Digital Rocks Portal (Digital Porous Media): connecting data, simulation and community

Maša Prodanović[1,*], *Maria* Esteva[2], *James* McClure[3], *Bernard* C. Chang[1], *Javier* E. Santos[4], *Anuradha* Radhakrishnan[1], *Ankita* Singh[5], and *Hasan* Khan[6]

[1]Hildebrand Department of Petroleum and Geosystems Engineering, The University of Texas at Austin, Austin, TX, USA
[2]Texas Advanced Computing Center, The University of Texas at Austin, Austin, TX, USA
[3]Advanced Research Computing, Virginia Tech, VT, USA
[4]Center for Nonlinear Studies & Earth and Environmental Sciences, Los Alamos National Laboratory, Los Alamos, NM, USA
[5]University of New South Wales alumna, Sydney, Australia
[6]Petroleum Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia

Abstract. Digital Rocks Portal (DRP, https://www.digitalrocksportal.org) organizes and preserves imaged datasets and experimental measurements of porous materials in subsurface, and beyond, with the mission to connect them to simulation and analysis, as well as educate the research community. We have over 150 projects represented in more than 200 publications, and an active community that reuses the data, most recently in multiple machine learning applications for automating image analysis as well as the prediction of transport. Such automation is at the core of ability to many formation evaluation tasks in near-real time. We present benchmark datasets that have played a role in recent machine learning prediction successes in the field. We further discuss the vision for further research advances, educational materials, as well as growth and sustainability plan of this digital rock physics community resource. In particular, we are in the process of expanding into a broader repository of engineered porous materials, specifically those for energy storage and the portal will transition to Digital Porous Media (DPM).

1 Introduction

The field of digital rock physics combines imaged porous media (such as rocks and soil) to simulations and other research methods such as machine learning (ML) and data analysis to understand transport and deformation in subsurface. Applications include formation evaluation, management of groundwater resources, carbon sequestration, enhanced oil recovery, and contaminant transport [1]. Typical geological systems are composed of a broad spectrum of porous media with properties such as permeability varying by orders of magnitude within an individual system. The advances in high-resolution imaging techniques (x-ray tomography , scattered electron and optical microscopy) have provided a wealth of 2D and 3D datasets that reveal the microstructure of rocks, soil or model media on scales ranging from nanometers to centimeters, alas analysis and use of such images in simulation is computationally demanding [2]. Our practical questions nevertheless require upscaling to the field scale. Therefore, efficient connection of curated digital rocks data and scalable computing and machine learning (ML) applications is necessary for applications [3].

We here describe the organization of Digital Rocks Portal (DRP), a well-adopted portal for curation and visualization of large imaged datasets of geomaterials. DRP's mission is to organize and preserve imaged datasets and experimental measurements of porous materials in subsurface, connect them to research methods We report on our efforts in educating the research community by organizing workshops and mini-courses in data preservation, analysis and visualization.

Deep learning methods and convolutional neural networks (CNNs) have many applications in image processing and flow prediction [4]. For this, reliable "ground truth" benchmark datasets are needed because (1) supervised deep learning methods require a large amount of validated data to train models; and (2) the capabilities of the trained classifiers must be assessed quantitatively. Applying learning algorithms to the large volume of data stored within DRP is currently rendering scientific advances. We successfully applied deep learning algorithms on predicting 3D velocity fields in porous media images [5,6]. The implementation was done in TensorFlow [7] using high performance resources (HPC) at Texas Advanced Computing Center (TACC). The work in [5] sampled a wide variety of images from DRP as well as adding a benchmark collection [8], has specifically enabled deep learning training with geologic porous media whose heterogeneity (including fractures/cracks) likely surpasses any engineered porous materials, and, should data from energy storage materials become available, the results should immediately transfer. This exemplifies a workflow that can enable future real time permeability estimation based on imaged datasets, as the computational time of a deep learning prediction is in seconds compared to hours on HPC.
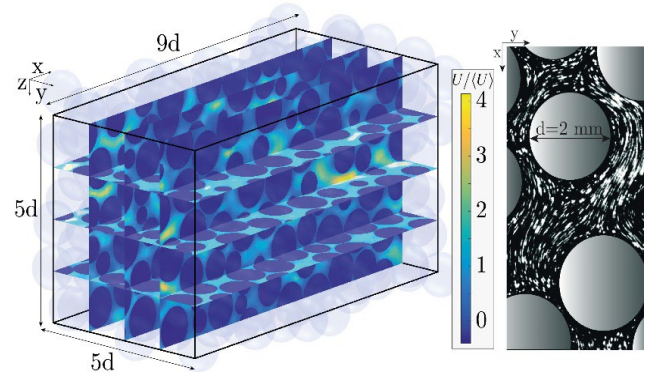
Visualization algorithms of cross-sections, volumes and surfaces of 3D data (e.g. marching cubes algorithm) are well-
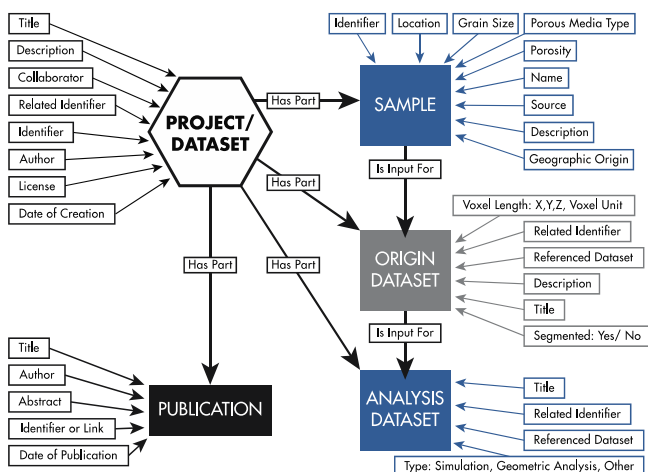
---

* Corresponding author: masha@utexas.edu

known and available in a number of advanced visualizing platforms (e.g. ParaView, MayaVi, ImageJ, Dragonfly, to name a few that are open source). However, the complexity of the pore structure and processes captured within pose a challenge in visualizing the data, both from memory and computational perspective. That said, while this challenge is meant to reuse data from geosciences and subsurface engineering, all of the concepts that it is meant to educate students about are very much applicable to any scientific visualization endeavor: for instance, in biomedical field, heart or blood vessels in a human body, are often imaged using computed tomography (or simulated based on those images) and have particulate flow (blood flow, that is) within. In addition to the materials from the recent visualization contest reusing DRP datasets, we provide Jupyter Notebooks examples visualizing a 3D velocity field as well as porous media surfaces in multi-fluid situation.

Transitioning to renewable energy from the hydrocarbon-based sources that currently meet 65% of energy demands will require significant engineering effort on several fronts: (1) mitigating the effects of air pollution with solutions such as subsurface carbon dioxide sequestration; and (2) scaling up renewable energy based on the maturation of advanced energy devices, geothermal, and other solutions. Complex microstructure is a common element for many of these physical problems. We discuss our sustainability plan that includes the paid membership fees scheme for Digital Rocks Portal, thus ensuring its financial stability, (2) strengthening our visualization and analysis services and (3) generalizing and broadening the user base of DRP to Digital Porous Media (DPM) portal to energy storage materials. Geometric approaches developed to enhance digital rock images have excellent crossover potential to other complex material structures [9]. High performance methods to enhance experimental data collection workflows are also broadly generalizable [10].

for porous media as opposed to platforms such as Mendeley Data [12], Dryad [13] or Energy Data Exchange [14] that store data of any origin. The portal is specifically equipped with tools tailored to organize, visualize, analyze, and publish digital images of porous materials, Its data model shown in Fig.1. is the foundation for curating, describing and displaying the datasets. The data model enables imaged data curation workflows including adding provenance information and metadata used for rendering the images, and organizing the dataset in relation to its corresponding research activity such as observations, experiments, and simulations. Thus, the model is the basis for the datasets representations on the landing pages that displays provenance as material entities (samples) and processes (experimental or simulation) from which data derives. This tree like representation, which was evaluated through user studies [15], facilitates understanding of and access to complex and large datasets [16]. Example projects are shown in Fig.2. and Fig.3.
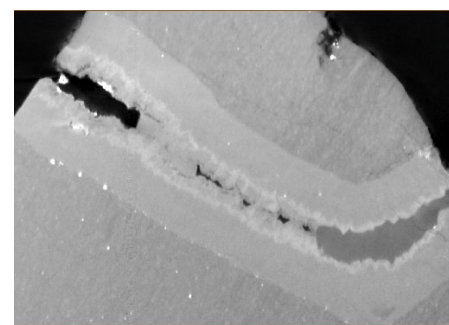

Fig.2. Example DRP project 3D experimental fluid velocity field within a sphere packing (model granular or soil structure) obtained particle velocimetry [17]. To our knowledge, this is the only open 3D velocimetry dataset, and has been the bases for the winning visualization entry in DRP visualization contest (see below).

Fig.3. Example DRP project [18] with 3D x-ray microtomography images as well as XRF mineral maps of a partially mineralized vein contained from a core sample (depth of ~10,400 ft) in the Upper Wolfcamp formation. The sample was studied as part of the effort of the Department of Energy's Center for Nanoscale Controls on Geologic CO_2.

DRP does not assume a specific image format at this point. This is because subsurface porous media heterogeneity requires a variety of imaging modalities ranging from microscopy to x-ray microtomography to medical computed tomography with no commonly accepted data formats. As a result, DRP has built-in capabilities for translating and visualizing raw data after import, which distinguishes it from other imaged data repositories.

2 Portal organization

Launched in September 2015 the DRP [11] is an established data repository, trusted by the community. DRP is specialized

Fig.1. DRP Data model is foundation for the database linking web content and files on HPC storage.

Data curation is assisted through the publication pipeline interface that carefully guides users through the process. Metadata collected during this process is essential for full documentation of the dataset. Images by themselves contain structural information collected at a certain time, but not more than that and this step is essential. Fig.4. and Fig.5. exemplify this information for an example project. Note that only part of the information collected is shown for brevity. Prior to publication and before assigning a digital object identifier (DOI), the datasets are verified for completeness and accuracy by a live curator, again unlike many data portals that allow upload of compressed archives. Current list of curators includes the majority of the authors of this publication and we are open to volunteers. DOIs are provided through The University of Texas Libraries subscription to DataCite Fabrica [19].

Fig.4. Part of the metadata collected for an example dataset [20] including voxel length that relates the image to sample physical size.

Fig.5. Second part showing successful "ingestion" of two datasets of the same project as shown in Fig.4. ([20]) that are part of of the same experiment (as proved by the image cross-sections. Action button allows download, shows a video of all slices (in 3D), and

basic image statistics such as histogram. "Action" button also has a button that allows connection to HPC cluster on TACC that opens ParaView for visualization. This software stack is however currently being updated.

DRP is implemented within the reliable HPC environment deployed and maintained by Texas Advanced Computing Center, and data storage is supported by The University of Texas System Research Cyberinfrastructure (UTRC) in the storage resource Corral. This infrastructure assures data security and persistence, interoperability with different HPC systems for visualization and analysis (TACC Analysis Portal, [21] currently on Lonestar6 and Stampede2 clusters) and continuing operation of the software stack, and thus it contributes to the portal's sustainability.

3 Portal impact and user engagement

3.1. Metrics

Since 2015, DRP has published 154 projects and 200 related publications [22], ranging in size from 10MB to 728GB. DRP accepts international contributions. Our current query country of the institutional affiliation of the first author of the current DRP projects we find 48% from the U.S., 18.2% from UK, 9.5% from Australia, and the reminder from China, Belgium, Brazil, Canada, South Korea, Japan, Russia and Switzerland.

Since its launch, we have been keen in promoting DRP and in understanding its impact. We implemented Search Engine Optimization (SEO) strategies such as using schema.org vocabularies and Google-friendly metadata in the repository landing page. Using Google Analytics, we monitor an increase in total visits over the years as well as consistent spikes in usage, which coincide with academic events that we undertake (workshops, presentations, publications). We also implemented a data download count (of individual datasets) to each of our dataset pages: it is available under link "Usage Information" embedded in each project landing webpage. Through Google Scholar Alerts we follow mentions and citations to the repository and to individual datasets.

3.2 Digital Object Identifiers (DOIs) and increasing visibility of data

Journal publications have had DOIs for a long time, but providing them for research data or code is a relatively new adoption thanks to efforts such as Findable, Accessible, Interoperability, and Reusable (FAIR) data efforts launched in 2016 [23]. We have found that authors have variable or incomplete referencing styles for datasets in papers, whether they are referencing their own data or reusing someone else's. Many, for instance, include only the name of the platform where the data is posted. Note that including the dataset DOI and a proper reference improves the findability, as **search engines presently do not find data as efficiently as they do papers (if at all).** One reason for this is that datasets come in variable formats and are not easily searchable for information the way .pdf or text formats are. We thus next describe strategies that readers can use to improve other authors finding their data.

Increasingly, journals request that you publish your paper along with the data that supports your findings. For example, the AGU journals follow this FAIR data practice. Curating data for publication demands significant efforts on the part of

the researchers that have to organize and describe it so that others can recreate their results or reuse the data for a different purpose. At the same time, data publications can bring significant rewards in terms of citations and exposure of your work. To help you boost the discoverability and consequently the reuse of your data there are some basic strategies that you can follow:

To enhance discoverability:

- Add all the information (metadata) that DRP asks for, including your ORCID [24] ID (that uniquely identifies you in case you change affiliation or someone else shares your name) if you have it, and who funded your project.
- In your data description use terms that you consider people will search for when looking for this type of data. Highlight the unique characteristics of your dataset and suggest how your data could be useful to others and for what purposes.
- All these metadata are indexed by search engines such as Google, improving the possibilities that your data is discovered by broader audiences.

To facilitate data reuse:

- Add documentation in the form of readme files or reports to clearly explain how you obtained and processed the data. This stimulates the understandability of your publication and consequent data reuse.
- You can include code snippets as "Non-image data" to your datasets, or better yet, publish the code on GitHub and add it as a related publication in DRP.
- If you have them prior to publishing the data, add papers and proceedings that discuss the data as related publications. After your dataset is published you can request DRP curators to add your new publications by sending an email to masha@utexas.edu.

3.3 Newsletters, project highlights and educational material

Engaging the research community is necessary if the data will become reused and thus actively contribute to science and engineering. This can be done through (classical) conference presentations, but also contests, newsletters and social media activities that promote data sharing, and reuse, data cyberinfrastructure, and overall data education through the published materials.
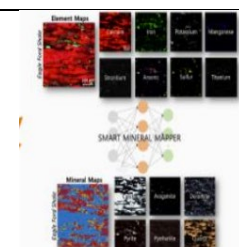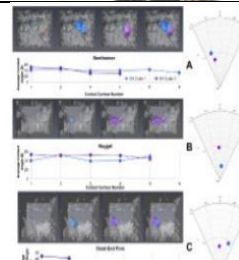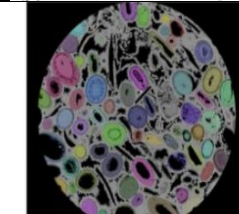
To that end, we launched newsletters in October 2020 and have since published them at the rate of two to three per academic semester. See [22] for links to all newsletters as well as links to a various news articles written about DRP. YouTube playlist on Digital Rocks Portal Visualization [25] contains various visualizations of the datasets, as well as a growing list of interviews we recorded to highlight projects on DRP.

Data reuse contests is another activity whose goal was to stimulate engagement and demonstrate the research potential of these datasets. organized a successful a virtual visualization mini-course, followed by a visualization challenge in early 2021 with three categories (video, static image and 3D printed porous microstructure). This is described in the next section.
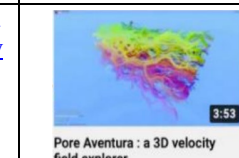
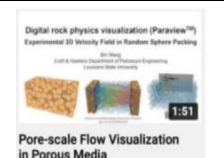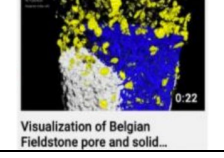### 3.4. Visualization mini course and contest

In October 2020, a porous media visualization competition was organized which consisted of a mini course (taught using Jupyter Notebooks), followed by a challenge with three categories with monetary awards to promote data reuse and create visualization templates for porous materials (known for their complexity). The task was to reuse any 3D dataset from the DRP to create a static image, video, or 3D printed visualization. Porous media are challenging to visualize due to complexity of pore/grain/fluid surfaces and interfaces, and this creates a resource for those who wish to learn advanced 3D visualization. Resources and materials were stored in DRP to help the participants. The events were sponsored by Southern Big Data Hub, Object Research Systems, Kitware and Dassault Systèmes. Tables 1 through 3 show the list of projects and the datasets that were reused from DRP.
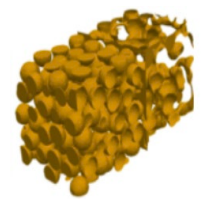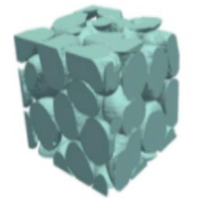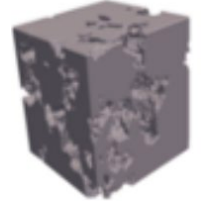
Table 1. Winners of the static image category

Title & winner	Dataset reused	Visualization
Smart mineral mapper, **Julie Kim,** Princeton University	Eagle Ford shale: synchrotron-based element and mineral maps	
Contact Angle Evolution, **Laura Dalton,** West Virginia University and National Energy Technology Laboratory	Bentheimer and nugget residual saturation micro-computed tomography data	
Identifying savonnieres carbonate ooliths, **Miral Tawfik**, Pennsylvania State University	Savonnières carbonate	

Table 2. Winners of the video category

Title & winner	Dataset reused	Visualization
Pore Aventura: 3D velocity field explorer, Dr. Mathieu Souzy, University of Twente	Experimental 3D velocity field in random sphere packing	

Flow visualization in porous media, **Bin Wang**, Louisiana State University	Experimental 3D velocity field in random sphere packing	Pore-scale Flow Visualization in Porous Media
Fluid displacement in a fracture, **Honggeun Jo**, The University of Texas at Austin	Fractures with variable wettability and roughness	Fluid displacement in a fracture
Belgian fieldstone pores and solids, **David Rippner**, University of California, Davis	Belgian fieldstone	Visualization of Belgian Fieldstone pore and solid...

Table 3. Winners of the 3D print category

Title & winner	Dataset reused	Visualization
Sphere pack juxtaposed with the non-wetting fluid at 40% saturation, **Javier Santos,** The University of Texas at Austin	Finney packing of spheres	
Visualization of sand/pore and a sphere packing surface, Yanbin Gong, University of Wyoming	Uniform quartz-Silver nanoparticle injection experiment and Finney packing of spheres	
Visualization of sandstone and carbonate pore space surface, Dr. Lingyun Kong, Energy and Environment Research Center	Doddington sandstone and Massangis Jaune carbonate	

4 Benchmark datasets

In order to push the state of the art for machine learning (ML) forward and provide comparisons between ML techniques, it would be very useful to have a large number of labeled data (the results of expensive full-physics simulations) to build models that work in the complexity of real-world pore structures. In other fields, large and diverse datasets [26,27] have sparked revolutions in ML by providing a common focal point that allows techniques to be compared against one another. Broadly speaking, this can be attributed to two factors. Firstly, unprecedented size and scale of data pushes the state of the art forward. Secondly, the public availability of these datasets makes easier the comparison of different modeling techniques, so that the relative advantages and disadvantages of different approaches can be more clearly

assessed. As an example, in porous media, the choice of fluid solver and numerical implementation thereof is a confounding variable with unknown effects on ML model characteristics.

We are working for DRP to become the a resource where benchmark tests that represent challenging, complex porous media in the context of image analysis, 3D simulation, empirical functional forms, and machine learning can be found. This diversity should include properties such as porous media lithology, boundary conditions, geometric resolution, and physical processes simulated. A large-scale collection of pore-scale data is critical for developing advanced ML algorithms that generalize to unseen samples from many real-life sources. Additionally, it (1) fosters advancement in our understanding of microscopic physics, (2) enables creation of new upscaling relationships, (3) provides a point of comparison to benchmark other simulators of physical processes, and (4) guides ML, which is still an up-and-coming field.

To that end, we have recently added a catalog of 3D geometries [8,28], as well as the companion collection of simulations [29,30], and structural properties of samples. The dataset includes 1736 flow and electrical simulation results on 217 digital samples, which required more than 500 core years of computation time. This data can be used for many purposes, such as constructing empirical models, validating new simulation codes, and developing machine learning algorithms that closely match the extensive purely-physical simulation. The dataset consists of simulations based on 217 digital rock samples, each containing two different domain sizes 256^3 and 480^3. For each digital sample, there are five nanoconfinement simulations, one fluid flow simulation, ten geometric features, electrical conductivity data, topological descriptors, and .png files visualizing cross-sections of the 3D data. The companion SCA paper (Chang et al.) discusses how the simulations are used for training neural networks for prediction of electrical properties.

There are also benchmark datasets posted by other authors in the area of image analysis algorithms. Three sandstone and carbonate collections [31–33] have been used for improving resolution of micro-tomography images, especially in the presence of microporosity. Further, there is an extensive dataset [34] for comparison of both traditional and ML image denoising algorithms.
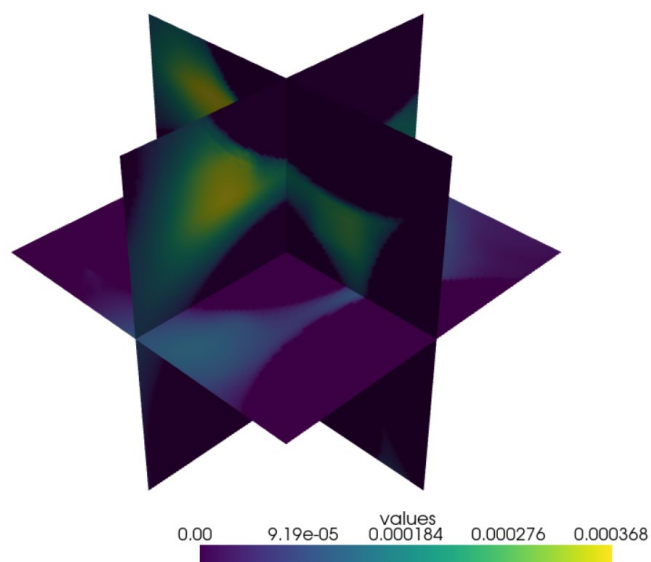
5 Visualization examples

We created two workflows for 3D visualization of surfaces and vector (velocity) fields and publish them in conjunction with this paper on GitHub of Dr. Prodanovic *(for the purpose of the review, they are available as a pdf alongside this manuscript)*. The workflows download and use the datasets from the portal.
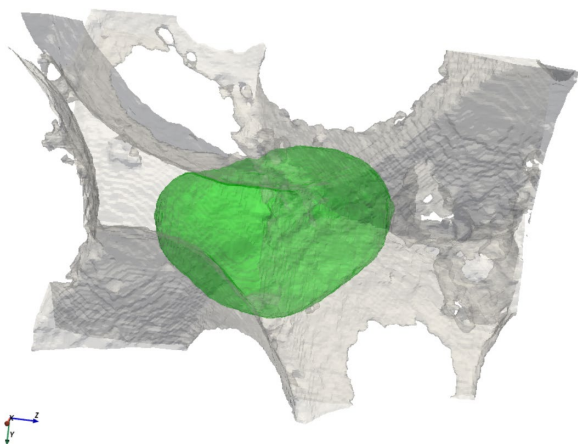
In these examples we relied on Python-based PyVista visualization software and provide both regular Python code (that is somewhat more robust, as it allows PyVista to open a separate window with the plot) as well as Jupyter Notebook. Jupyter Notebooks are excellent teaching tool and embed the results into web browser. That makes them versatile, but an extension is needed to make it work. The figures below show

a glimpse into what is available within the worksheets; they can be adjusted to any other data on DRP.

Notably, while these worksheets are currently set up to download the data and visualize it, **they also work on TACC Visualization and Analysis portal** [21]**, in particular on Lonestar6 HPC environment that can access DRP data storage directly without any download, and can analyze large datasets beyond the capability of an individual workstation.**

Fig. 6. Subset of the entire DRP velocimetry dataset from Fig.2 visualizing absolute values of velocities in PyVista.

Fig. 7. Surface visualization in PyVista of an oil blob (in transparent green) within a Ketton limestone sample (transparent gray). The high resolution x-ray microtomography experimental data [35] was used in the example.

6 Related software

The data in Digital Rocks Portal can be characterized and visualized using various data analysis software, and previous section provides one example. Beyond that, open source software such as ImageJ, ParaView, Visit or DragonFly is available, as well as commercial options (e.g. MATLAB, PerGeos).

Transport properties of porous materials, however, require simulation with different type of solvers. If segmented (binarized) images are available, then images can be direct input into simulation that produces a scalar or vector field of interest. The following table includes some open source example libraries.

Table 4. Example open source simulation software using DRP data.

Software	Application / method	Platform
LBPM [36]	Velocity fields, multiphase flow, lattice Boltzmann method	Unix, Parallel, C/C++ and MPI
Palabos [37]	Velocity fields, multiphase flow, lattice Boltzmann method	Unix, Parallel, C/C++ and MPI
MPLBM-UT [38]	Pre-processing and post-processing tools for single- and multi- phase flow with Palabos	Windows/Unix, Python
LSMPQS 1.0 [39]	Capillarity dominated fluid displacement, level set method	Unix, Serial C & Fortran + OpenMPI
NIST [40] Laplace Solver	Generalized Laplace Equation Solver (e.g. utilized for el. conductivity)	Unix, Parallel
PoresPy [41]	Image analysis and pore network extraction, many algorithms	Unix/Windows, Python, Serial
OpenPNM [42,43]	Pore network modeling (extract network with PoresPy)	Unix/Windows, Python, Serial

7 Sustainability model

Sustainability is understood as a combination of technical and administrative strategies to ensure that users can rely on ongoing data infrastructure services that promote research innovation. While we support open science and specifically open data, neither is unfortunately free and a business model has to ultimately be developed for sustainability. The portal is on the way to scale up to a go-to location for benchmark datasets, and as such cannot be maintained in an ad-hoc manner and requires consistent financing to provide required services.

7.1. User fees

We formulated a business plan and begin adoption in 2021. Until that point users publishing data in DRP where not charged any fees as initial development (2015-2018) was supported by an NSF grant, and to this day rather limited support is available through Digital Porous Media Industry Affiliate Program at The University of Texas at Austin. In May and June 2021 DRP initiated a request for comments about establishing a funding model based on a) one-time publication fees or b) a DRP subscription plan to start in September of 2021. The community discussion was initiated via the DRP newsletter to all users (available on [22]) as well as through personal and email communications to the biggest DRP data providers. Documentation on DRP user agreement changes [44] has complete details of the proposed fee structure. Responses were publicly posted without personal identifiable information in DRP newsletter from July 13, 2021, linked to [22]. Five groups that publish regularly expressed support for the fee structure after the necessity was explained. The push back from the rest of the researchers came from the fact that they do not have open data fees built into their budgets. The fee structure is in Table 2 and the new user agreement is available on [11].

Table 5. The user fee structure implemented in September 2021.

Individual (one time per project)	< 2 GB	2-500 GB	500 GB – 1 TB
Open data publication fee	Free	$1,500	$ 3,000
Institutional membership (annual)	**< 500 GB**	**500 GB - 1TB**	**1 – 3TB**
Academic/non-profit	$ 5,000	$ 7,500	$ 10,000
Company/private	$ 7,500	$ 10,000	$ 12,500

7.2. Donations and sponsorships

Anyone can donate to or support Digital Rocks Portal: this is tax-deductible and easily done via credit card using UT Austin's Giving platform. Contributions of $100 or more are thanked with a promotional mug, mouse pad or a calendar. Table 6 below provides details on sponsorship levels.

Table 6. Donation and sponsorship levels.

Level	Type	Details
$20+	Any time giving	Tax deductible and easily done via credit card.
$500	Newsletter sponsor	You and/or your company will be featured in next newsletter.
$1,500	Project sponsor	The donation will: (1) make one of the Folk McBride high resolution thin sections examples available for research, (2) permanently place a thank-you note to you/your company on the project landing page, and (3) make you an owner of promotional items with that

		project imagery. Check this presentation for suggested datasets. Any other (existing or new) project can be sponsored under the same terms.
$3,000 or more	Student/fellow sponsor	Sponsor a student fellow (suggested $3000-5000) who will help upload data from partner institutions and research groups and create advanced visualizations to accompany them in the semester following the donation.

7.3. Expanding the user base

Last but not the least, many energy storage and materials engineering samples are porous and require image analysis, simulation and ML techniques that are comparable to digital rocks physics. Technology transfer and change of the name to Digital Porous Media is imminent.

Acknowledgements

References

[1] M.J. Blunt, Multiphase flow in permeable media: a pore-scale perspective, Cambridge University Press, Cambridge, United Kingdom ; New York, NY, 2017.

[2] A. Mehmani, R. Verma, M. Prodanović, Pore-scale modeling of carbonates, Marine and Petroleum Geology. 114 (2020) 104141. https://doi.org/10.1016/j.marpetgeo.2019.104141.

[3] Y.D. Wang, M.J. Blunt, R.T. Armstrong, P. Mostaghimi, Deep learning in pore scale imaging and modeling, Earth-Science Reviews. 215 (2021) 103555. https://doi.org/10.1016/j.earscirev.2021.103555.

[4] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks, in: F. Pereira, C.J.C. Burges, L. Bottou, K.Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 25, Curran Associates, Inc., 2012: pp. 1097–1105. http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf (accessed August 15, 2019).

[5] J.E. Santos, Y. Yin, H. Jo, W. Pan, Q. Kang, H.S. Viswanathan, M. Prodanović, M.J. Pyrcz, N. Lubbers, Computationally Efficient Multiscale Neural Networks

Applied to Fluid Flow in Complex 3D Porous Media, Transp Porous Med. (2021). https://doi.org/10.1007/s11242-021-01617-y.

[6] J.E. Santos, D. Xu, H. Jo, C.J. Landry, M. Prodanović, M.J. Pyrcz, PoreFlow-Net: a 3D convolutional neural network to predict fluid flow through porous media, Advances in Water Resources. (2020) 103539. https://doi.org/10.1016/j.advwatres.2020.103539.

[7] TensorFlow, TensorFlow. (n.d.). https://www.tensorflow.org/ (accessed August 15, 2019).

[8] J. Santos, Y. Yin, M. Prodanović, H. Khan, N. Lubbers, 3D Collection of Binary Images, (2021). https://doi.org/10.17612/NXGK-K611.

[9] Y. Niu, Y. Da Wang, P. Mostaghimi, J.E. McClure, J. Yin, R.T. Armstrong, Geometrical-Based Generative Adversarial Network to Enhance Digital Rock Image Quality, Phys. Rev. Applied. 15 (2021) 064033. https://doi.org/10.1103/PhysRevApplied.15.064033.

[10] J.E. McClure, J. Yin, R.T. Armstrong, K.C. Maheshwari, S. Wilkinson, L. Vlcek, Y. Da Wang, M.A. Berrill, M. Rivers, Toward Real-Time Analysis of Synchrotron Micro-Tomography Data: Accelerating Experimental Workflows with AI and HPC, in: J. Nichols, B. Verastegui, A. 'Barney' Maccabe, O. Hernandez, S. Parete-Koon, T. Ahearn (Eds.), Driving Scientific and Engineering Discoveries Through the Convergence of HPC, Big Data and AI, Springer International Publishing, Cham, 2020: pp. 226–239. https://doi.org/10.1007/978-3-030-63393-6_15.

[11] M. Prodanovic, M. Esteva, M. Hanlon, G. Nanda, P. Agarwal, Digital Rocks Portal, (2015). https://doi.org/10.17612/P7CC7K.

[12] Mendeley Data, (n.d.). https://data.mendeley.com/ (accessed September 1, 2019).

[13] Dryad, (n.d.). http://datadryad.org/ (accessed February 22, 2017).

[14] Energy Data Exchange (EDX), (n.d.). https://edx.netl.doe.gov/ (accessed February 19, 2017).

[15] S. Sweat, A. Ranganath, M. Esteva, M. Prodanović, User Guided Design: Building Confidence in Engineering Data Publication, in: University of Texas at Austin, Denver, CO, 2016. https://doi.org/10.15781/T27940X4W.

[16] S. Fuchs, D. Crandall, M. Gill, J. Moore, B. Kutchko, Foamed Cement API RP 10B-4, (2018). https://doi.org/10.17612/p7gd4q.

[17] M. Souzy, H. Lhuissier, Y. Méheust, T. Le Borgne, B. Metzger, Experimental 3D Velocity Field in Random Sphere Packing, (2020). https://doi.org/10.17612/HDP8-0149.

[18] C. Peters, S. Hajireziae, J. Kim, D. Crandall, Mineralized fracture in a mudrock shale, (2021). https://doi.org/10.17612/P4WH-W011.

[19] DataCite Fabrica, DataCite Fabrica. (n.d.). https://doi.datacite.org/ (accessed May 13, 2022).

[20] H. Khan, M. Prodanovic, D. Dicarlo, Particulate straining in carbonate proxy, Digital Rocks Portal, 2018. https://doi.org/10.17612/P7D96T.

[21] TACC Visualization Portal, (n.d.). https://vis.tacc.utexas.edu/.

[22] Digital Rocks Portal About section, (n.d.). https://www.digitalrocksportal.org/about/.

[23] M.D. Wilkinson, M. Dumontier, Ij.J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L.B. da Silva Santos, P.E. Bourne, J. Bouwman, A.J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C.T. Evelo, R. Finkers, A. Gonzalez-Beltran, A.J.G. Gray, P. Groth, C. Goble, J.S. Grethe, J. Heringa, P.A.C. 't Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S.J. Lusher, M.E. Martone, A. Mons, A.L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M.A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A.

Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, B. Mons, The FAIR Guiding Principles for scientific data management and stewardship, Scientific Data. 3 (2016) 160018. https://doi.org/10.1038/sdata.2016.18.

[24] ORCID, (n.d.). https://orcid.org/ (accessed May 14, 2022).

[25] Digital Rocks Portal: Visualization - YouTube, (n.d.). https://www.youtube.com/playlist?list=PL7QcRCNp33XA7BBvpNbwsDZjLB-mezkUu (accessed May 16, 2022).

[26] ImageNet, (n.d.). https://image-net.org/index (accessed October 20, 2021).

[27] Quantum-Machine.org: Home, (n.d.). http://quantum-machine.org/ (accessed May 16, 2022).

[28] J.E. Santos, M.J. Pyrcz, M. Prodanović, 3D Dataset of binary images: A collection of synthetically created digital rock images of complex media, Data in Brief. 40 (2022) 107797. https://doi.org/10.1016/j.dib.2022.107797.

[29] J. Santos, Chang, Bernard, Gigliotti, Alex, Kang, Qinjun, Lubbers, Nicholas, Viswanathan, Hari, Prodanovic, Masa, 3D Dataset of Simulations, (2021). https://doi.org/10.17612/93PD-Y471.

[30] J.E. Santos, Chang, Bernard, Gigliotti, Alex, Y. Yin, W. Song, Prodanovic , Masa, Q. Kang, Lubbers, Nicholas, H.S. Viswanathan, A Dataset of 3D Structural and Simulated Transport Properties of Complex Porous Media, Nature Scientific Data. Under Review (n.d.).

[31] Y.D. Wang, R. Armstrong, P. Mostaghimi, A Diverse Super Resolution Dataset of Digital Rocks (DeepRock-SR): Sandstone, Carbonate, and Coal, (2019). https://doi.org/10.17612/S3M9-E024.

[32] Y. Niu, Armstrong, Ryan, Mostaghimi, Peyman, Unpaired super-resolution on micro-CT sandstone by using cycle-consistent generative adversarial network, (2020). https://doi.org/10.17612/VZTT-YX38.

[33] N. Alqahtani, P. Mostaghimi, R. Armstrong, A Multi-Resolution Complex Carbonates Micro-CT Dataset (MRCCM), (2021). https://doi.org/10.17612/3T36-Q704.

[34] M. Tawfik, Z. Karpyn, S.X. Huang, scCO2-Brine-Glass Dataset for Comparing Image Denoising Algorithms, (2022). https://doi.org/10.17612/A1QA-2A25.

[35] K. Singh, M.J. Blunt, High resolution X-ray micro-tomography datasets for in-situ effective contact angle analysis in carbonate rocks, (2018). https://doi.org/10.17612/P7D95F.

[36] J.E. McClure, Z. Li, M. Berrill, T. Ramstad, The LBPM software package for simulating multiphase flow on digital images of porous rocks, Comput Geosci. 25 (2021) 871–895. https://doi.org/10.1007/s10596-020-10028-9.

[37] Palabos CFD and Complex Physics, (n.d.). http://www.palabos.org/ (accessed January 16, 2014).

[38] J.E. Santos, A. Gigliotti, A. Bihani, C. Landry, M.A. Hesse, M.J. Pyrcz, M. Prodanović, MPLBM-UT: Multiphase LBM library for permeable media analysis, SoftwareX. 18 (2022) 101097. https://doi.org/10.1016/j.softx.2022.101097.

[39] R. Verma, M. Prodanovic, Level Set Method based Progressive Quasi-Static (LSMPQS) software 1.0, 2018. https://github.com/rahulverma88/lsmpqs-1.0.

[40] National Institute of Standards and Technology, NIST. (n.d.). https://www.nist.gov/ (accessed May 17, 2022).

[41] J.T. Gostick, Versatile and efficient pore network extraction method using marker-based watershed segmentation, Phys. Rev. E. 96 (2017) 023307. https://doi.org/10.1103/PhysRevE.96.023307.

[42] OpenPNM | Home | An Open Source Pore Network Modeling Package, (n.d.). http://openpnm.org/ (accessed November 1, 2015).

[43] J. Gostick, M. Aghighi, J. Hinebaugh, T. Tranter, M.A. Hoeh, H. Day, B. Spellacy, M.H. Sharqawy, A. Bazylak, A. Burns, W. Lehnert, A. Putz, OpenPNM: A Pore Network Modeling

Package, Computing in Science & Engineering. 18 (2016)
60–74. https://doi.org/10.1109/MCSE.2016.49.

[44] Digital Rocks Portal User Agreement Documentation, (n.d.).
https://www.digitalrocksportal.org/user-agreement-changes/.