PROPOSAL FOR THE 3RD ROUND OF THE JUPYTER COMMUNITY WORKSHOP

Accepted for funding. See: https://github.com/JCMSK/2022_JCW/blob/main/README.md

Your Name and Affiliation

Serena Bonaretti, Transparent MSK Research, Maastricht, The Netherlands (https://tmskr.github.io/), on behalf of the nine members of the Musculoskeletal Imaging Community listed in the paragraph about participants, who collaboratively worked on this proposal

Email

serena.bonaretti.research@gmail.com

Brief title for your event

Building the Jupyter Community in Musculoskeletal Imaging Research

Describe the goal/topic of the event

For years, we, members of the musculoskeletal (MSK) imaging community, have been looking for ways to standardize our research computational workflows, make them reproducible and open, and share them to foster collaborative research. Currently, our computational workflows are mainly made of separate subunits that use various proprietary software, closed in-house algorithms, and different file formats, thus challenging collaborations, reducing efficiency, and potentially decelerating discoveries for relevant musculoskeletal conditions, such as osteoarthritis and osteoporosis. We are thrilled that we have identified Jupyter notebook as the perfect solution to our problem. The intrinsic narrative nature of Jupyter notebook is ideal for our workflows, which mainly consist of loading medical images, processing and analyzing the images, and visualizing the results as both labelled images and graphs with descriptive statistics. Thanks to Jupyter notebook, we will be able to create comprehensive workflows in a simple yet powerful environment, share and compare algorithms, and integrate workflows into our publications.

The Jupyter Community Workshop is a unique opportunity for us to start translating our aspirations into a concrete change for our community. The main goal of the workshop will be to start building an active and enthusiastic MSK Imaging Jupyter Community to create open and reproducible workflows for MSK imaging research. Our specific aims will be to: 1) learn the basics of JupyterLab, Jupyter notebook, and SimpleITK (package for medical imaging) in python, and 2) create MSK-oriented Jupyter notebook templates that we can use, adapt, and share inside and outside our community.

How does the workshop involve strategic work related to the central open-standards, protocols, abstractions and architecture, and open-source subprojects of Project Jupyter?

We are convinced that the outcome of the workshop will have a tremendous impact on our community and potentially on the larger Jupyter ecosystem for the following reasons:

1. During the workshop we will start creating templates of computational workflows for MSK imaging research in Jupyter notebook using python. Instantiations of these templates will become essential components of collaborative research and publications. We envision that other research

communities will be able to adapt these templates for their specific needs. Specifically, we plan to start two templates, regarding:

- Data description. We will create a Jupyter notebook template that: 1) reads subject data from a tabular file, extracts demographics (e.g. sex and age), and creates summary statistics; and 2) loads medical image information, extracts and summarizes image characteristics (e.g. resolution and dimension). Instantiations of this template will eventually be integrated in the *Materials* session of our papers;
- Image analysis workflows. We will initiate a Jupyter notebook template that: 1) loads MSK images acquired with different imaging modalities (e.g. Computed Tomography and Magnetic Resonance); 2) allocates space for algorithms; 3) visualizes results; and 4) prints out dependencies for future reproducibility of computational environments. In addition, during breakout sessions, we will start developing algorithms for specific implementations that will be part of the point 2) (see the draft of the workshop program below). Instantiations of these notebooks will eventually be attached to the paragraphs *Methods and Results* of our papers. For this task we are inspired by the Brain Imaging community who has been developing open software in python for many years (https://nipy.org/) and by the geophysics community who is creating a modular and scalable ecosystem for the next generation of open-source analysis tools (Pangeo).

2. We envision that adopting Jupyter notebook as our main tool for computational workflows will encourage us to create new tools to compare algorithms, define new standards for workflows, design systems to organize notebooks, and establish solutions to assign digital object identifiers (DOIs) to notebooks. In addition, we anticipate that our achievements will be of interest for (but not limited to) other imaging communities, who could potentially adapt the notebook templates to their specific needs and implementations.

How will the workshop lead to the growth and sustainability of the Jupyter community? How will it grow the size and health of the core Jupyter project contributor community? Of the broader Jupyter ecosystem of which Jupyter is a part?

We aim to grow the MSK Jupyter community starting with the workshop itself. We will make sure to involve people from different subfields of the MSK imaging community and we will promote our efforts to reach other research communities. As requested, within 2 weeks following the end of the workshop we will write a guest blog post on the Jupyter Blog with a description of the activities, results, and outcomes of the workshop. In addition, within 4 months of the end of the workshop we will write a short communication describing our experience and achievements during the workshop. We will publish the preprint on bioRxiv, and we will submit the manuscript to an open access journal. Back at our institutions, we will share acquired knowledge with students, colleagues, and other groups. In addition, we will promote the community, its work, and future directions during conferences and social media (e.g. Twitter). Finally, all material will be accessible on GitHub, freely available to anyone.

What types of people would attend the event and why does their participation align with the goals of Jupyter Workshops in general and your particular goals for this workshop?

Participants of the workshop will be members of the MSK imaging community who are willing to learn and develop tools for the Jupyter ecosystem. We aim to include participants with diverse background skills, i.e. processing and analysis of magnetic resonance (MR) and computed tomography (CT and HR-pQCT) images, finite element simulations of bones and joints, and expertise in osteoporosis and osteoarthritis. We expect that the majority of the participants will not be familiar with Jupyter notebook but extremely eager to learn, develop, and promote open and reproducible MSK research using Jupyter notebook. During the workshop, the learning experience will be guided by S. Bonaretti with the support of M. Barzegari. They have extensive experience with Jupyter notebook, python, Binder, and the Jupyter ecosystem for open source and reproducible research. We are currently looking for other Jupyter notebook experts in our community that will be able to support our learning experience. The workshop will be strategically placed in the week between two major international MSK conferences (International Workshop on Osteoarthritis Imaging (IWOAI) and European Society of Biomechanics (ESB)) held in Europe this summer, and we will aim to attract participants to attend the workshop before or after one of these conferences. The following participants have confirmed their presence at the workshop and have provided brief biographies and personal motivations in the last question of the proposal:

1. Serena Bonaretti, Founder and Research Scientist at Transparent MSK Research (The Netherlands)
2. Andrew Burghardt, Technical Director of the Quantitative Micro-Imaging Facility at University of California, San Francisco (USA)
3. Kathryn Stok, Senior Lecturer of Mechanobiology at the University of Melbourne (Australia)
4. Sarah Manske, Assistant Professor of Radiology (Image Science) at University of Calgary (Canada)
5. Mojtaba Barzegari, PhD student in Biomechanics at Leuven University (Belgium)
6. Lorenzo Grassi, Research Scientist in Biomechanics at Lund University (Sweden)
7. Enrico Schileo, Research scientist in Biomechanics at Rizzoli Orthopaedic Institute, Bologna (Italy)
8. Fulvia Taddei, Head of Bioengineering and Computing Laboratory at Rizzoli Orthopaedic Institute, Bologna (Italy)
9. Julio Carballido Gamio, Assistant Professor of Radiology at University of Colorado, Denver (USA)

Upon receiving confirmation of funding, we will promote the workshop through emails to the participants of IWOAI and ESB, to our collaborators, and on social media (e.g. Twitter).

How does the event reach users/contributors that are underserved or underrepresented in our community?

We will do our best to create an inclusive environment where participants will bring different competences and perspectives. We aim to have participants with these characteristics:

1. 20% Jupyter users, 80% new users
2. 50% women, 50% men
3. 50% junior researchers (until Postdoc), 50% senior researchers (after postdoc)
4. 50% European, 50% international.

Approximate event dates (no later than August 2020)

The workshop will be held on July 7-9, 2020. These dates are in the week between two major international MSK conferences held in Europe:

1. 14th International Workshop on OsteoArthritis Imaging (IWOAI), held in Rotterdam, The Netherlands, on July 1-4, 2020 (https://www.iwoai.org/);
2. 26th Congress of the European Society of Biomechanics (ESB) , held in Milan, Italy, on July 12-15, 2020 (https://esbiomech.org/conference/esb2020/).

Proposed venue

The workshop will be held in Maastricht, The Netherlands. Maastricht is 2 hours by train from Rotterdam, where IWOAI will take place, and is one-and-a-half-hour flight from Milan, where ESB will occur.

What is your proposed budget? Include a total amount you will not exceed and a high-level line-item breakdown of expense categories. Is there any other funding being provided from other sources?

We have investigated several options to contain our budget. The cheapest solution we have found is The Student Hotel (https://www.thestudenthotel.com/maastricht/), which will provide a classroom, meals (except dinners) and bedrooms for three days for 20 people for a total amount of about 8,500 euros. This solution is equivalent to a free classroom at Maastricht University to which we have to separately add costs for catering (coffee breaks and lunches) and hotel rooms. We think that being in the same environment both during work hours and free time will favor our discussions, collaborations, and sense of community. In addition, we estimate a total of 1,120 euros for two dinners for 20 people. Finally, we estimate a total of 7,660 euros for transportation. This corresponds to 6 people coming from the Maastricht region, 10 people coming from Europe, and 4 people coming from outside Europe. Two of the confirmed participants coming from outside Europe will cover their trip to Europe and will only need transportation inside Europe. Thus, the final amount we would like to request is about 17,260 euros, which corresponds to about 19,160 USD (exchange rate = 1.11 on Google on Dec 12). More details about our estimated budget are at the end of this proposal.

We are currently negotiating support by Helis Academy (https://helisacademy.com/en), a Dutch industrial-academic partnership supporting education in the Life Sciences and Health. In case of their approval, we will be happy to reduce our budget request.

All events funded by Jupyter require the adoption and enforcement of our Code of Conduct. Do you agree to serve as our proxy to address any COC complaints that may arise at the event?

Yes

Is there any other information you'd like for us to consider?

We would like to let you know who we are, share the preliminary program, and how we calculated the preliminary budget.

Who we are

- Serena Bonaretti: After more than six years in the Departments of Radiology at University of California, San Francisco and Stanford University, I have recently founded Transparent MSK

Research to develop and promote open and reproducible research in the musculoskeletal imaging community. My last release is pyKNEEr, an image analysis workflow specifically designed for open and reproducible research on femoral knee cartilage. It is written in python and uses Jupyter Notebooks as a user-interface (https://sbonaretti.github.io/pyKNEEr/). In addition, I promote transparent research with workshops, seminars, and a YouTube channel. In January 2020, I will conduct the workshop "The Basics of Python and Jupyter Notebook for Medical Image Analysis" during OpenMR Benelux (https://openmrbenelux.github.io/page-openmrb-2020/), and my last video "How to: Binder" was endorsed on Twitter by one of the creators of Binder (https://twitter.com/choldgraf/status/1200081533853294592).

- Andrew Burghardt: I am a research scientist and engineer in the Department of Radiology and Biomedical Imaging at the University of California, San Francisco and Co-Director of the Quantitative Micro-Imaging Facility (QMIF). My research program is focused on the technical development of quantitative imaging techniques to study musculoskeletal disease in humans, including osteoporosis, rheumatoid arthritis, rare bone diseases, and skeletal growth and development. My group is responsible for running centralized quality assurance, analysis, and computational tasks for large multicenter studies funded by NIH. Furthermore I consult to a number of investigators in the field on implementing standard image acquisition and analysis protocols using High Resolution peripheral Quantitative Computed Tomography (HR-pQCT). I am highly motivated to transition my group workflows from a highly niche, proprietary platform to an open development platform based on python and Jupyter Notebooks.
- Kathryn Stok: I am a Senior Lecturer in the Department of Biomedical Engineering at the University of Melbourne, and an innovative biomedical engineer in quantitative microstructural imaging (micro-computed tomography) and biomechanics of cartilage and joint structures. I was awarded a PhD from ETH Zurich, Switzerland in 2008. This work involved devising novel imaging platforms for quantitative measurement of articular cartilage. In September 2016, I took up a position at the University of Melbourne to continue this work. I lead the Integrative Cartilage Research Group, and am a co-founder of the Melbourne School of Engineering Mechanobiology Lab. My work involves the development of novel, in vivo, imaging-based measurement technologies for musculoskeletal health in 3D at different length scales (www.integrativecartilageresearchgroup.com). I am looking forward to provide my contribution for open and reproducible musculoskeletal image analysis tools.
- Sarah Manske: I am an Assistant Professor in the Division of Image Science, Department of Radiology at the University of Calgary. My research utilizes multi-modal imaging to understand the development and progression of musculoskeletal disease. I am interested in developing open image analysis workflows to investigate bone damage in arthritis. My lab is starting to develop workflows using python and Jupyter notebook (https://github.com/ManskeLab) after many years of experience using proprietary software for high resolution- and micro-computed tomography.
- Mojtaba Barzegari: I am a PhD Researcher at KU Leuven in the field of computational bioengineering. My research is mainly focused on developing mathematical models and high-performance numerical simulations of tissue engineering systems. I use solely open source packages and frameworks in my research, and apart from technical advantages, this enables me to contribute more efficiently in open science community. I also share my work via reproducible

science disciplines (https://mbarzegary.github.io), one important part of which is a collection of Jupyter Notebooks with Python and C++ kernels.

- Lorenzo Grassi: I am a Research Scientist at the Department of Biomedical Engineering at Lund University, Sweden. My main research interest is predicting fracture risk for elderly people using clinical images. This involves image analysis and automatic generation of subject-specific finite element (FE) models. For most of the steps of the workflow, I have either in-house code or open-source packages to rely on. However, this fragmentation is hindering reproducibility of some studies as well as the adoption of the developed methods by other groups. I am looking at Jupyter notebooks as a way to include the tools developed to analyze bone quality for elderly people in a coherent framework accessible to everyone in the research community.
- Enrico Schileo: I am a senior research scientist at the Bioengineering and Computing Laboratory of the Rizzoli Orthopaedic Institute in Bologna. I heavily rely on workflows for the generation and solution of numerical models of bones. Sometimes, these models are integrated with other, image-based MSK workflows, to receive input (e.g. forces from musculoskeletal multibody models) or interpret output (e.g. morphological and densitometric analysis to clarify strength determinants). The need to develop and glue chunks of code has been a constant source of worries for me, and a clear hurdle to the clinical application of my research work. To foster open and reproducible research, together with colleagues in my lab we develop and share Bonemat (www.bonemat.org), a freeware that maps on a finite element mesh bone elastic properties derived from QCT. While Bonemat is highly cited and used in our research community, it is still just a tile of the puzzle. I am looking at Jupyter as a powerful framework to comprehensively develop, and document our MSK image based workflows, integrating or wrapping Bonemat too. A dedicated event to engage with other researchers would be ideal to share (thus improve!) MSK workflows with the research community.
- Fulvia Taddei: I am Head of the Bioengineering and Computing Lab of Istituto Ortopedico Rizzoli and I've been working for more than twenty years in the field of MSK modelling from Medical Images. I've always strongly supported the open-source approach believing it the only real possibility to promote reproducible research. We are developing and maintaining ALBA (Agile Library for Biomedical Application, https://github.com/IOR-BIC/ALBA): an open-source rapid development framework (in C++) on GitHub based on the VTK library among others. On this platform we have developed the Bonemat software (www.bonemat.org) for the automatic mapping of bone material properties on CT-based Finite Element Models, and NMSBuilder software (www.nmsbuilder.org) for the generation of subject-specific MSK models that exports directly towards OpenSim solver. However the workflow for the generation of MSK models is complex, and we only have, and have shared, some chunks of the whole process. I strongly support the initiative to develop a Jupyter platform in which other contributions may be exposed and connected to go forward in the direction of a complete MSK workflow shared among our research community.
- Julio Carballido-Gamio: I am an Assistant Professor in the Department of Radiology at the University of Colorado Anschutz Medical Campus. I have extensive experience in developing and implementing multiple algorithms for the analysis of MSK images of different anatomical regions and medical imaging modalities including MRI, QCT and HR-pQCT. I am extremely interested in

Jupyter Notebook for the development of an Image Analysis Workflow that enables the easy integration of existing and novel image analysis algorithms. This workflow will benefit: 1) clinical imaging research groups mainly focused on the application of existing techniques to the study of MSK pathologies, and 2) research groups developing new image analysis techniques. I am convinced that the Jupyter ecosystem will have a huge impact in the MSK imaging research community because it will facilitate the development of new image analysis algorithms and their transparent multi-site validation. Jupyter is the perfect platform towards the long-time desire of standardizing MSK image analysis pipelines, and the open source nature of Jupyter will improve existing collaborations and promote new ones.

Preliminary program

Tuesday, July 9, 2020 – Learning (S. Bonaretti, M. Barzegari):

- Introduction to Jupyter Lab, Jupyter notebook and SimpleITK (python package for medical imaging)
- First joint hackathon session to create the Jupyter notebook template on Data description

Wednesday, July 10, 2020 - Hackathon:

- Start creating Jupyter notebook templates for Image analysis workflows. We will divide into four groups with the following goals:
    - SPECTRA: To initiate a Jupyter notebook with algorithms for bone erosion in HR-pQCT images of arthritis
    - pyKNEEr: To build a Jupyter notebook to analyze tibial cartilage
    - Biomechanics group: To create a Jupyter notebook that includes algorithms for automatic analysis of human femurs from clinical CT data to predict bone strains and strength
    - Image registration: To create a template for multi-scale intra-subject gray-level-based image registration with normalized mutual information.

Thursday, July 11, 2020 - Hackathon:

- Hackathon + Final discussion of achievements, lessons learned, challenges, and future steps