doi.org/10.1109/SPMB47826.2019.9037859 ; Federal Government's License = Acknowledged. (Completed by Picone, Joseph on 01/26/2021) Full text Citation details

Houser, D. and Shadhin, G. and Anstotz, R. and Campbell, C. and Obeid, I. and Picone, J. and Farkas, T. and Persidsky, Y. and Jhala, N.. (2018). The Temple University Hospital Digital Pathology Corpus. *Proceedings of the IEEE Signal Processing in Medicine and Biology Symposium*. 1 to 7. Status = Deposited in NSF-PAR doi:10.1109/SPMB.2018.8615619 ; Federal Government's License = Acknowledged. (Completed by Picone, Joseph on 11/03/2019) Full text Citation details

Hunt, I., Husain, S., Simons, J., Obeid, I., & Picone, J. (2019). Recent Advances in the Temple University Digital Pathology Corpus. In I. Obeid & J. Picone (Eds.), Proceedings of the IEEE Signal Processing in Medicine and Biology Symposium (SPMB) (pp. 1–3). Retrieved from https://www.isip.piconepress.com/publications/unpublished/conferences/2019/ieee_spmb/dpath/.. Status = PUBLISHED.

Shawki, N., Elseify, T., Cap, T., Shah, V., Obeid, I., & Picone, J. (2020). A Deep Learning-Based Real-time Seizure Detection System. Proceedings of the IEEE Signal Processing in Medicine and Biology Symposium (SPMB) (pp. 1–4). Philadelphia, Pennsylvania, USA.. Status = AWAITING_PUBLICATION.

Licenses

Other Conference Presentations / Papers

Other Products

Other Publications

Patent Applications

Technologies or Techniques

Thesis/Dissertations

Websites or Other Internet Sites

Participants/Organizations

What individuals have worked on the project?

Name	Most Senior Project Role	Nearest Person Month Worked

Name	Most Senior Project Role	Nearest Person Month Worked
Picone, Joseph	PD/PI	1
Farkas, Tunde	Co PD/PI	0
Obeid, Iyad	Co PD/PI	1
Persidsky, Yuri	Co PD/PI	0
Khalkhali, Vahid	Graduate Student (research assistant)	3
Battalora, Leo	Undergraduate Student	3
Campbell, Chris	Undergraduate Student	1
Duong, Thuc	Undergraduate Student	0
Elseify, Tarek	Undergraduate Student	1
Houser, Devin	Undergraduate Student	0
Hunt, Isabel	Undergraduate Student	1
Husain, Saiyeda	Undergraduate Student	2
Jakielaszek, Luke	Undergraduate Student	0
Jean-Paul, Shmyrde	Undergraduate Student	3
Liang, Dennis	Undergraduate Student	2
Makholia, Paras	Undergraduate Student	1
Mecca, Nicholas	Undergraduate Student	0
Shadhin, Golam	Undergraduate Student	0
Shaw, Skyler	Undergraduate Student	2
Simons, Julien	Undergraduate Student	3
Teperov, Josh	Undergraduate Student	2
Tulin, Nikita	Undergraduate Student	2
Vorwick, Lynn	Undergraduate Student	4
Wevodau, Zoe	Undergraduate Student	1

Full details of individuals who have worked on the project:

Joseph Picone
Email: joseph.picone@gmail.com
Most Senior Project Role: PD/PI
Nearest Person Month Worked: 1

Contribution to the Project: Project manager and technical lead. Responsible for making sure the data meets the needs of machine learning researchers.

Funding Support: None.

Change in active other support: No

International Collaboration: No
International Travel: No

Tunde Farkas
Email: tunde.farkas@temple.edu
Most Senior Project Role: Co PD/PI
Nearest Person Month Worked: 0

Contribution to the Project: Subject matter expert responsible for training the student data annotators and ensuring that the technology meets the needs of pathologists.

Funding Support: None.

Change in active other support: No

International Collaboration: No
International Travel: No

Iyad Obeid
Email: iobeid@temple.edu
Most Senior Project Role: Co PD/PI
Nearest Person Month Worked: 1

Contribution to the Project: Supported the development of annotation standards and training of the data annotators.

Funding Support: None.

Change in active other support: No

International Collaboration: No
International Travel: No

Yuri Persidsky
Email: yuri.persidsky@temple.edu
Most Senior Project Role: Co PD/PI
Nearest Person Month Worked: 0

Contribution to the Project: Subject matter expert responsible for training the student data annotators and ensuring that the technology meets the needs of pathologists.

Funding Support: None.

Change in active other support: No

International Collaboration: No
International Travel: No

Vahid Khalkhali
Email: vahid.khalkhali@temple.edu
Most Senior Project Role: Graduate Student (research assistant)
Nearest Person Month Worked: 3

Contribution to the Project: algorithm software

Funding Support: None

International Collaboration: No
International Travel: No

Leo Battalora
Email: leo.battalora@gmail.com
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 3

Contribution to the Project: system administration (IT); hardware installation and deployment; IT support for personnel working on the project.

Funding Support: None.

International Collaboration: No
International Travel: No

Chris Campbell
Email: tuf46810@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 1

Contribution to the Project: system administration (IT); hardware installation and deployment; IT support for personnel working on the project.

Funding Support: None.

International Collaboration: No
International Travel: No

Thuc Duong
Email: tug98850@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 0

Contribution to the Project: system administration (IT); hardware installation and deployment; IT support for personnel working on the project.

Funding Support: None.

International Collaboration: No
International Travel: No

Tarek Elseify

Email: tug35668@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 1

Contribution to the Project: web design; baseline system development

Funding Support: None.

International Collaboration: No
International Travel: No

Devin Houser
Email: tuf89323@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 0

Contribution to the Project: Managed data collection at the hospital; operate the digital scanning equipment for hospital pathologists; organize medical reports; anonymize the data.

Funding Support: None.

International Collaboration: No
International Travel: No

Isabel Hunt
Email: tuj85563@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 1

Contribution to the Project: Managed data collection at the hospital; operate the digital scanning equipment for hospital pathologists; organize medical reports; anonymize the data.

Funding Support: None other than this award.

International Collaboration: No
International Travel: No

Saiyeda Husain
Email: tuh11491@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 2

Contribution to the Project: Managed data collection at the hospital; operate the digital scanning equipment for hospital pathologists; organize medical reports; anonymize the data.

Funding Support: None

International Collaboration: No
International Travel: No

Luke Jakielaszek
Email: tug52339@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 0

Contribution to the Project: web design; baseline system development

Funding Support: None

International Collaboration: No
International Travel: No

Shmyrde Jean-Paul
Email: tuh26880@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 3

Contribution to the Project: computer system administrator / IT support

Funding Support: None

International Collaboration: No
International Travel: No

Dennis Liang
Email: tul14986@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 2

Contribution to the Project: Web developer responsible for the dissemination of information about the project

Funding Support: None

International Collaboration: No
International Travel: No

Paras Makholia
Email: tuj81485@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 1

Contribution to the Project: Web developer responsible for the dissemination of project information

Funding Support: None

International Collaboration: No
International Travel: No

Nicholas Mecca
Email: tuf89560@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 0

Contribution to the Project: system administration (IT); hardware installation and deployment; IT support for personnel working on the project.

Funding Support: None

International Collaboration: No
International Travel: No

Golam Shadhin

Email: tug69453@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 0

Contribution to the Project: Managed data collection at the hospital; operate the digital scanning equipment for hospital pathologists; organize medical reports; anonymize the data.

Funding Support: None

International Collaboration: No
International Travel: No

Skyler Shaw
Email: tuk32485@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 2

Contribution to the Project: system administration (IT); hardware installation and deployment; IT support for personnel working on the project.

Funding Support: None.

International Collaboration: No
International Travel: No

Julien Simons
Email: tuk39975@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 3

Contribution to the Project: Managed data collection at the hospital; operate the digital scanning equipment for hospital pathologists; organize medical reports; anonymize the data.

Funding Support: None.

International Collaboration: No
International Travel: No

Josh Teperov
Email: josh.teperov@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 2

Contribution to the Project: system administration (IT); hardware installation and deployment; IT support for personnel working on the project.

Funding Support: None.

International Collaboration: No
International Travel: No

Nikita Tulin
Email: tug47034@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 2

Contribution to the Project: Database developer, data annotation, operations manager

Funding Support: None

International Collaboration: No
International Travel: No

Lynn Vorwick
Email: tug70217@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 4

Contribution to the Project: software engineer; image recognition system development

Funding Support: None

International Collaboration: No
International Travel: No

Zoe Wevodau
Email: tuk59951@temple.edu
Most Senior Project Role: Undergraduate Student
Nearest Person Month Worked: 1

Contribution to the Project: Database developer / data annotator

Funding Support: None

International Collaboration: No
International Travel: No

What other organizations have been involved as partners?

Name	Type of Partner Organization	Location
Fox Chase Cancer Center (FCCC)	Other Nonprofits	Philadelphia, PA

Full details of organizations that have been involved as partners:

Fox Chase Cancer Center (FCCC)

Organization Type: Other Nonprofits
Organization Location: Philadelphia, PA

Partner's Contribution to the Project:
In-Kind Support
Collaborative Research

More Detail on Partner and Contribution: FCCC, as described in the report, helped train our students on data annotation and will be providing data (pathology slides).

Were other collaborators or contacts involved? If so, please provide details.

Connolly, Denise C <Denise.Connolly@fccc.edu>
Wu, Hong <Hong.Wu@fccc.edu>

Impacts

What is the impact on the development of the principal discipline(s) of the project?

The impact of the project has been limited because we haven't made our first release of a large chunk of data that can be used for machine learning experiments. We have made pilot releases of the data to favorable reviews, and we have published, so the community is aware of the project. We expect to have more impact in the next reporting period.

What is the impact on other disciplines?

Once our data is released, it will be of value to several communities, including machine learning, pathology and bioengineering. We expect to host formal competitions centered around the data to further increase the impact of the data.

What is the impact on the development of human resources?

Our pool of annotators come from fields such as biochemistry that don't always get exposed to engineering problems. The students on the annotation team acquire a wide range of skills including Linux computing, Python programming and data science in general. Often this makes them want to pursue careers in the computational sciences upon graduation. It definitely impacts their view of medicine and science even if they stay in their discipline.

What was the impact on teaching and educational experiences?

Since all of the data is indexed in a database, the database can be used by pathologists to study cases similar to their current cases, and it can be used by medical students for training in manual interpretation of pathology slides.

What is the impact on physical resources that form infrastructure?

As part of this project, we have constructed the first computer network at Temple University that spans the main campus, a HIPAA-restricted research network, and the hospital's operational HIPPA-controlled network. It took a long time to make this happen, but this is one of the few networks of its types. Our students can move data across each of these networks without being physically located at the hospital. This is not a great technical achievement, but simply a demonstration of persistence. It involved solving a number of complex firewall issues. It puts us in a very unique position with respect to data collection.

What is the impact on institutional resources that form infrastructure?
Nothing to report.

What is the impact on information resources that form infrastructure?

As explained above, we have worked very closely with campus and hospital IT teams to achieve a very unique virtual network that allows us to move HIPPA-compliant data. Many other research groups at Temple will benefit from the path we have opened with this project.

What is the impact on technology transfer?
Nothing to report.

What is the impact on society beyond science and technology?
Nothing to report.

What percentage of the award's budget was spent in a foreign country?

None.

Changes/Problems

Changes in approach and reason for change

Obviously, COVD-19 has impacted our ability to collect data at the hospital. It is still unclear when undergraduate student workers will be allowed to return to the hospital. From March 2020 through December 2020, students have not been able to

work at the hospital. We have focused on analysis and preparation of the existing data we have collected, doing so using a secure VPN connection to the hospital.

Actual or Anticipated problems or delays and actions or plans to resolve them

We are still evaluating the COVID-19 situation. This is why we requested and were granted a one-year extension. Student workers will probably not be able to return to the hospital until they are vaccinated. The timetable for that is uncertain at present. In the meantime, we are doing as much remotely through a secure VPN connection as possible.

Changes that have a significant impact on expenditures

Our undergraduate labor charges have been much lower than expected due to the hospital access issues previously described. We are still within our budget guidelines, but expenditures are occurring at a slower than expected rate. Hence, we requested and received an extension.

Significant changes in use or care of human subjects
Nothing to report.

Significant changes in use or care of vertebrate animals
Nothing to report.

Significant changes in use or care of biohazards
Nothing to report.

Change in primary performance site location
Nothing to report.

Special Requirements

Responses to any special reporting requirements specified in the award terms and conditions, as well as any award specific reporting requirements.