2020

Smithsonian
DIGITIZATION PROGRAM OFFICE

It's hard to encapsulate a year such as 2020 without resorting to well-worn terms such as "unprecedented," "trying times," or "New Normal." For the Digitization Program Office, the term that best describes the year is "steadfast." The DPO team remained unwavering in its efforts to create the digital building blocks that bring Smithsonian collections to the world, even as that world closed down all around us.

The year began with pride and excitement after the successful February release of the Smithsonian's Open Access collection, when three million digital images, several thousand 3D models, and 14 million metadata records were made available to the world without restrictions. But as we moved into March, the reality of the pandemic took hold, the Smithsonian shut down, and its staff began teleworking. Like everyone else, we had to adapt quickly.

This proved particularly true for our mass digitization team. With digital imaging work on hold due to facility closures, this team accelerated its informatics efforts to focus on collections records digitization and metadata enhancement, bringing new and "tried and true" automated workflows, tools, and resource management to the task. Their efforts will yield a number of big "wins" for the Smithsonian: broader and deeper access to our digitized collections, more quality information for the public, and greater potential for those using digitized collections as research datasets.

Our 3D team remained steadfastly focused on creating robust 3D infrastructure for the Smithsonian. Their work will increase the amount of 3D digitization we can undertake when we get back on site, ensure that the Smithsonian's 3D data is sustainable into the long-term, and allow us to use Smithsonian 3D models on new and emerging online platforms. As an example of the latter, the 3D team was able to place a number of 3D models in augmented reality (AR) experiences with internal and external partners in 2020, and they will build on these experiences in the future.

Our Policy and Analysis program continued to conduct the annual digitization assessment, which provides metrics on digitization progress across the Smithsonian, and oversaw other assessment activities such as Unit Digitization Plans and a Data Management Plan focusing on risk assessment.

None of our work is possible without our wider Smithsonian family. The team in the Office of the Chief Information Officer (OCIO) deserves accolades for its extraordinary efforts during the pandemic that enabled all of us to begin teleworking immediately after the Institution shut down. The DPO is an office in OCIO and has long been the recipient of the talent and expertise that exists across its programs. This year was no different.

We also thank Bill Tompkins and his team at the National Collections Program, who collaborate with DPO's mass digitization, digitization assessment efforts, and so much more. The NCP's work and outreach to the collections community is a role model for many of the efforts underway in DPO.

The Smithsonian museums, archives and libraries ("units") deserve considerable thanks for helping enable digitization efforts across the Institution. From planning through implementation, we relied on their professional expertise and support to bring more of their collections online.

I am writing this note on January 7, 2021, one day after the siege on our nation's Capitol, still in the midst of a pandemic, and with 2021 looking as uncertain as 2020. As unsettling as this is, the DPO team is committed to creating and delivering Smithsonian digitized collections in ways that support the mission of the Institution and its goals to meet the moment: fostering conversations about race and social injustice, supporting educational needs, and making the Smithsonian a more responsive and digitally engaged institution.

Diane Zorich
Director, Digitization Program Office, OCIO

Despite the challenges that 2020 posed, the DPO's 3D team had a productive year that kicked off with an all-hands effort around the Smithsonian's Open Access Initiative, an unprecedented release of Smithsonian content and the largest single release of open access 3D models from a museum. In support of this Initiative, the 3D team migrated 2,248 3D models of Smithsonian objects from legacy tools to the Smithsonian's 3D viewer, Voyager, which was developed in-house. This migration enables Smithsonian 3D content to be integrated into central Smithsonian IT infrastructure such as EDAN and the EDAN Drupal module. Smithsonian 3D content now can be displayed on si.edu or any Drupal-based Smithsonian web property, and it is available on our open access APIs. In just the first month after the launch of the Smithsonian's Open Access Initiative, Smithsonian 3D content was viewed over 400,000 times, and over 40,000 3D models were downloaded.

The 3D team's support of the Open Access Initiative went beyond the contribution of 3D assets and IT infrastructure upgrades. The team also brought eight partners to the launch, including the Khronos Group, a leading 3D graphic industry consortium of over 150 companies; Autodesk Tinkercad, Autodesk's free online 3D software suite and a popular platform used by educators; Duke University's MorphoSource, a leading academic digital repository of 3D content; and Sketchfab, a prominent platform for viewing and sharing 3D digital assets.

The 3D team also collaborated with both the National Museum of Natural History and the National Museum of American History. These projects included the publication of dozens of 3D models of corals and related animals, the publication of 3D models of objects from the "Girlhood, It's Complicated" exhibit, and the safe and successful 3D digitization of objects for the Latino Center's upcoming "Pleibol" exhibit. The 3D team partnered with Instagram to release a set of AR filters and with Tinkercad on a #TinkerTogether Design Challenge, both based on Smithsonian 3D objects. The team continued outreach and public engagement speaking at IST Archiving, DrupalCon Global, The Pipeline Conference, DrupalGovCon, and ToScA.

Efforts to develop open-source software to help manage, process, and distribute the Smithsonian's growing collection of 3D content continued over the course of the year. Jamie Cope joined the team as lead software developer, managing the development of both Voyager, the DPO's online 3D viewer, and Cook, the DPO's automated 3D data processing platform. Cook was critical to successfully migrating 3D content in preparation for the launch of the Smithsonian Open Access initiative. The Voyager 3D viewer fully replaced our legacy 3D model viewers and has been enhanced with beta support for AR. The Packrat project, which will create a system to manage Smithsonian 3D digital assets, saw a refresh in 2020. With the new lead developer on board, the team set about architecting the system with a focus on long-term sustainability of the Smithsonian's 3D assets.
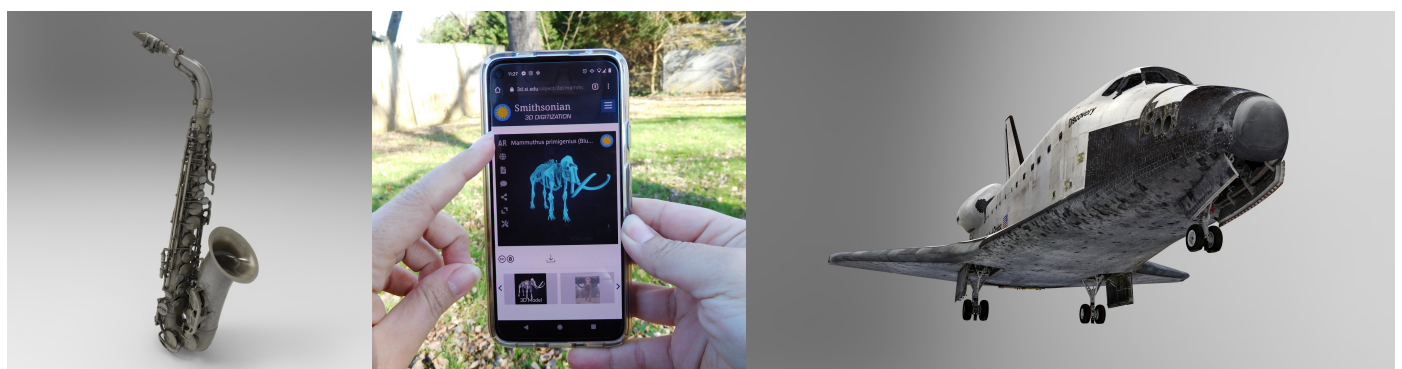
2020 started with the continuation of our programmatic digitization efforts initiated in 2019, which were based on museum prioritizations documented in the Unit Digitization Plans (UDPs). As always, we appreciate the efforts of our sister DPO program, Policy and Analysis, in supporting each museum's development and management of their UDPs. To that end, the beginning of the year saw us continue and complete projects with the Smithsonian Gardens' Druse Collection, the National Museum of African Art's Egyptian Postcard Collection, the National Museum of American History's Numismatics coin collection, and three projects at the National Museum of Natural History in the Departments of Botany, Paleobiology, and Entomology. All told, these projects contributed to the creation of half a million additional digitized objects and specimens across the Institution.

The COVID-19 pandemic made this year particularly challenging since the Mass Digitization program, which relies on access to collections, could not conduct operations as planned. In response to the COVID challenge, the DPO Mass Digitization program executed a strategic pivot to collections records digitization, a move that is on par with the scale and importance of object imaging in addressing critical Institutional needs.

To that end, the Mass Digitization Program accelerated the growth of its Informatics Program to meet the challenge of digitizing the Institution's vast amount of collections-based analog data (card catalogs, registers, accession records, etc). Using our well-refined project processes and workflows, we initiated a series of collections records digitization pilot projects across the Institution to include NMNH's Paleobiology Department, NMAH's collections registers, CHSDM's card catalog, among others. With the various pilots' successful planning and implementation well underway by year's end, we expect to see great results from these initiatives as we transition from smaller pilot projects to large scale production projects in 2021.

In addition to ramping up collections records digitization, the Mass Digitization Informatics program completed two important projects. The first was a Mass Georeferencing pilot project, conducted in collaboration with two NMNH departments, that focused on enhancing and improving the ease of use of georeferenced data enrichment. The second effort was an OCR/Machine Learning project which gave us the blueprint for planning transcription projects that use iterative passes of OCR, Machine Learning, and human transcription to get the best results from each resource while leveraging cost savings from automated processes.

Despite the significant slowdown the pandemic has caused us all, the Mass Digitization Program's work to date has brought the total number of digitized Smithsonian collections to 4.5 million since the program's inception in 2014.
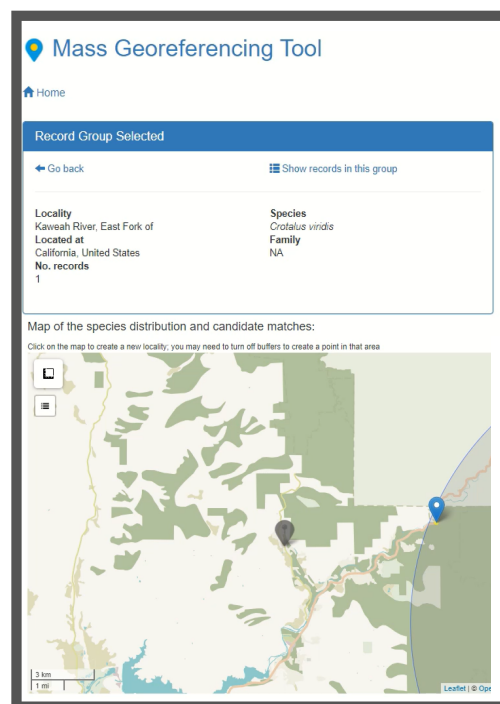
Once again, the Smithsonian's SD 600 collecting units responded to the annual call for metrics to represent the physical and digital state of the Smithsonian's collections. In FY2019, the museum, archive, and library units together created or enhanced a total of 851,000 standard digital descriptions (+8% over FY2018) and 888,000 standard digital surrogates (+9% over FY2018). These describe and represent our 155 million museum object/specimen collections, 163,000 cubic feet of archives and 2.2 million library volumes.

To learn more, **visit the Smithsonian Dashboard**.

The Smithsonian has a number of systems and processes in place that store and maintain our digital assets (such as documents, still or moving images, sound recordings, presentations, 3D scans or data generated from scientific instruments as well as digital descriptions or metadata) for the long-term. In order to ensure that all digital assets are properly cared for, in the summer of 2020 DPO issued a survey to 26 units and research centers to identify any digital assets that may be at risk of loss because they are not currently stored and managed in a Smithsonian collections information system like TMS, ArchivesSpace, DAMS, or another "system of record". Gathering this information will help us to locate and quantify potentially at-risk digital assets so steps can be taken to safeguard them. Survey results and a final report are expected in 2021.
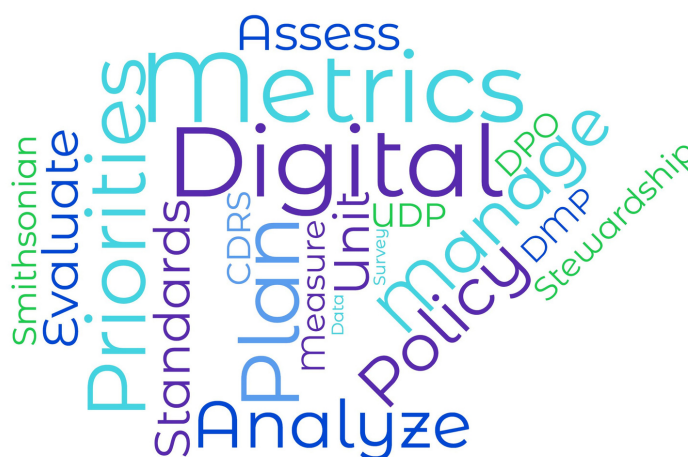
Special Initiative - Audio Transcription/ Captioning Pilot

In 2020 the DPO began a pilot project with SI Archives (now in SLA) on over 80 digital recordings to do a side-by-side comparison of four different methods for transcribing and captioning digital audio recordings. The methods we are comparing are: 1) Human Transcription & Captioning with Human Verification; 2) Automated Transcription & Captioning with Human Verification; 3) SI Transcription Center's TC Sound crowdsourcing platform and 4) DIY method using an online tool such as YouTube Studio. Variables we are comparing include: Cost, Accuracy, Quality, SI Staff Time, File Delivery, Professionalism/Satisfaction with Vendor/Method. The results of this pilot are expected in 2021 and will help inform efforts underway to make Smithsonian collections more accessible to all of our visitors

| **33M**	**127K**	**1.5M**
Objects and Specimens with Digital Record	Archival Cubic Feet with Digital Record	Library Volumes with Digital Record
5.6M	**26.0K**	**59.6K**
Objects and Specimens with Digital Image	Archival Cubic Feet with Digital Image	Library Volumes with Digital Image

With support from OCIO's F&A, OA, and OGC, OCIO Development assists DPO with more than just gifts: collaborations, partnerships, content on 3rd party platforms, business ventures, and generally all interactions with third parties except contracting and purely programmatic relationships. The launch of Smithsonian Open Access and the COVID-19 pandemic dominated development work at the DPO in FY2020, as they did everywhere else at the Smithsonian. The blog post **"How to Build Partnerships for Museum Digitization in 2020"** shares DPO Development's lessons learned with a wider audience. Although the blog was written pre-pandemic, its lessons still hold true in the New Normal.

How to Build Partnerships for Museum Digitization in 2020

By Jaap Otte on Fri, 02/07/2020 - 17:07

Smithsonian Open Access proved to be a big draw for corporations and foundations who recognized the transformative nature of making millions of collections images as well as 3D models available without restrictions and were eager to partner with us in demonstrating various uses of the Open Access data. The effort described below is the result of a close collaboration between DPO and the larger OCIO, Open Access, OA, OGC, OCON, and OPA. Although launch collaborations were not structured as gifts to the Smithsonian, several existing donors and partners eagerly participated, helping us to successfully launch and deepening the relationship. Building on DPO's network in the tech and artistic communities, we were able to contribute about half of the collaborators for the launch event on February 28, 2020. The launch also resulted in several new relationships with corporations large and small. In addition to launch demos, Google and AWS also provided major support towards Open Access and the long-term cloud storage costs of Open Access data. Attracted by the Smithsonian's new Open Access content and the Smithsonian brand, we were able to start discussions after the launch with several large companies about the creation of digital experiences, major donations, and business opportunities. We expect these to come to fruition in FY2021.

We used the first few weeks and months after the museums closed for COVID-19 to connect with our donors and partners on the human level and, subsequently, to assess how discussions about donations and collaborations had been impacted and might be able to proceed under the new circumstances. FARO, a major partner of the 3D Program since 2012, and Materialise, a 3D software provider, generously agreed to renew key multi-year donations. The Smithsonian's and the world's focus on digital content as a result of both COVID-19 and Smithsonian Open Access resulted in new opportunities and provided additional impetus to ongoing discussions with donors and partners. COVID-19 also helps us to increasingly be One Smithsonian when it comes to working with donors and partners.
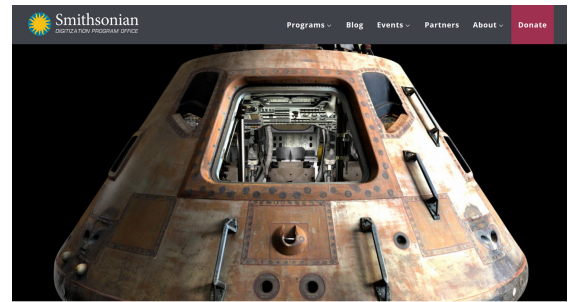
Jamie Cope, 3D Program Lead Developer

Jamie Cope is a computer engineer by trade and entrepreneur at heart with more than 15 years of experience employing skills in usability, data visualization, AR/VR, and simulation across diverse fields such as medicine, biology, education and training, and entertainment. He has developed 3D data pipelines for large-scale model sets and several custom systems for cleaning, storing, and analyzing user-generated data. He is passionate about user-centered design to solve real-world impactful problems. Jamie is the lead developer for the Smithsonian's Digitization Program Office 3D Program.

Jon Tyson, 3D Program Software Architect for Packrat

Jon Tyson is a seasoned IT professional with 30 years of experience. He has worked in every hands-on engineering role including junior engineer, senior engineer, team lead, software architect, and chief engineer. He has more than 17 years of diverse management experiences, including time spent as a front-line engineering manager, engineering director, CTO, and CEO. Jon spent 9 years at Autodesk, working as an engineer and tech lead on graphics and then as an engineering director for an online collaboration platform for the management of complex construction projects.As CEO of 2 online retail companies, Jon also filled the roles of product and project manager for the suite of software solutions employed by these companies; he translated domain-focused business problems into engineering specifications and then directed the work of a team of engineers to build successful solutions.Jon lives in San Francisco, has two children aged 21 and 18, and also performs as a professional musician in a number of musical projects.

Keturah Kiehl, Policy & Analysis Support

Keturah joined the DPO part-time to assist with the functions of the Policy & Analysis program. When Keturah is not working at the DPO, she is working on her PhD in Classics at Catholic University or participating in historical reenactments of various battles from American history.

New Opportunities

In 2020 we said farewell to Kathy Adams, long-time project management consultant (and unofficial cat herder) for DPO's 3D and Policy and Analysis programs. Kathy and her husband are moving to the warmer climes and more open spaces of South Carolina. We wish her well and thank her for her help and innumerable kindnesses.

Digitization Program Office Advisory Committee (DPAC)

Members advise the DPO about its program and plans and serve as conduits of information from DPO to their units. Members serve three-year terms and are drawn from across the Smithsonian, representing all units and every type of profession that exists across the Institution.
https://sinet.sharepoint.com/sites/DPO/Sitepages/DPAC.aspx

In 2020, we said goodbye and thanked the following members who completed their terms:
- Effie Kapsalis, Senior Program Officer for Digital Strategy, American Women's History Initiative | Smithsonian Open Access
- Beth Stern, Director, OCIO Office of Research Computing

We also welcomed the following new members, who will serve until 2023:
- Alise Fisher, Science Press Secretary, OPA
- Dave Walker, Audiovisual Archivist, CFCH

DPAC Subcommittees:

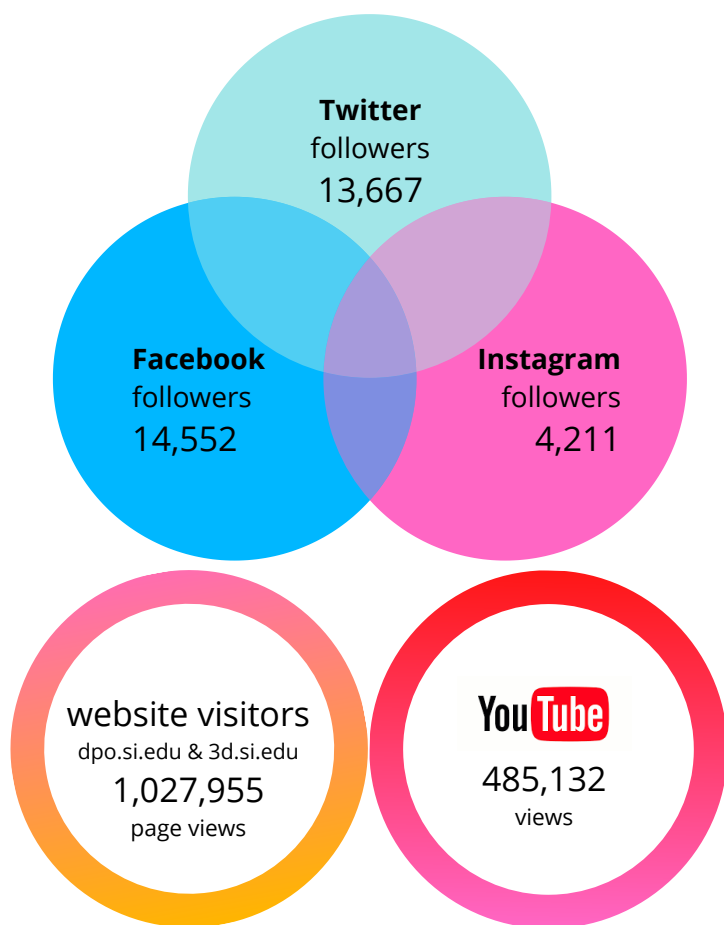**3D Digitization Subcommittee** (Vince Rossi, Chair)
The subcommittee met intermittently during the year to review demos and updates to the Packrat asset management system, a critical component of the 3D pipeline under development by the 3D team.

Informatics Subcommittee (Luis Villanueva, Chair)
The committee met on several occasions to preview and comment on the Mass Digitization Program's Mass Georeferencing and OCR with Machine Learning projects.

DPO staff served on several internal Smithsonian Committees including:

- Collections Information Management Committee (CIMC)
- Smithsonian Institution Archives & Special Collections Council (SIASC)
- Audio Visual Archivists Institutional Leadership (AVAIL)
- Audio Visual Mass Preservation Initiative (AVMPI)
- Open Access Infrastructure Team
- Open Access Advisory Committee
- Open Access Partnerships Team
- Collections Information Systems (CIS) Pool Fund Review Committee
- Pool Fund Allocation Committee
- NMNH Informatics Task Force
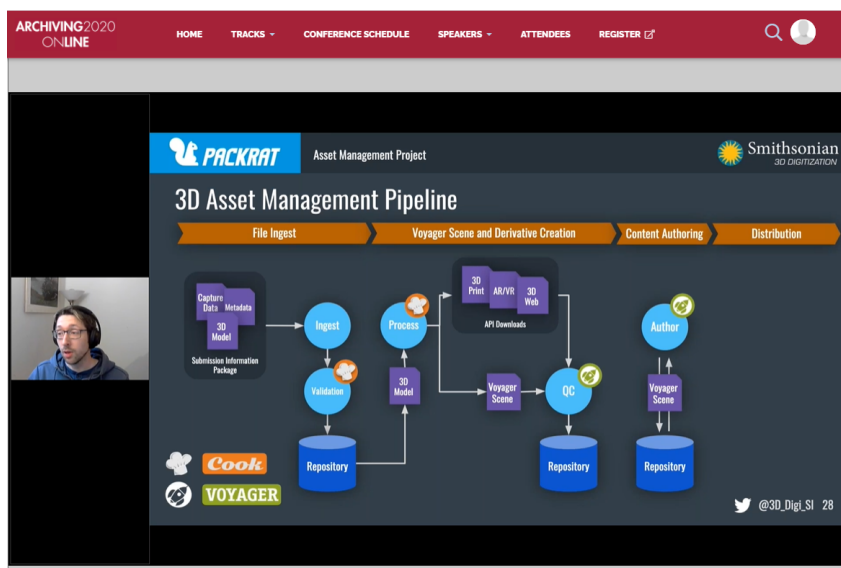- DAMS Quarterly Users Group Meeting
- GUID Working group

Twitter
followers
13,667

Facebook
followers
14,552

Instagram
followers
4,211

website visitors
dpo.si.edu & 3d.si.edu
1,027,955
page views

You Tube
485,132
views

Top Tweets

Smithsonian 3D
@Smithsonian3D

We are excited to announce the test launch of new #3D augmented reality #AR experiences! Our museums may be temporarily closed, but you can visit 10 beta AR experiences from your mobile device. dpo.si.edu/blog/voyager-ar #musetech @NMNH

GIF

4:01 PM · Nov 24, 2020

263   111 people are Tweeting about this

Top twitter by likes (261 likes, 142 RTs, 99,486 impressions):

SI Mass Digitization
@SIxDIGI

The @SIxDIGI program of @SmithsonianDPO has helped the Smithsonian units digitize 4.2M specimens and objects with 3.6M images since 2013!
Take a look at the Mass Digitization Dashboard we are now sharing with the public:
shiny.si.edu/massdigi/

2:13 PM · May 12, 2020

36   20 people are Tweeting about this

OUTREACH

DPO staff presented their work at a dozen virtual conferences, workshops, and classroom venues during the year. Some of these opportunities were Smithsonian-based, but many were professional conferences in both the cultural and technology sections, such as the Museum Computer Network annual conference, Tomography for Scientific Advancement (ToScA), and Imaging Science and Technology Archiving conference. Staff also lectured in classes and public programs at Georgetown University, George Washington University, Harvard, Bowdoin College, and University of Florida.

Jon Blundell speaking at the Archiving Online 2020 virtual conference on the 3D asset management pipeline 'PackRat'

DPO's work was showcased in over three dozen blog posts and print/online news media such as hyperallergic, Engadget, SlashGear, smithsonianmag.com, adweek, Forbes, PBS.org, yahoo!finance, PLOSciComm, and Nature. Below is a partial sampling of some high profile news stories.

03/03/20: **Forbes** Magazine: Editor's Pick - Smithsonian Releases NASA Space Shuttle Discovery 3D Models You Can Download Free

12/09/20: **Engadget**: Instagram adds AR exhibits from the Smithsonian to its camera

12/09/20: **tech@Facebook**: Home is where the ARtifacts are

02/26/20: **slashgear.com**: Smithsonian Open Access initiative releases loads of 3D models to public domain

03/16/20: **PBS News story**: 19 immersive museum exhibits you can visit from your couch

02/27/20: **Hyperallergic**: 1,700 3D Models of Cultural Heritage Objects Now Available in Public Domain

12/23/20: **PLOS SciComm**: Top 9 Discoveries in Human Evolution, 2020 Edition

05/14/20: **Published Paper:** NATURE - Snapshots of human anatomy, locomotion, and behavior from Late Pleistocene footprints at Engare Sero, Tanzania

10/08/20: **Smithsonian Magazine**: How an Expedition to the Galápagos Islands Saved One of the World's Largest Natural History Museums

Coral Collection

DPO's 3D team and NMNH, in collaboration with The Hydrous, unveiled a new project aimed at igniting discovery through digitization. Anyone with internet connection can access NMNH's Coral Collection to discover and explore beautiful and fascinating coral ecosystems at **3d.si.edu/corals**.

Coral Collection features over one hundred high resolution 3D scans of specimens, including corals and reef dwelling species, curated through an interactive and educational digital platform. The Hydrous is a 501(c)3 non-profit organization whose mission is to create 'open access oceans,' and their team of marine scientists created rich educational content for this digital platform that is divided into five learning modules: Ecosystem Engineers, Classifying Corals, Shape & Structure, Depth & Distribution, and Coral to Community.

All of the 3D models in this project are freely available for use via Smithsonian Open Access, and are downloadable by selecting "Voyager Link: View model" on each specimen.

| Original Name	Synonomized	Catalog number	Module	Link to model
Acropora cervicornis		USNM 1171477	Classifying Corals	View model »
Acropora cervicornis		USNM 74016	Depth & Distribution	View model »
Acropora digitifera		USNM 48086	Shape & Structure	View model »
Acropora palmata		USNM 3968	Depth & Distribution	View model »
Acropora prolifera		USNM 1149322	Depth & Distribution	View model »
Acropora secale		USNM 268	From Coral to Community	View model »
Acropora valenciennesi		USNM 47056	Shape & Structure	View model »
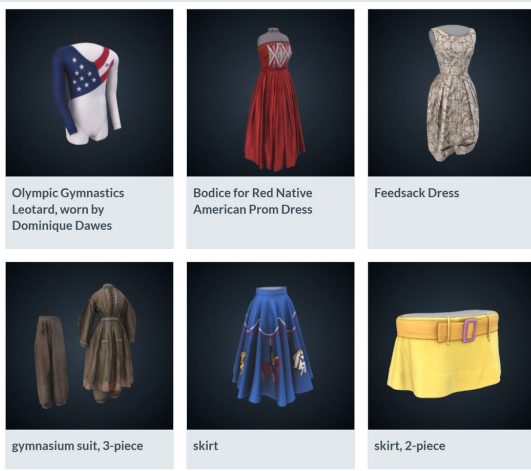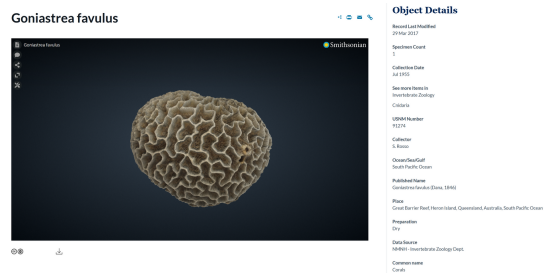
Costumes from Girlhood (It's complicated)

In celebration of the opening of "Girlhood (It's complicated)" at the National Museum of American History, the DPO 3D team launched a collection of 21 3D models of garments from the exhibit, along with interactive educational content. These garments hold rich autobiographical stories, and can now be explored in detail (down to the stitching!) while you read about the individuals who wore these items and ignited social change.
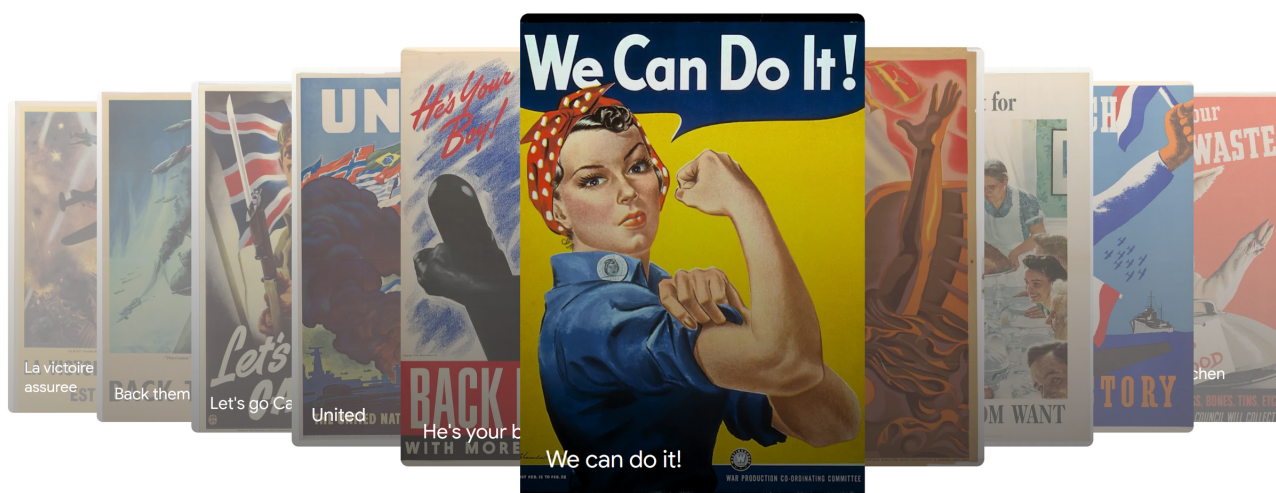
Instagram AR Stories

The 3D Team collaborated with five Smithsonian museums and Facebook to create augmented reality (AR) effects for ten iconic objects.

Instagram users can now bring these objects into their backyard or living room, learn facts about them, and share photos and videos with their friends through Instagram Stories, Reels, and in their Instagram feed. This personal and immersive experience also lets users view the objects close-up and at angles not available in our museums.
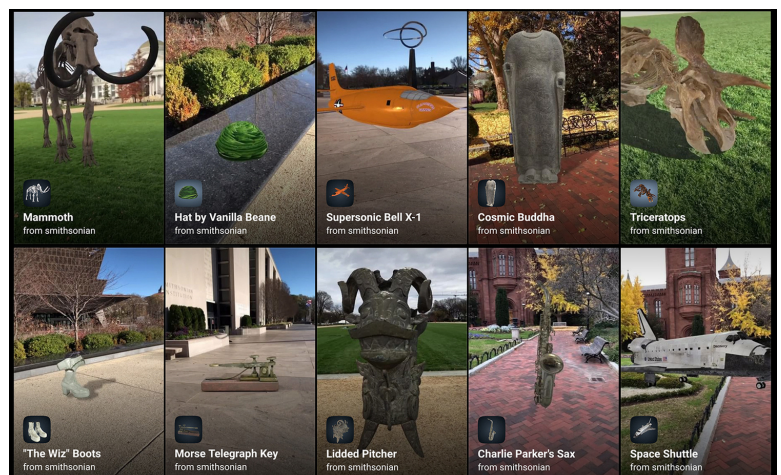
- Library of Congress and AfricanFossils.org adopted the Smithsonian open source 3D Voyager viewer https://gcn.com/articles/2020/02/21/loc-3d-printing.aspx & https://africanfossils.org/search

- The NMAH Poster Collection
  This collection of WWI and WWII posters, which was digitized last year, was recently launched on the Google Arts and Culture platform. This project was a collaboration between the Mass Digitization team, the National Museum of American History, and Google Arts and Culture.
  https://artsandculture.google.com/partner/smithsonian-national-museum-of-american-history
  WWII collaboration with other institutions: https://artsandculture.google.com/project/smithsonian-world-war-ii/art

AWARDS

- The 3D open source Voyager viewer was a finalist in the 2020 GLAMi awards' category for Best Exhibition/Collections Extension. https://mw20.museweb.net/glami-award-finalists/. (First place went to the Smithsonian's Open Access Initiative, making for an all-Smithsonian sweep in this category!)

- Luis J. Villanueva wins his second Ebbe Nielsen Challenge Award
  The Global Biodiversity Information Facility (GBIF) awarded a Third Place prize in their 2020 Ebbe challenge to Mass Digitization Informatics Program Officer Luis J. Villanueva for his Mass Georeferencing Tool.

  The Ebbe Nielsen Challenge is a competition for projects that "leverage biodiversity data and tools from the GBIF network to advance open science." This is the second Ebbe Nielsen Challenge award for Luis, who received second place in 2018 for a tool he created that helps units identify and correct issues in digital records.
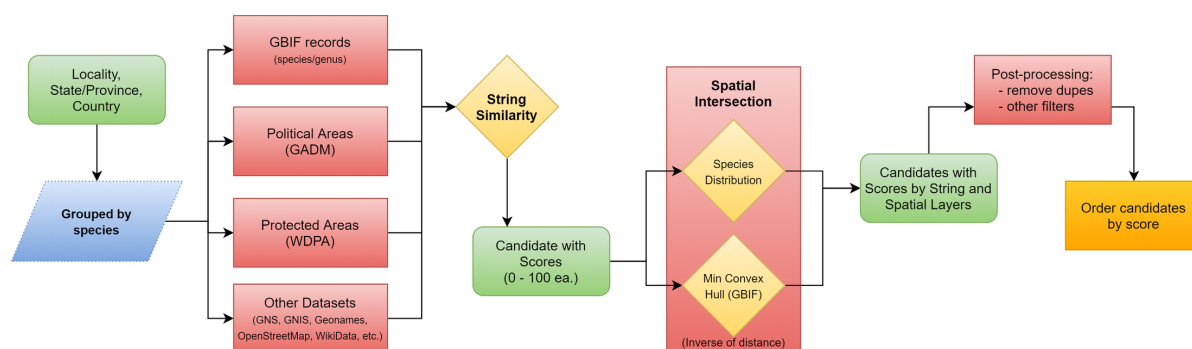
DPO is an office in the Smithsonian's Office of the Chief Information Officer. Our colleagues in OCIO assisted us in unheralded ways throughout 2020, as they do every year. We are particularly grateful to the following divisions and branches: Network and Voice Services, Network Security, Desktop Services, DAMS Support, Help Desk, Research Computing, System Architecture and Product Assurance, Collections System Support, the Web team, and Libraries and Archives System Support. Special thanks to the OCIO Finance & Administration team, who ease our way through Federal policies and procedures on a daily basis.

Digitization Program Office Staff

Diane Zorich
Director, Digitization Program Office

Nathan Ian Anderson
Mass Digitization Program Officer

Jon Blundell
3D Program Officer

Jamie Cope
3D Program Lead Developer

Megan Dattoria
3D Program Officer

Keturah Kiehl
Policy & Analysis Support

Jeanine Nault
Mass Digitization Program Officer

Jaap Otte
Director of Development, OCIO

Anaís Perez
3D Specialist

Ken Rahaim
Supervisor, Mass Digitization Program

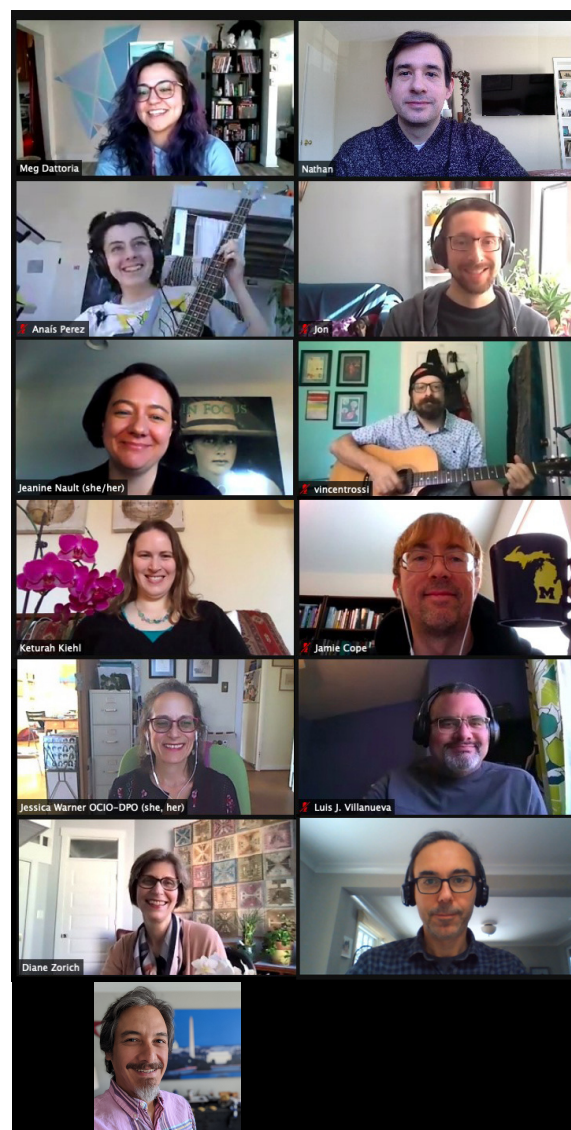Vincent Rossi
Supervisor, 3D Program

Jon Tyson
3D Program Software Architect for Packrat

Luis J. Villanueva
Mass Digitization Informatics Program Officer

Jessica Warner
Senior Policy & Analysis Program Officer

Additional Image Credits

Cover: upper right image of a 3D rendering of 'Bodice for Red Native American Prom Dress' courtesy National Museum of American History; lower right images of Xylocopa bees courtesy National Museum of Natural History.
Page 3: upper images from left to right; 3D rendering of coral collection courtesy National Museum of Natural History, 3D rendering of gymnasium suit from the 1800s courtesy National Museum of American History, 3D rendering of skull specimens courtesy National Museum of Natural History. Lower images from left to right; 3D rendering of alto saxophone owned and played by Charlie Parker courtesy National Museum of African American History and Culture, wooly mammoth skeleton in augmented reality on a cellphone courtesy National Museum of Natural History, 3D rendering of Space Shuttle Discovery courtesy National Air and Space Museum
Page 4: images descending in order from right; NMNH Entomology digitization; production image of bee specimen xylocopa from the NMNH Entomology Dept.; Luis Villanueva's mass georeferencing tool screenshots; production image of fossilized marine invertebrates from the NMNH Paleobiology Dept.

Annual Report Design: Nathan Ian Anderson.