Moving beyond the desktop: Prospects for Practical Bioimage Analysis via the Web

Wei Ouyang[1#], Kevin W. Eliceiri[2] and Beth A. Cimini[3]

[1]Department of Applied Physics, Science for Life Laboratory, KTH Royal Institute of Technology, Stockholm, Sweden
[2]Morgridge Institute for Research and Center for Quantitative Cell IMaging, University of Wisconsin-Madison, Madison, WI
[3]Imaging Platform, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

#Corresponding Author: wei.ouyang@scilifelab.se

Abstract

The rapid advancements in biological imaging and the increasing complexity of image data call for a reevaluation of traditional bioimage analysis methods and how these methods are accessed. In our perspective, we argue that transitioning from desktop-based tools to web-based bioimage analysis offers immense opportunities for enhanced accessibility, improved collaboration, and streamlined workflows. We briefly describe the potential advantages, including lower computational requirements and overcoming challenges such as installation issues and limited reproducibility. Additionally, we discuss the current landscape of web-based tools, the challenges in implementation, and the importance of community involvement in driving this transition. By embracing web-based bioimage analysis, we believe the life sciences community can unlock new possibilities for accelerating biological research.

Main text

1. Introduction

Bioimage analysis has become an indispensable tool in modern biological research, playing a crucial role in advancing our understanding of complex biological systems. The rapid development of imaging instrumentation has resulted in the generation of massive image data sets, capturing information from the molecular to the organismal level. Concurrently, advances in artificial intelligence (AI) and computer vision have revolutionized bioimage analysis by enabling the automated extraction of quantitative data from these large-scale image collections.

Furthermore, bioimage analysis has become integral to spatial multi-omics research, where the integration of imaging data with transcriptomics, proteomics, and metabolomics data allows for a more comprehensive understanding of biological systems.

Despite the significant advancements in bioimage analysis, researchers still face challenges when it comes to collaboration, data sharing, and the implementation of diverse analysis approaches. The ever-increasing size and complexity of image datasets necessitate efficient storage, retrieval, and processing solutions. At present, the exchange of image data often relies on institutional storage services, cloud storage tools like Dropbox, or occasionally public shared image databases combined with ontologies for annotation. However, these methods can be cumbersome and may not fully address the requirements for seamless collaboration and data sharing.

Furthermore, diverse research teams and institutions often use different imaging modalities, file formats, and analysis pipelines, which can create hurdles in the seamless exchange of data and knowledge. The single-user, single-desktop paradigm that dominates much of the current bioimage analysis landscape can hinder collaboration, as it requires specialized software and hardware resources that may not be universally available to all researchers.

In this paper, we explore the prospects for moving beyond the desktop and embracing web-based approaches to overcome these challenges and enhance the practicality of bioimage analysis for researchers worldwide.

2. Current Approaches to Bioimage Analysis Collaboration

Conventional methods of sharing image files and analysis workflows often involve sending files via email or using cloud storage tools like Dropbox or Google Drive. However, these methods can be inadequate for handling large datasets shared among many users (Figure 1). Dedicated image data servers, such as OMERO(1), have been developed to address this issue by providing more efficient storage and management of large-scale datasets.

For public access, especially after publication, researchers commonly use shared databases such as the Image Data Repository (IDR)(2), the BioImage Archive(3) and The Cell Image Library(4). These platforms facilitate the exchange of information and promote interoperability between analysis tools by providing standardized metadata and file formats.

Ontologies, such as the EDAM Bioimaging Ontology(5), are also used for standardization of metadata, which greatly facilitates the exchange of information and helps the interoperability between analysis tools. The use of standardized file formats and metadata enables more efficient data sharing and collaboration while reducing the need to move data around.

Overall, current approaches to bioimage analysis collaboration rely on a combination of conventional file-sharing methods, dedicated image data servers, shared databases, and ontologies for metadata standardization. These methods often require moving data around, which is why they rely on standardized file formats for better compatibility and efficiency.

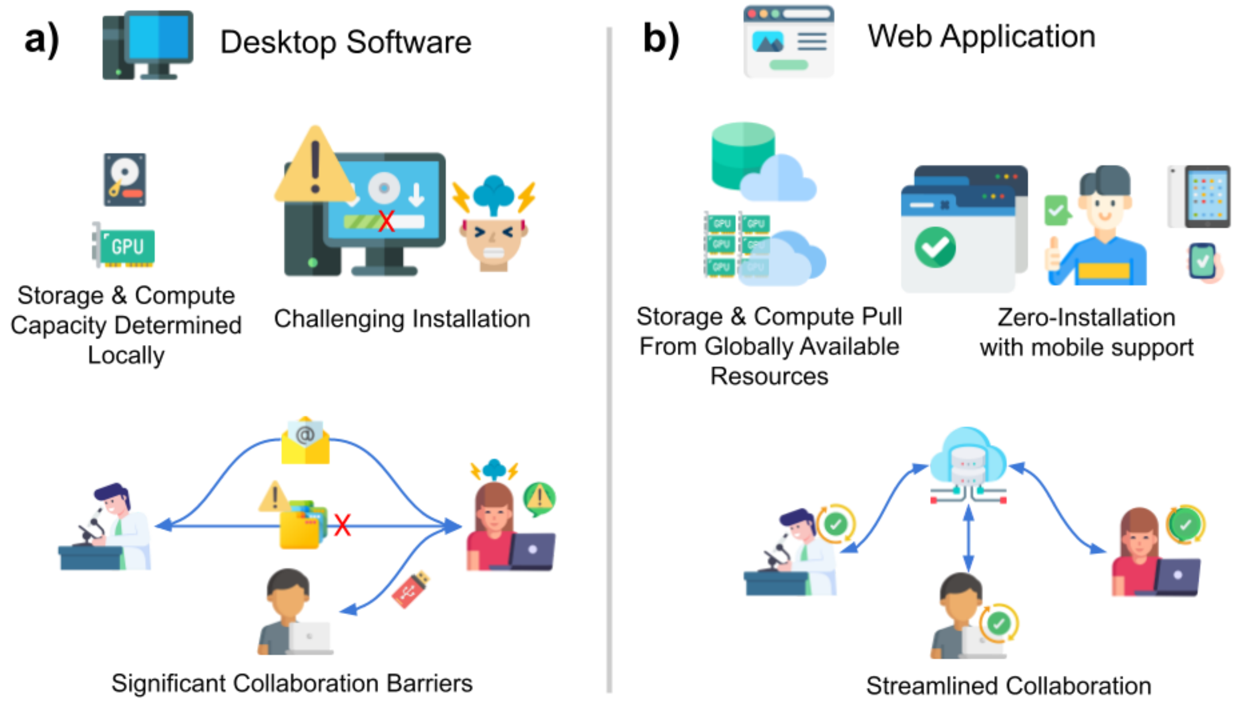

Image credit: Flaticon.com

Figure 1: Desktop vs Web bioimage analysis tools
Comparison of conventional desktop-based bioimage analysis software (a) with the emerging web-based bioimage analysis software (b). When using conventional software, one must work to handle installation of software on a local computer, the hardware requirements, and the difficulties in collaboration and reproducibility due to varying software versions and platform dependencies. Web-based tools, while they have their own disadvantages, have significantly improved ease of access, reduced computational requirements for the end user, and enhanced collaboration through shared resources and data. The figure emphasizes the potential benefits of transitioning from desktop to web-based bioimage analysis tools in life sciences research.

3. Web-Based Opportunities for Bioimage Analysis

Web-based bioimage analysis offers a promising alternative to conventional desktop-based analysis software. This approach encompasses native web tools as well as cloud-utilizing tools to enable researchers to initiate and carry out image analysis entirely through a web interface. Native web tools are applications built specifically for the web, allowing users to perform tasks such as image visualization, processing, and analysis directly within their web browsers. Accessing web resources refers to leveraging online databases, libraries, and repositories to enhance bioimage analysis workflows. Cloud resources enable researchers to offload computationally intensive tasks to remote servers, providing scalable computational power, storage, and specialized hardware for a wide range of applications.

One notable advantage of web-based tools is the ability to exchange data and analysis workflows through clickable URLs instead of explicitly downloading and uploading image files. This feature greatly improves user experiences, especially when working with tissue scans and large image files, as it simplifies data sharing and allows for streamlined collaboration among researchers.

Web-based bioimage analysis has the potential to address many of the limitations associated with the single-user, single-desktop paradigm by offering a more flexible, accessible, and collaborative platform for researchers. In the subsequent sections, we will discuss the benefits, existing tools, and challenges related to web-based bioimage analysis, as well as how the research community can transition from desktop-based to web-based analysis tools.

4. Benefits of Web-Based Bioimage Analysis

Web-based bioimage analysis offers a range of benefits that can improve the overall user experience, accessibility, and efficiency of biological research. One notable advantage is the lower computational comfort required for researchers who primarily focus on imaging rather than analysis (Figure 2). By offering user-friendly interfaces and streamlined workflows, web-based tools can help bridge the gap between image acquisition and data analysis, making it easier for researchers with limited computational expertise to engage in meaningful bioimage analysis.

An analysis of the image.sc forum, a popular online resource for bioimage analysis, reveals that a significant portion of discussions revolves around installation issues. Analysis performed for this paper showed that approximately 7-8% of recent posts (approximately 1 in every 12-14) in the Scientific Community Image Forum(6) (image.sc) contain the word "install". Furthermore, installation topics come up at least once in ~19% of the ~500 most recently posted threads, and are mentioned in the opening post 9.9% of the time. This suggests that about 1 in 10 threads in the forum are likely started due to an installation issue, and 1 in 5 threads end up discussing it at some point.

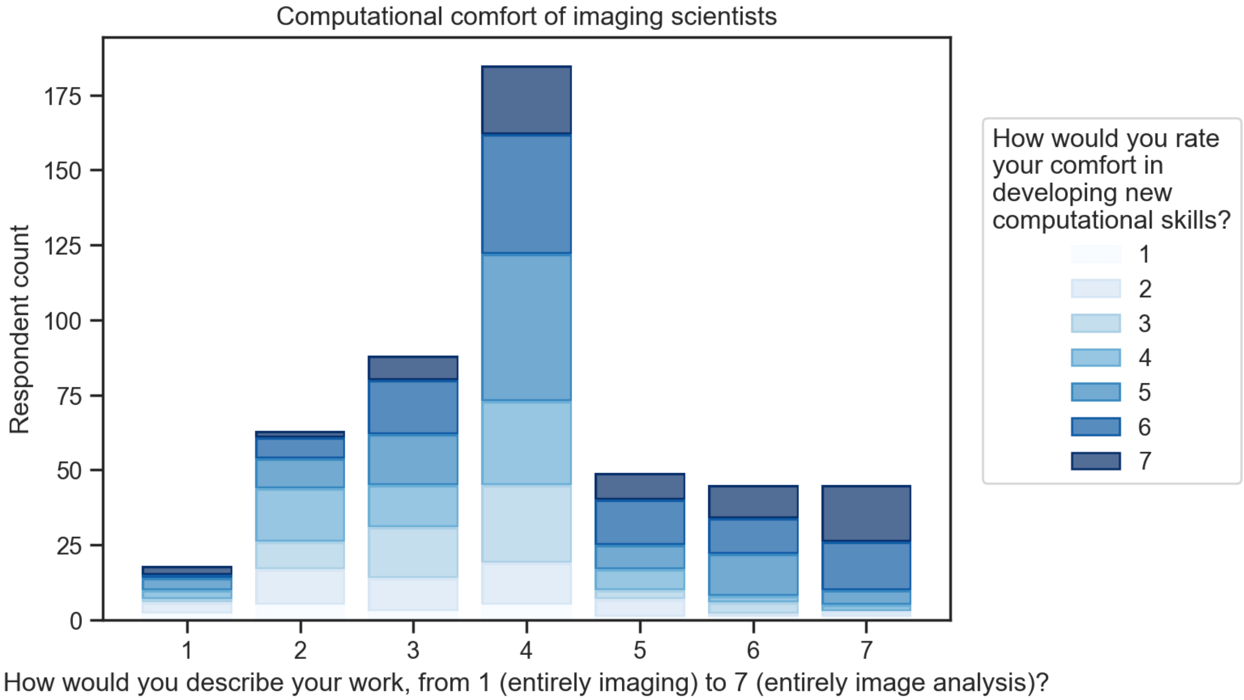

Figure 2: Computational comfort of scientists in a variety of work types. Approximately 500 respondents who use scientific imaging in the life or physical sciences were asked to describe their work on a scale from "entirely imaging" and "entirely image analysis"; in parallel, they were asked to report their comfort developing new computational skills. While the highest comfort levels are present in all groups, they are more common in users who spend more time performing analysis, and the lowest levels of comfort are represented almost entirely in the lower bins of computational work level. Data from Sivagurunathan, et al 2023 (in preparation).

The high frequency of installation-related discussions highlights the challenges researchers face when setting up and using traditional desktop-based bioimage analysis tools. By contrast, web-based tools can significantly alleviate these challenges by moving the installation burden entirely to the developer or by offering easy installation and updates, eliminating the need for complex configuration processes, and ensuring cross-platform compatibility.

5. Existing Tools for Web-Based Bioimage Analysis

Various tools have now been developed to facilitate web-based bioimage analysis. These tools can be categorized into scalable storage image formats, browser-based visualization and computing, and cloud or server-based image analysis platforms.

Scalable storage image formats, such as OME-NGFF(7), enable remote visualization in tools like Fiji(8)'s MoBIE(9). OME-NGFF is suitable for storing massive datasets in chunks and serving them using object storage, such as S3 or HTTP servers. Browser-based visualization tools, including Kitware's ITK VTK Viewer(10), Viv(11), (12)OpenSeadragon, TissUUmaps(13), and webknossos(14), offer easy-to-use and performant visualization due to their utilization of browser GPU standards, such as WebGL or WebGPU. These tools are ideally suited for accessing chunked storage formats like OME-NGFF or OME-Zarr(15).

Browser-based computing tools, such as Piximi(16), ImJoy and ImageJ.JS(17), can perform computation locally in web browsers. While useful and secure, since computation happens locally, their performance is limited by memory and compute constraints (e.g., 4GB memory per tab in Chrome) and lack of direct access to GPU accelerators, such as CUDA.

Cloud or server-based image analysis platforms, including ImJoy (with Jupyter(18)/BinderHub(19) backend), DeepCell Kiosk(20), Terra(19, 21), Galaxy(22), KNIME(23), CDeep3M(24), Bisque(25), BAND(26), Cytomine(27), Colab notebooks, and ZeroCostDl4Mic(28), provide a more scalable way of storing, sharing, and processing image data. These platforms are especially helpful for AI-based analysis pipelines, given the ongoing trend towards foundation models(29), such as Segment Anything(30), diffusion models(31), or transformer-based large language models. In cloud environments, containers are often used to create reproducible environments, orchestrated by cluster software like Kubernetes.

Despite the many advantages of these web-based tools, some challenges remain. For example, browser-based tools may have limited performance due to memory and compute constraints. Additionally, users may be hesitant to upload their image data to servers run by third-party service providers due to privacy concerns. Sometimes, hybrid computing is applied to take advantage of both the flexibility of browser-based computation and the powerfulness of server-side computation, e.g. in the Segment Anything demo application. Nonetheless, the continued development of web-based bioimage analysis tools offers promising opportunities to overcome the limitations of traditional desktop-based approaches.

6. Challenges in Implementing Web-Based Bioimage Analysis

While web-based bioimage analysis offers many benefits, there are several challenges to overcome when implementing these solutions:

1) Algorithmic challenges: Algorithms need to be tailored to fit the browser environment, which is typically single-threaded and memory-constrained. Optimization is often required to run efficiently within the browser. Remote servers can be used to complement storage and compute limitations.
2) Cross-language and algorithmic accessibility: With WebAssembly, algorithms implemented in languages such as C/C++ can be compiled to run in the browser. However, not all languages are readily compatible with web-based platforms, requiring additional effort for adaptation.
3) Deep learning integration: Browsers support deep learning frameworks like TensorFlow.js, some even leveraging WebGPU. However, due to browser constraints,

remote servers or local compute engines (e.g., Jupyter servers) are often necessary for more demanding tasks.

4) Handling big data: Although browsers can store data, they are better suited as clients for data handling when dealing with large datasets. Efficient data management strategies are needed to accommodate big data in web-based bioimage analysis.
5) Performance optimization: Striking a balance between limited local compute power and storage in the web browser, and reducing data transmission with remote servers, is crucial. Utilizing browser-based compute engines like WebWorkers and WebGPU can help optimize performance.
6) Efficient data transfer: Data compression techniques and browser-based compute can help reduce latency and data transmission. Efficient data transfer methods are essential for maintaining responsive web-based bioimage analysis tools.
7) Security concerns: In addition to conventional security measures such as encryption, edge computing using the browser can help alleviate privacy concerns. Ensuring data security and user privacy is vital when implementing web-based bioimage analysis tools.
8) Funding challenges: It may be challenging to gauge the cost needed to support the server infrastructure when hosting one's own server, especially if wide adoption increases traffic. Most projects are funded on short-term grant models, making it challenging to secure the future of a server long term, especially as funding mechanisms for maintenance are scarce relative to those for tool creation.
9) Additional challenges: Usability, maintenance, and support for web-based tools are also important considerations. Ensuring that tools are user-friendly and well-documented, while providing ongoing support and updates, will contribute to the success of web-based bioimage analysis platforms.

Addressing these challenges requires a collaborative effort from the research community and developers to create robust, efficient, and user-friendly web-based bioimage analysis tools.

7. Transitioning from Desktop-Based to Web-Based Analysis Tools

The transition from desktop-based to web-based bioimage analysis tools requires several key steps, including adapting existing tools, coding efforts, and fostering community involvement and support.

1) Adapting existing tools and approaches for web-based use: To make the most of current bioimage analysis tools, developers need to adapt them to work within the web environment. This process may involve refactoring code, optimizing algorithms for browser-based execution, and integrating tools with web-based platforms and services. For running desktop tools remotely in a cloud environment, it is essential to ensure that they can be executed in "headless" mode and provide batch processing capabilities with a scripting interface. New software should adopt distributed software design principles, such as separating the user interface from the computational part, allowing the UI to run

in the browser while computation occurs on a remote server. Remote Procedure Calls can be used to achieve this purpose.

2) Required coding efforts: Moving to web-based tools often necessitates learning new programming languages and frameworks, such as JavaScript, WebAssembly, or WebGL. Developers must also become familiar with cloud-based services and remote computing resources that can be leveraged for scalable storage and computation.
3) Community involvement and support: A successful transition to web-based bioimage analysis tools requires strong community involvement, including researchers, developers, and industry partners. Collaboration is essential for developing and maintaining web-based tools, sharing best practices, and addressing common challenges. Open-source development and the sharing of code, data, and resources can foster a more inclusive and efficient research community.

By addressing these aspects, the research community can work together to create a new generation of powerful, accessible, and collaborative web-based bioimage analysis tools that overcome the limitations of traditional desktop-based approaches.

8. Conclusion

Web-based bioimage analysis has the potential to revolutionize the way researchers collaborate, share data, and analyze complex biological images. By moving beyond the desktop, web-based tools can offer improved accessibility, scalability, and reproducibility, ultimately democratizing resources for resource-limited areas and enhancing collaboration across various user groups.

Overcoming the challenges associated with implementing web-based bioimage analysis tools requires concerted efforts from the research community, developers, and industry partners. By adapting existing tools and approaches, learning new programming languages and frameworks, and embracing community involvement and support, we can work towards a web-based paradigm that benefits all stakeholders.

As we transition from desktop-based to web-based analysis tools, it is crucial to maintain a focus on user experience, performance optimization, and security. The development of efficient, user-friendly, and secure web-based bioimage analysis platforms is a collective responsibility, and we must continue to innovate and share our experiences to facilitate progress.

In conclusion, web-based bioimage analysis holds immense promise for the future of biological research. By working together to overcome challenges and build robust, accessible tools, we can unlock the full potential of web-based bioimage analysis and contribute to groundbreaking discoveries in the life sciences.

References

1. C Allan; J-M Burel; J Moore; C Blackburn; M Linkert; S Loynton; D Macdonald; WJ Moore; C Neves; A Patterson; M Porter; A Tarkowska; B Loranger; J Avondo; I Lagerstedt; L Lianas; S Leo; K Hands; RT Hay; A Patwardhan; C Best; GJ Kleywegt; G Zanetti; JR Swedlow. OMERO: flexible, model-driven data management for experimental biology. *Nat Methods* 9, 245–253 (2012)
2. E Williams; J Moore; SW Li; G Rustici; A Tarkowska; A Chessel; S Leo; B Antal; RK Ferguson; U Sarkans; A Brazma; REC Salas; JR Swedlow. The Image Data Resource: A Bioimage Data Integration and Publication Platform. *Nat Methods* 14, 775–781 (2017)
3. M Hartley; GJ Kleywegt; A Patwardhan; U Sarkans; JR Swedlow; A Brazma. The BioImage Archive - Building a Home for Life-Sciences Microscopy Data. *J Mol Biol* 434, 167505 (2022)
4. Digital collections: Cell image library, https://library.ucsd.edu/dc/collection/bb42043698
5. M Kalaš; L Plantard; J Lindblad; M Jones; N Sladoje; MA Kirschmann; A Chessel; L Scholz; F Rössler; LN Sáenz; EG de Mariscal; J Bogovic; A Dufour; X Heiligenstein; D Waithe; M-C Domart; M Karreman; R Van de Plas; R Haase; D Hörl; L Paavolainen; IV Madunić; D Karaica; A Muñoz-Barrutia; P Sampaio; D Sage; S Munck; O Golani; J Moore; F Levet; J Ison; A Gaignard; H Ménager; C Zhang; K Miura; J Colombelli; P Paul-Gilloteaux. EDAM-bioimaging: the ontology of bioimage informatics operations, topics, data, and formats (update 2020). (2020)
6. CT Rueden; J Ackerman; ET Arena; J Eglinger; BA Cimini; A Goodman; AE Carpenter; KW Eliceiri. Scientific Community Image Forum: A discussion forum for scientific image software. *PLoS Biol* 17, e3000340 (2019)
7. J Moore; C Allan; S Besson; J-M Burel; E Diel; D Gault; K Kozlowski; D Lindner; M Linkert; T Manz; W Moore; C Pape; C Tischer; JR Swedlow. OME-NGFF: a next-generation file format for expanding bioimaging data-access strategies. *Nat Methods* 18, 1496–1498 (2021)

8. J Schindelin; I Arganda-Carreras; E Frise; V Kaynig; M Longair; T Pietzsch; S Preibisch; C Rueden; S Saalfeld; B Schmid; J-Y Tinevez; DJ White; V Hartenstein; K Eliceiri; P Tomancak; A Cardona. Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9, 676–682 (2012)

9. C Pape; K Meechan; E Moreva; M Schorb; N Chiaruttini; V Zinchenko; H Martinez Vergara; G Mizzon; J Moore; D Arendt; A Kreshuk; Y Schwab; C Tischer. MoBIE: a Fiji plugin for sharing and exploration of multi-modal cloud-hosted big image data. *Nat Methods* 20, 475–476 (2023)

10. M McCormick; P Elliott; S Jourdain; S Wittenburg; B Major; W Ouyang; L Abdala; D Smyth; I Popovych; P Hernandez-Cerdan; J Bilkey; B Kerautret; I Rosenblum. Kitware/itk-vtk-viewer: v14.41.3. (2023)

11. T Manz; I Gold; NH Patterson; C McCallum; MS Keller; BW Herr 2nd; K Börner; JM Spraggins; N Gehlenborg. Viv: multiscale visualization of high-resolution multiplexed bioimaging data on the web. *Nat Methods* 19, 515–516 (2022)

12. openseadragon: An open-source, web-based viewer for zoomable images, implemented in pure JavaScript. Github

13. L Solorzano; G Partel; C Wählby. TissUUmaps: interactive visualization of large-scale spatial gene expression and tissue morphology data. *Bioinformatics* 36, 4363–4365 (2020)

14. KM Boergens; M Berning; T Bocklisch; D Bräunlein; F Drawitsch; J Frohnhofen; T Herold; P Otto; N Rzepka; T Werkmeister; D Werner; G Wiese; H Wissler; M Helmstaedter. webKnossos: efficient online 3D data annotation for connectomics. *Nat Methods* 14, 691–694 (2017)

15. J Moore; D Basurto-Lozada; S Besson; J Bogovic; J Bragantini; EM Brown; J-M Burel; X Casas Moreno; G de Medeiros; EE Diel; D Gault; SS Ghosh; I Gold; YO Halchenko; M Hartley; D Horsfall; MS Keller; M Kittisopikul; G Kovacs; A Küpcü Yoldaş; K Kyoda; A le Tournoulx de la Villegeorges; T Li; P Liberali; D Lindner; M Linkert; J Lüthi; J Maitin-Shepard; T Manz; L Marconato; M McCormick; M Lange; K Mohamed; W Moore; N Norlin; B Özdemir; G Palla; C Pape; L Pelkmans; T Pietzsch; S Preibisch; M Prete; N Rzepka; S Samee; N Schaub; H Sidky;

AC Solak; DR Stirling; J Striebel; C Tischer; D Toloudis; I Virshup; P Walczysko; AM Watson; F Wong; KA Yamauchi; O Bayraktar; M Haniffa; N Hotaling; S Onami; LA Royer; S Saalfeld; O Stegle; FJ Theis; JR Swedlow. OME-Zarr: a cloud-optimized bioimaging file format with international community support. *bioRxiv* (2023)

16. Piximi, https://www.piximi.app/

17. W Ouyang; F Mueller; M Hjelmare; E Lundberg; C Zimmer. ImJoy: an open-source computational platform for the deep learning era. *Nat Methods* 16, 1199–1200 (2019)

18. T Kluyver; B Ragan-Kelley; F Pérez; B Granger; M Bussonnier; J Frederic; K Kelley; J Hamrick; J Grout; S Corlay; P Ivanov; D Avila; S Abdalla; C Willing. Jupyter Notebooks -- a publishing format for reproducible computational workflows. In: Positioning and Power in Academic Publishing: Players, Agents and Agendas. F Loizides, B Schmidt, eds. , IOS Press (2016)

19. Project Jupyter; M Bussonnier; J Forde; J Freeman; B Granger; T Head; C Holdgraf; K Kelley; G Nalvarte; A Osheroff; M Pacer; Y Panda; F Perez; B Ragan-Kelley; C Willing. Binder 2.0 - Reproducible, interactive, sharable environments for science at scale. In: Proceedings of the 17th Python in Science Conference (2018)

20. D Bannon; E Moen; M Schwartz; E Borba; T Kudo; N Greenwald; V Vijayakumar; B Chang; E Pao; E Osterman; W Graf; D Van Valen. DeepCell Kiosk: scaling deep learning-enabled cellular image analysis with Kubernetes. *Nat Methods* 18, 43–45 (2021)

21. Broad Institute of MIT and Harvard, Microsoft, and Verily. Terra.Bio, https://terra.bio/#

22. Galaxy Community. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Res* 50, W345–51 (2022)

23. MR Berthold; N Cebron; F Dill; TR Gabriel; T Kötter; T Meinl; P Ohl; C Sieb; K Thiel; B Wiswedel. KNIME: The Konstanz Information Miner. In: Studies in Classification, Data Analysis, and Knowledge Organization (GfKL 2007). Springer (2007)

24. MG Haberl; C Churas; L Tindall; D Boassa; S Phan; EA Bushong; M Madany; R Akay; TJ

Deerinck; ST Peltier; MH Ellisman. CDeep3M-Plug-and-Play cloud-based deep learning for image segmentation. *Nat Methods* 15, 677–680 (2018)

25. K Kvilekval; D Fedorov; B Obara; A Singh; BS Manjunath. Bisque: a platform for bioimage analysis and management. *Bioinformatics* 26, 544–552 (2010)

26. EMBL. EMBL Bioimage ANalysis Desktop (BAND), https://band.embl.de/#/eosc-landingpage

27. R Marée; L Rollus; B Stévens; R Hoyoux; G Louppe; R Vandaele; J-M Begon; P Kainz; P Geurts; L Wehenkel. Collaborative analysis of multi-gigapixel imaging data using Cytomine. *Bioinformatics* 32, 1395–1401 (2016)

28. L von Chamier; RF Laine; J Jukkala; C Spahn; D Krentzel; E Nehme; M Lerche; S Hernández-Pérez; PK Mattila; E Karinou; S Holden; AC Solak; A Krull; T-O Buchholz; ML Jones; LA Royer; C Leterrier; Y Shechtman; F Jug; M Heilemann; G Jacquemet; R Henriques. Democratising deep learning for microscopy with ZeroCostDL4Mic. *Nat Commun* 12, 2276 (2021)

29. R Bommasani; DA Hudson; E Adeli; R Altman; S Arora; S von Arx; MS Bernstein; J Bohg; A Bosselut; E Brunskill; E Brynjolfsson; S Buch; D Card; R Castellon; N Chatterji; A Chen; K Creel; JQ Davis; D Demszky; C Donahue; M Doumbouya; E Durmus; S Ermon; J Etchemendy; K Ethayarajh; L Fei-Fei; C Finn; T Gale; L Gillespie; K Goel; N Goodman; S Grossman; N Guha; T Hashimoto; P Henderson; J Hewitt; DE Ho; J Hong; K Hsu; J Huang; T Icard; S Jain; D Jurafsky; P Kalluri; S Karamcheti; G Keeling; F Khani; O Khattab; PW Koh; M Krass; R Krishna; R Kuditipudi; A Kumar; F Ladhak; M Lee; T Lee; J Leskovec; I Levent; XL Li; X Li; T Ma; A Malik; CD Manning; S Mirchandani; E Mitchell; Z Munyikwa; S Nair; A Narayan; D Narayanan; B Newman; A Nie; JC Niebles; H Nilforoshan; J Nyarko; G Ogut; L Orr; I Papadimitriou; JS Park; C Piech; E Portelance; C Potts; A Raghunathan; R Reich; H Ren; F Rong; Y Roohani; C Ruiz; J Ryan; C Ré; D Sadigh; S Sagawa; K Santhanam; A Shih; K Srinivasan; A Tamkin; R Taori; AW Thomas; F Tramèr; RE Wang; W Wang; B Wu; J Wu; Y Wu; SM Xie; M Yasunaga; J You; M

Zaharia; M Zhang; T Zhang; X Zhang; Y Zhang; L Zheng; K Zhou; P Liang. On the Opportunities and Risks of Foundation Models, http://arxiv.org/abs/2108.07258, (2021)

30. A Kirillov; E Mintun; N Ravi; H Mao; C Rolland; L Gustafson; T Xiao; S Whitehead; AC Berg; W-Y Lo; P Dollár; R Girshick. Segment Anything, http://arxiv.org/abs/2304.02643, (2023)

31. J Ho; A Jain; P Abbeel. Denoising Diffusion Probabilistic Models, http://arxiv.org/abs/2006.11239, (2020)

32. W McKinney. Data Structures for Statistical Computing in Python. In: Proceedings of the 9th Python in Science Conference. S van der Walt, J Millman, eds. , SciPy (2010)

33. M Waskom. seaborn: statistical data visualization. *J Open Source Softw* 6, 3021 (2021)

Acknowledgements

The authors utilized the language model ChatGPT developed by OpenAI for assistance in the structuring and drafting of this paper.

The authors thank [anyone?]

Code availability

Code supporting this manuscript can be found at https://github.com/COBA-NIH/WebImageAnalysisPerspective. Pandas(32) and seaborn(33) were used in graph creation.

Contributions

WO drafted the initial manuscript, to which BAC and KWE added revisions. WO and BAC created figures.

Funding

This work was supported by the European Commission through the Horizon Europe program (AI4LIFE project, grant agreement 101057970-AI4LIFE), the SciLifeLab & Wallenberg

Data Driven Life Science Program (grant: KAW 2020.0239), the National Institutes of Health (NIH P41 GM135019 to BAC and KWE). BAC (2020-225720) and KWE also acknowledge funding from the Chan Zuckerberg Initiative DAF, an advised fund of the Silicon Valley Community Foundation.

Ethics declaration

W.O. is affiliated with Amun AI AB, a commercial company that builds, delivers, supports and integrates AI-powered data management systems for academic, biotech and pharmaceutical industries. The remaining authors declare no competing interests associated with the manuscript.