Reimagining Scientific Discovery: Collective Intelligence Through the Convergence of Citizen Science and Computational Systems

Saiph Savage
Northeastern University
Boston, Massachusetts, USA

Lea Shanley
International Computer Science Institute
Berkeley, California, USA

Lucy Fortson
University of Minnesota
Minneapolis, Minnesota, USA

Corey Jackson
University of Wisconsin–Madison
Madison, Wisconsin, USA

Kevin Crowston
Syracuse University
Syracuse, New York, USA

Tanya Berger-Wolf
The Ohio State University
Columbus, Ohio, USA

Haley Griffin
The Computing Community Consortium
Washington D.C., USA

Abstract

Scientific discovery is increasingly constrained not by a lack of data, but by the challenge of turning massive, diverse, and constantly evolving datasets into knowledge that is meaningful and actionable. While advances across artificial intelligence (AI), computing and data sciences have created exciting new modes for scientific inquiry, these systems alone remain insufficient because they struggle with the contextual interpretation, local knowledge, or ethical judgment that complex scientific problems require. Our work proposes that the convergence of citizen science methods with AI, distributed computing and data sciences creates a new form of collective intelligence grounded in human-machine teaming, in which human communities and computational systems are intentionally designed to complement, correct, and extend one another. Rather than viewing citizen science merely as crowdsourced data collection or computational tools as autonomous infrastructure, we advocate for human-machine scientific systems in which citizen scientists, including distributed publics, participate as monitors, interpreters, decision-makers, and co-producers of knowledge. Through reimagining how scientific work is organized and interpreted, this contribution speaks to the CI 2026 theme of Connections by showing how scientific discovery can be redesigned as a connected process of human-machine collaboration across disciplines, communities, and computational systems.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CI/HCOMP Conference'26, Alexandria, VA, USA
© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.

CCS Concepts

• **Human-centered computing** → **Collaborative and social computing theory, concepts and paradigms**; • **Information systems** → *Crowdsourcing*; • **Computing methodologies** → *Artificial intelligence*; • **Social and professional topics** → Socio-technical systems.

Keywords

Citizen science, collective intelligence, human-machine teaming, artificial intelligence, scientific discovery, socio-technical systems.

1 Introduction

Scientific knowledge is increasingly produced through systems that combine human judgment, computational capabilities, and diverse forms of participation rather than through individual experts working in isolation [30, 43, 65]. Recent advances in computing and data science have expanded the scale and speed of scientific work by enabling pattern recognition, environmental simulation, data processing, and hypothesis generation [10, 25, 27, 28, 64, 72]. At the same time, citizen science has demonstrated that members of the public can make meaningful contributions to scientific discovery [6, 37, 59, 61]. Platforms such as Foldit, iNaturalist, and Zooniverse show how citizen scientists can contribute to domains ranging from protein folding and biodiversity monitoring to astrophysics, environmental research, and disaster response [4, 14, 32, 56, 60, 62, 67, 73]. Similarly, health-focused initiatives such as PatientsLikeMe and Cochrane Crowd demonstrate how public participation can generate valuable insights for disease surveillance and population health research [8, 44, 66].

However, despite their complementary strengths, advances in computing and citizen science have largely evolved as separate research agendas [23]. Computational systems are often developed to automate scientific processes, while citizen science initiatives are frequently designed as mechanisms for distributed data collection [12, 37]. This separation represents a missed opportunity. Across

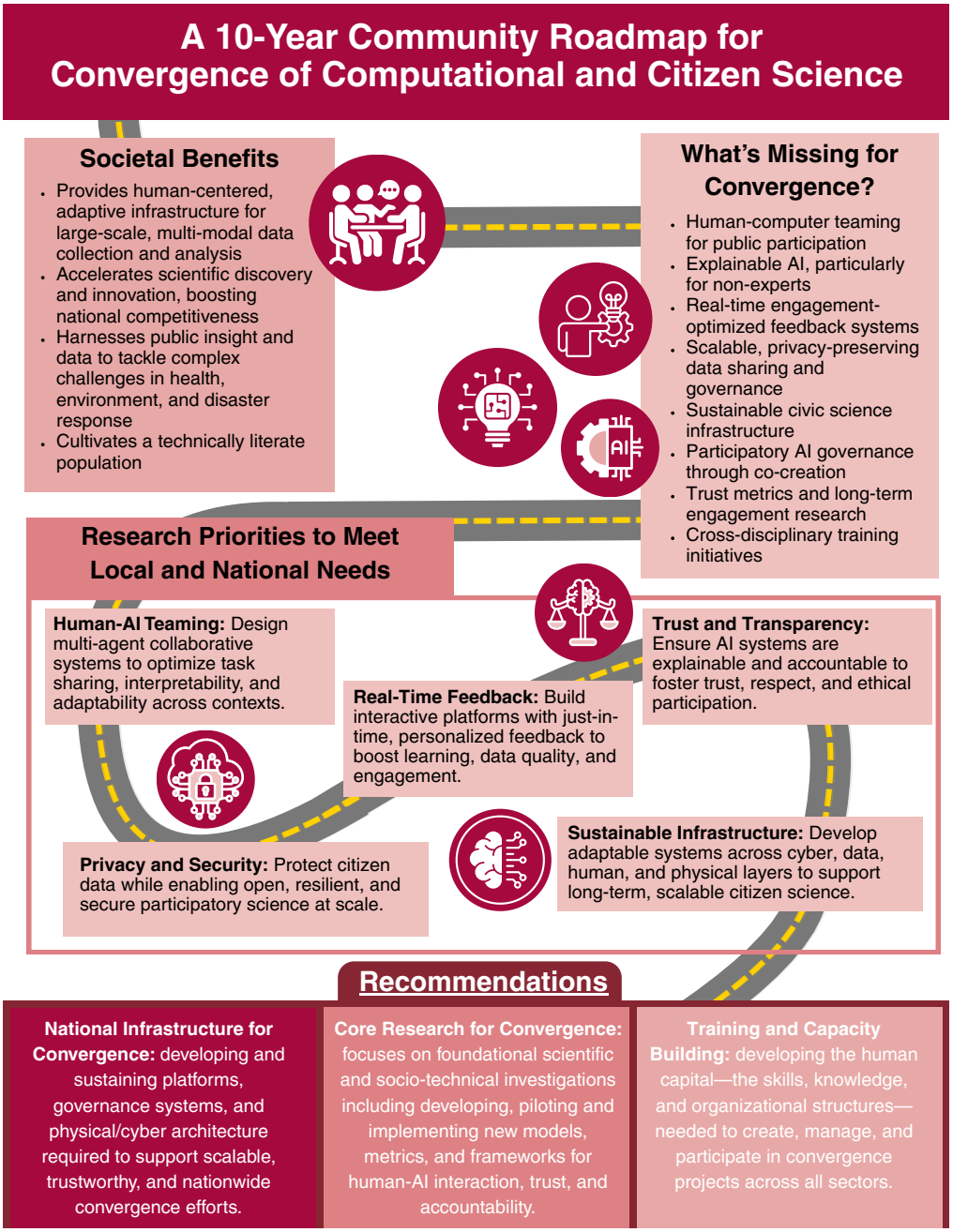

Figure 1: Overview of our roadmap for the convergence of computational systems, citizen science for collective intelligence taken from a recent visioning workshop of the Computing Research Association's Computing Community Consortium [23].

scientific domains, researchers increasingly face challenges whose scale, complexity, and interdisciplinarity exceed the capabilities of any individual expert, institution, or computational system acting alone [29**?**]. Addressing these challenges requires new ways of organizing scientific work that intentionally and ethically combines the strengths of computational systems, scientific institutions, and diverse publics [11, 21, 40, 49, 70].

This paper draws on recent reports and research papers published by the Computing Community Consortium and other researchers [22, 23] and proposes that the convergence of computing, data science and citizen science creates the foundation for a new generation of collective intelligence systems for scientific discovery. We build this argument around different interconnected concepts. The first and central concept is citizen science. Citizen science

refers to research systems in which members of the public contribute to scientific inquiry by collecting observations, monitoring environments, classifying data, interpreting findings, or helping generate new knowledge [6, 19, 61]. Citizen science is particularly valuable when scientific problems require broad geographic coverage, sustained participation, local expertise, or forms of lived experience that professional scientists alone cannot easily access [33, 55, 61]. Importantly, citizen science increasingly has within it "distributed publics": diverse communities of people outside formal scientific institutions who possess relevant knowledge, observations, experiences, and judgment [1, 3, 34]. These contributors may participate through organized citizen science projects or through existing community practices and networks [6, 9, 55]. Examples include residents documenting air quality, patients sharing health experiences, farmers monitoring agricultural conditions, and fishermen observing environmental change [9, 13, 16]. Recognizing distributed publics as an important component of citizen science expands both who can participate in science and what forms of knowledge can contribute to discovery [34, 46]. The second concept is collective intelligence [39, 68]. While citizen science focuses on public participation in research [6, 61], collective intelligence provides a framework for understanding how diverse actors combine their capabilities to solve complex problems [39, 68, 69]. Collective intelligence refers to systems in which people, computational tools, organizations, and institutions coordinate their complementary strengths to produce outcomes that no participant could achieve independently [15, 38]. This perspective shifts attention from simply recruiting volunteers toward designing systems that effectively integrate human judgment, scientific expertise, local knowledge, and computational capabilities [5, 15].

Our central contribution is a framework for the convergence of citizen science, computing, and collective intelligence. We argue that these domains form a shared scientific infrastructure in which citizen scientists contribute expertise, contextual knowledge, and scientific judgment, while computational systems provide automation, coordination, large-scale analysis, and hypothesis generation. Scientific discovery emerges through hybrid human-machine workflows that combine these complementary capabilities throughout the knowledge-production process. Building on this vision, we identify five drivers for enabling human-machine collective intelligence in scientific discovery. Figure 1 presents our related proposed roadmap.

2 The Convergence of Citizen Science, Computing, and Collective Intelligence

Driver 1: Human-Machine Teaming

The foundation of human-machine collective intelligence is the design of workflows that leverage complementary capabilities. Computational systems excel at speed, scale, automation, and pattern recognition, while humans contribute contextual understanding, creativity, ethical reasoning, domain expertise, and local knowledge [17, 48, 63, 71]. Future citizen science systems should therefore focus not on replacing participants with automation but on designing workflows in which humans and machines contribute different forms of expertise throughout the scientific process [21].

Driver 2: Feedback and Interactivity

Collective intelligence depends on continuous learning between participants and systems [2, 45, 53]. Contributors must understand how their work influences scientific outcomes, how information is evaluated, and how decisions are made [51, 55]. Computational systems can strengthen these feedback loops by routing tasks dynamically, identifying uncertainty and disagreement, supporting participant learning, and facilitating collaboration across communities [47]. Effective feedback mechanisms transform citizen science from a collection of isolated contributions into an adaptive collective intelligence system [15].

Driver 3: Trust and Governance

Human-machine collective intelligence requires participants to trust both the computational systems and the institutions that govern them [2, 58]. Participants should thus understand how data are processed, how decisions are made, and how scientific conclusions are generated [37, 50]. Equally important, contributors should have meaningful opportunities to influence platform governance, scientific priorities, data policies, and system design [18, 26, 31]. Collective intelligence systems are strongest when participants are contributors and stakeholders in the infrastructure in which they operate.

Driver 4: Privacy and Security

Large-scale scientific participation creates privacy and security challenges, particularly when environmental, geographic, social, or health data are involved [20, 50, 52, 54]. These challenges are likely to intensify as citizen-generated data become increasingly integrated with AI systems and distributed scientific computing infrastructures [24]. Future infrastructures must support secure participation, privacy-preserving data sharing, and protection against adversarial manipulation. Approaches such as federated analysis, differential privacy, secure aggregation, and confidential computing may enable participants and institutions to contribute to shared models without requiring centralized access to raw data [35, 36, 41]. Privacy and security are essential for sustaining public participation in human-machine scientific workflows.

Driver 5: Sustainable Infrastructure

The convergence of citizen science and computing requires long-term socio-technical infrastructure [13, 42]. Yet citizen science continues to operate across a fragmented landscape of platforms, data repositories, technical standards, and project-specific tools [57, 61]. Interoperability is particularly important as citizen-generated observations become inputs to machine learning, integrated data systems, and emerging scientific digital twins [7, 61]. Sustainable collective intelligence systems require not only technical platforms but also training programs, governance structures, incentives for participation, and mechanisms for recognizing public contributions [6, 55]. Long-term investment in these infrastructures is necessary if human-machine collective intelligence is to become a durable component of scientific practice [30**?**].

2.1 Conclusion

Taken together, the drivers define a comprehensive research agenda for the convergence of citizen science with advanced computing and AI. Research carried out through the five drivers will lead to a socio-technical ecosystem where citizen participation and computational systems reinforce each other. Success will depend not only on technical innovation, but also on responsible governance, sustained infrastructure, and continuous human engagement.

References

[1] Arun Agrawal. 1995. Dismantling the divide between indigenous and scientific knowledge. *Development and change* 26, 3 (1995), 413–439.

[2] Zeynep Akata, Dan Balliet, Maarten de Rijke, Frank Dignum, Virginia Dignum, Guszti Eiben, Antske Fokkens, Davide Grossi, Koen V. Hindriks, Hoosier H. Huang, Catholijn M. Jonker, Luis Moniz Pereira, Myrthe Tielman, Frank van Harmelen, Han van Hoof, Birna van Riemsdijk, Aimee van Wynsberghe, Rineke Verbrugge, Bart Verheij, and Piek Vossen. 2020. A Research Agenda for Hybrid Intelligence: Augmenting Human Intellect with Collaborative, Adaptive, Responsible, and Explainable Artificial Intelligence. *Computer* 53, 8 (2020), 18–28. doi:10.1109/MC.2020.2996587

[3] Karin Bäckstrand. 2003. Civic science for sustainability: reframing the role of experts, policy-makers and citizens in environmental governance. *Global environmental politics* 3, 4 (2003), 24–41.

[4] Steven P Bamford, Robert C Nichol, Ivan K Baldry, Kate Land, Chris J Lintott, Kevin Schawinski, Anže Slosar, Alexander S Szalay, Daniel Thomas, Mehri Torki, et al. 2009. Galaxy Zoo: the dependence of morphology and colour on environment. *Monthly Notices of the Royal Astronomical Society* 393, 4 (2009), 1324–1352.

[5] Susanne Beck and Marion Poetz. 2024. Artificial intelligence and the changing costs and benefits of engaging in open and collaborative science. *The Journal of Open Commons & Organizing (JOCO)* 2, 2 (2024), 28–31.

[6] Rick Bonney, Caren B Cooper, Janis Dickinson, Steve Kelling, Tina Phillips, Kenneth V Rosenberg, and Jennifer Shirk. 2009. Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience* 59, 11 (2009), 977–984.

[7] Anne Bowser, Caren Cooper, Alex De Sherbinin, Andrea Wiggins, Peter Brenton, Tyng-Ruey Chuang, Elaine Faustman, Mordechai Haklay, and Metis Meloche. 2020. Still in need of norms: the state of the data in citizen science. *Citizen Science: Theory and Practice* 5, 1 (2020).

[8] Jed R Brubaker, Caitlin Lustig, and Gillian R Hayes. 2010. PatientsLikeMe: empowerment and representation in a patient-centered social network. In *CSCW'10; Workshop on research in healthcare: past, present, and future.*

[9] Jeffrey A Burke, Deborah Estrin, Mark Hansen, Andrew Parker, Nithya Ramanathan, Sasank Reddy, and Mani B Srivastava. 2006. Participatory sensing. (2006).

[10] Giuseppe Carleo, Ignacio Cirac, Kyle Cranmer, Laurent Daudet, Maria Schuld, Naftali Tishby, Leslie Vogt-Maranto, and Lenka Zdeborová. 2019. Machine learning and the physical sciences. *Reviews of Modern Physics* 91, 4 (2019), 045002.

[11] Luigi Ceccaroni, James Bibby, Erin Roger, Paul Flemons, Katina Michael, Laura Fagan, and Jessica L. Oliver. 2019. Opportunities and Risks for Citizen Science in the Age of Artificial Intelligence. *Citizen Science: Theory and Practice* (Nov 2019). doi:10.5334/cstp.241

[12] Debora Irene Christine and Mamello Thinyane. 2021. Citizen science as a data-based practice: A consideration of data justice. *Patterns* 2, 4 (2021).

[13] Cathy C Conrad and Krista G Hilchey. 2011. A review of citizen science and community-based environmental monitoring: issues and opportunities. *Environmental monitoring and assessment* 176, 1 (2011), 273–291.

[14] Seth Cooper, Firas Khatib, Adrien Treuille, Janos Barbero, Jeehyung Lee, Michael Beenen, Andrew Leaver-Fay, David Baker, Zoran Popović, and Foldit Players. 2010. Predicting protein structures with a multiplayer online game. *Nature* 466, 7307 (2010), 756–760.

[15] Hao Cui and Taha Yasseri. 2024. AI-enhanced collective intelligence. *Patterns* 5, 11 (2024).

[16] Finn Danielsen, Martin Enghoff, Eyðfinn Magnussen, Tero Mustonen, Anna Degteva, Kia K Hansen, Nette Levermann, Svein D Mathiesen, Øystein Slettemark, C Bieling, et al. 2017. Citizen science tools for engaging local stakeholders and promoting local and traditional knowledge in landscape stewardship. *The science and practice of landscape stewardship* (2017), 80–98.

[17] Dominik Dellermann, Adrian Calma, Nikolaus Lipusch, Thorsten Weber, Sascha Weigel, and Philipp Ebel. 2021. The future of human-AI collaboration: a taxonomy of design knowledge for hybrid intelligence systems. *arXiv preprint arXiv:2105.03354* (2021).

[18] Annelies Duerinckx, Katrin Vohland, Kevin Kaufmann, Anna Tarrant, Karen Masters, and Marisa Ponti. 2024. Co-creating Artificial Intelligence: Designing and Enhancing Democratic AI Solutions Through Citizen Science. *Citizen Science: Theory and Practice* 9, 1 (2024), 43. doi:10.5334/cstp.732

[19] Melissa Eitzel, Jessica Cappadonna, Chris Santos-Lang, Ruth Duerr, Sarah Elizabeth West, Arika Virapongse, Christopher Kyba, Anne Bowser, Caren Cooper, Andrea Sforzi, et al. 2017. Citizen science terminology matters: Exploring key terms. *Citizen science: Theory and practice* 2, 1 (2017), 1–20.

[20] Paul B English, Maxwell J Richardson, and Catalina Garzón-Galvis. 2018. From crowdsourcing to extreme citizen science: participatory research for environmental health. *Annual review of public health* 39 (2018), 335–350.

[21] Teppo Felin, Mari Sako, and Jessica Hullman. 2026. AI and actor-specific decisions. *AI & SOCIETY* (2026), 1–19.

[22] Lucy Fortson, Kevin Crowston, Laure Kloetzer, and Marisa Ponti. 2024. Artificial intelligence and the future of citizen science. *Citizen Science: Theory and Practice* 9, 1 (2024), 32–32.

[23] Lucy Fortson, Lea Shanley, Tanya Berger-Wolf, Kevin Crowston, Corey Jackson, Saiph Savage, and Haley Griffin. 2026. *Grand Challenges for the Convergence of Computational and Citizen Science Research.* CCC Workshop Report. Computing Community Consortium (CCC), Computing Research Association. https://cra.org/ccc/wp-content/uploads/sites/2/2026/02/Grand-Challenges-for-the-Convergence-of-Computational-and-Citizen-Science-Research.pdf.

[24] Dilek Fraisl, Linda See, Steffen Fritz, Mordechai Haklay, and Ian McCallum. 2025. Leveraging the collaborative power of AI and citizen science for sustainable development. *Nature Sustainability* 8, 2 (2025), 125–132.

[25] Shuyu Gan, James Mooney, Pan Hao, Renxiang Wang, Mingyi Hong, Qianwen Wang, and Dongyeop Kang. 2026. Scaling Unverifiable Rewards: A Case Study on Visual Insights. In *Findings of the Association for Computational Linguistics: ACL 2026.* 34537–34569.

[26] Claudia Göbel, Christian Nold, Aleksandra Berditchevskaia, and Mordechai Haklay. 2019. How does citizen science "do" governance? Reflections from the DITOs project. *Citizen Science: Theory and Practice* 4, 1 (2019).

[27] Alon Halevy, Peter Norvig, and Fernando Pereira. 2009. The unreasonable effectiveness of data. *IEEE intelligent systems* 24, 2 (2009), 8–12.

[28] Benjamin Herfort, Sven Lautenbach, João Porto de Albuquerque, Jennings Anderson, and Alexander Zipf. 2021. The evolution of humanitarian mapping within the OpenStreetMap community. *Scientific Reports* 11, 1 (Feb. 2021). doi:10.1038/s41598-021-82404-z

[29] Anthony JG Hey, Stewart Tansley, Kristin Michele Tolle, et al. 2009. *The fourth paradigm: data-intensive scientific discovery.* Vol. 1. Microsoft research Redmond, WA.

[30] Tony Hey, Stewart Tansley, and Kristin M Tolle. 2009. Jim Gray on eScience: a transformed scientific method. *The fourth paradigm* 17 (2009).

[31] Katharina Hölscher, Julia M Wittmayer, V Igno Notermans, Madeleine Cléa Montanari, Antonella Passani, and Annelli Janssen. 2025. Mainstreaming citizen science in policy: Adaptations needed in policy and how to achieve them in five European countries. *Environmental Science & Policy* 171 (2025), 104148.

[32] Henry F. Houskeeper, Isaac S. Rosenthal, Katherine C. Cavanaugh, Camille Pawlak, Laura Trouille, Jarrett E. K. Byrnes, Tom W. Bell, and Kyle C. Cavanaugh. 2022. Automated satellite remote sensing of giant kelp at the Falkland Islands (Islas Malvinas). *PLOS ONE* 17, 1 (Jan. 2022), e0257933. doi:10.1371/journal.pone.0257933

[33] Sarah E. Huebner, Zachary Miller, Elizabeth Deneen, Lucy Fortson, and Craig Packer. 2024. Human Supervision is Key to Achieving Accurate AI-assisted Wildlife Identifications in Camera Trap Images. *Citizen Science: Theory and Practice* 9, 1 (2024), 38. doi:10.5334/cstp.752

[34] Sheila Jasanoff. 2004. Science and citizenship: a new synergy. *Science and public policy* 31, 2 (2004), 90–94.

[35] Wenqi Li, Fausto Milletarì, Daguang Xu, Nicola Rieke, Jonny Hancox, Wentao Zhu, Maximilian Baust, Yan Cheng, Sébastien Ourselin, M Jorge Cardoso, et al. 2019. Privacy-preserving federated brain tumour segmentation. In *International workshop on machine learning in medical imaging.* Springer, 133–141.

[36] Zilinghan Li, Aditya Sinha, Yijiang Li, Kyle Chard, Kibaek Kim, and Ravi Madduri. 2025. Experiences building enterprise-level privacy-preserving federated learning to power ai for science. In *2025 IEEE 7th International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (TPS-ISA).* IEEE, 382–391.

[37] Roman Lukyanenko, Andrea Wiggins, and Holly K Rosser. 2020. Citizen science: An information quality research frontier. *Information Systems Frontiers* 22, 4 (2020), 961–983.

[38] Christoph Lutz, Gemma Newlands, and Mohammad Hossein Jarrahi. 2025. Hybrid Intelligence. In *Handbook of Human-Centered Artificial Intelligence.* Springer, 1–33.

[39] Thomas W Malone, Robert Laubacher, and Chrysanthos Dellarocas. 2009. Harnessing crowds: Mapping the genome of collective intelligence. (2009).

[40] Eva C. McClure, Michael Sievers, Christopher J. Brown, Christina A. Buelow, Ellen M. Ditria, Matthew A. Hayes, Ryan M. Pearson, Vivitskaia J. D. Tulloch, Richard K. F. Unsworth, and Rod M. Connolly. 2020. Artificial Intelligence Meets Citizen Science to Supercharge Ecological Monitoring. *Patterns* 1, 7 (Oct. 2020), 100109. doi:10.1016/j.patter.2020.100109

[41] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. 2017. Communication-efficient learning of deep networks

from decentralized data. In *Artificial intelligence and statistics*. Pmlr, 1273–1282.

[42] Greg Newman, Andrea Wiggins, Alycia Crall, Eric Graham, Sarah Newman, and Kevin Crowston. 2012. The future of citizen science: emerging technologies and shifting paradigms. *Frontiers in Ecology and the Environment* 10, 6 (2012), 298–304.

[43] Michael Nielsen. 2011. *Reinventing discovery: the new era of networked science*. Princeton University Press.

[44] Dooley G. Elliott J. Steele E. Shemilt I. Mavergames C. Wisniewski S. McDonald S. Murano M. Glanville J. Foxlee R. Beecher D. Ware J. & Thomas J. Noel-Storr, A. 2021. An evaluation of Cochrane Crowd found that crowdsourcing produced accurate results in identifying randomized trials. *Journal of clinical epidemiology* (2021), 133. https://doi.org/10.1016/j.jclinepi.2021.01.006

[45] Carsten Østerlund, Kevin Crowston, Corey Brian Jackson, Mahboobeh Harandi, Sara Bahaadini, Scott Coughlin, Aggelos K. Katsaggelos, Shane L. Larson, Neda Rohani, Joshua Smith, Laura Trouille, and Michael Zevin. 2024. Supporting Human and Machine Co-Learning in Citizen Science: Lessons From Gravity Spy. *Citizen Science: Theory and Practice* 9, 1 (2024), 42. doi:10.5334/cstp.738

[46] Deja Perkins, Dillon Mahmoudi, Erica Henry, and Caren Cooper. 2026. Geography of race and income shape spatial data gaps in two national participatory science projects. *People and Nature* 8, 5 (2026), 1485–1498. arXiv:https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1002/pan3.70307 doi:10.1002/pan3.70307

[47] Marisa Ponti and Alena Seredko. 2022. Human-Machine-Learning Integration and Task Allocation in Citizen Science. *Humanities and Social Sciences Communications* 9 (2022), 48. doi:10.1057/s41599-022-01049-z

[48] Evangelos Pournaras. 2020. Collective learning: A 10-year odyssey to human-centered distributed intelligence. In *2020 IEEE International Conference on Autonomic Computing and Self-Organizing Systems (ACSOS)*. IEEE, 205–214.

[49] Janet Rafner, Miroslav Gajdacz, Gitte Kragh, Arthur Hjorth, Anna Gander, Blanka Palfi, Aleksandra Berditchevskiaia, Francois Grey, Kobi Gal, Avi Segal, Mike Wamsley, Joshua Miller, Dominik Dellermann, Mordechai Haklay, Pietro Michelucci, and Jacob Sherson. 2022. Mapping Citizen Science through the Lens of Human-Centered AI. *Human Computation* 9, 1 (Nov. 2022), 66–95. doi:10.15346/hc.v9i1.133

[50] Alison Reiheld and Pamela L Gay. 2019. Coercion, consent, and participation in citizen science. *arXiv preprint arXiv:1907.13061* (2019).

[51] Mark O. Riedl. 2019. Human-Centered Artificial Intelligence and Machine Learning. *Human Behavior and Emerging Technologies* 1, 1 (2019), 33–36. doi:10.1002/hbe2.117

[52] Donna M. Schaeffer and Patrick C. Olson. 2021. Cybersecurity issues in citizen science. In *2021 IEEE International Symposium on Technology and Society (ISTAS)*. 1–1. doi:10.1109/ISTAS52410.2021.9629198

[53] David Schwartz, Dov Te'Eni, and Inbal Yahav. 2023. Reciprocal Human Machine Learning (RHML): Human-AI Collaboration Based on Theories of Dyadic Learning. In *Proceedings of the AAAI Symposium Series*, Vol. 1. AAAI Press.

[54] Ali Sepas, Ali Haider Bangash, Omar Alraoui, Khaled El Emam, and Alaa El-Hussuna. 2022. Algorithms to anonymize structured medical and healthcare data: a systematic review. *Frontiers in Bioinformatics* 2 (2022), 984807.

[55] Jennifer L Shirk, Heidi L Ballard, Candie C Wilderman, Tina Phillips, Andrea Wiggins, Rebecca Jordan, Ellen McCallie, Matthew Minarchek, Bruce V Lewenstein, Marianne E Krasny, et al. 2012. Public participation in scientific research: a framework for deliberate design. *Ecology and society* 17, 2 (2012).

[56] Brooke D Simmons, Chris Lintott, Steven Reece, Campbell Allen, Grant RM Miller, Rebekah Yore, David Jones, Sascha T Ishikawa, Tom Jardine-McNamara, Amy R Boyer, et al. 2022. Disaster, Infrastructure and Participatory Knowledge: The Planetary Response Network. *Citizen Science: Theory and Practice* 7, 1 (2022). https://doi.org/10.5334/cstp.392

[57] Karen Soacha-Godoy, Pranesh Prakash, Ana Álvarez, Adriana Giraldo Tejada, Alexandre López-Borrull, Xavier Salvador, Carlos Rodero, Diego Torres, Ivan Rodero, Mabel Ayure, et al. 2025. Research infrastructures in citizen science: state of knowledge and taxonomic framework as a pathway to sustainability. *Citizen Science: Theory and Practice* 10, 1 (2025), 21–21.

[58] Sonia Sousa, Jose Cravino, Paulo Martins, and David Lamas. 2023. Human-Centered Trust Framework: An HCI Perspective. *arXiv preprint arXiv:2305.03306* (2023). arXiv:2305.03306 [cs.HC] https://arxiv.org/abs/2305.03306

[59] Laura Trouille, Chris J. Lintott, and Lucy F. Fortson. 2019. Citizen science frontiers: Efficiency, engagement, and serendipitous discovery with human–machine systems. *Proceedings of the National Academy of Sciences* 116, 6 (2019), 1902–1909. arXiv:https://www.pnas.org/content/116/6/1902.full.pdf doi:10.1073/pnas.1807190116

[60] Grant Van Horn, Oisin Mac Aodha, Yang Song, Yin Cui, Chen Sun, Alex Shepard, Hartwig Adam, Pietro Perona, and Serge Belongie. 2018. The inaturalist species classification and detection dataset. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 8769–8778.

[61] Katrin Vohland, Anne Land-Zandstra, Luigi Ceccaroni, Rob Lemmens, Josep Perelló, Marisa Ponti, Roeland Samson, and Katherin Wagenknecht. 2021. The science of citizen science evolves. In *The science of citizen science*. Springer, 1–12.

[62] Mike Walmsley, Steven Bamford, Hugh Dickinson, Tobias Géron, Alexander J Gordon, Annette M.N. Ferguson, L. F. Fortson, Sandor J. Kruk, N. E. P. Lines, Chris J. Lintott, Karen L. Masters, Robert G. Mann, James Pearson, Hayley Roberts, Anna M M Scaife, Stefan Schuldt, Brooke D. Simmons, Rebecca J. Smethurst, Josh Speagle, and Kyle W. Willett. 2025. Galaxy Zoo Evo: 1 million human-annotated images of galaxies. https://api.semanticscholar.org/CorpusID:284311883

[63] Mike Walmsley, Micah Bowles, Anna M. M. Scaife, Jason Shingirai Makechemu, Alexander J. Gordon, Annette M. N. Ferguson, Robert G. Mann, James Pearson, Jürgen J. Popp, Jo Bautista, Julia Cáceres, James R. Dawson, and Daniel Contreras-Santos. 2024. Scaling Laws for Galaxy Images. arXiv:2404.02973 [cs.CV] doi:10.48550/arXiv.2404.02973

[64] Hanchen Wang, Tianfan Fu, Yuanqi Du, Wenhao Gao, Kexin Huang, Ziming Liu, Payal Chandak, Shengchao Liu, Peter Van Katwyk, Andreea Deac, Anima Anandkumar, Karianne Bergen, Carla P. Gomes, Shirley Ho, Pushmeet Kohli, Joan Lasenby, Jure Leskovec, Tie-Yan Liu, Arjun Manrai, Debora Marks, Bharath Ramsundar, Le Song, Jimeng Sun, Jian Tang, Petar Veličković, Max Welling, Linfeng Zhang, Connor W. Coley, Yoshua Bengio, and Marinka Zitnik. 2023. Scientific discovery in the age of artificial intelligence. *Nature* 620, 79727972 (Aug. 2023), 47–60. doi:10.1038/s41586-023-06221-2

[65] David Watson and Luciano Floridi. 2018. Crowdsourced science: sociotechnical epistemology in the e-research paradigm. *Synthese* 195, 2 (2018), 741–764.

[66] Paul Wicks, Michael Massagli, Jeana Frost, Catherine Brownstein, Sally Okun, Timothy Vaughan, Richard Bradley, and James Heywood. 2010. Sharing health data for better outcomes on PatientsLikeMe. *Journal of medical Internet research* 12, 2 (2010), e1549.

[67] Marco Willi, Ross T Pitman, Anabelle W Cardoso, Christina Locke, Alexandra Swanson, Amy Boyer, Marten Veldthuis, and Lucy Fortson. 2019. Identifying animal species in camera trap images using deep learning and citizen science. *Methods in Ecology and Evolution* 10, 1 (2019), 80–91.

[68] Anita Williams Woolley, Christopher F Chabris, Alex Pentland, Nada Hashmi, and Thomas W Malone. 2010. Evidence for a collective intelligence factor in the performance of human groups. *science* 330, 6004 (2010), 686–688.

[69] Wenjun Wu, Shaoting Tang, Yi Xu, and Tao Wang. 2026. *Collective Intelligence: From the Perspective of Complex Systems*. Springer Nature Singapore, Singapore, 45–78. doi:10.1007/978-981-95-8037-8_3

[70] Daniel Ben Zaken, Kobi Gal, Guy Shani, Avi Segal, and Darlene Cavalier. 2021. Intelligent Recommendations for Citizen Science. *Proceedings of the AAAI Conference on Artificial Intelligence* 35, 1717 (May 2021), 14693–14701. doi:10.1609/aaai.v35i17.17726

[71] M. Zevin, S. Coughlin, S. Bahaadini, E. Besler, N. Rohani, S. Allen, M. Cabero, K. Crowston, A. K. Katsaggelos, S. L. Larson, T. K. Lee, C. Lintott, T. B. Littenberg, A. Lundgren, C. Østerlund, J. R. Smith, L. Trouille, and V. Kalogera. 2017. Gravity Spy: Integrating Advanced LIGO Detector Characterization, Machine Learning, and Citizen Science. *Classical and Quantum Gravity* 34, 6 (2017), 064003. doi:10.1088/1361-6382/aa5cea

[72] Ying Zhou, Yintao Zhang, Zhichao Zhang, Zhimeng Zhou, and Feng Zhu. 2024. AI comes to the Nobel Prize and drug discovery. 101160 pages.

[73] Davide Zilli, Oliver Parson, Geoff V Merrett, and Alex Rogers. 2014. A hidden Markov model-based acoustic cicada detector for crowdsourced smartphone biodiversity monitoring. *Journal of Artificial Intelligence Research* 51 (2014), 805–827.