Electronic Supplementary Material (ESI) for Digital Discovery.

Supplementary Information - Autonomous laboratories for accelerated materials discovery: a community survey and practical insights

1. Methodology

Survey design

We began our survey design with the core research question: “Where might automation be used to accelerate the discovery process in research labs that work on materials discovery?” The survey was intended to (1) identify pain points in the research process that (a) consume a lot of time (b) cause researcher unhappiness (c) are unsafe (d) are boring and monotonous or (e) are difficult to reproduce; and (2) solicit researcher feedback on the value of automation to accelerate the materials discovery process.

As part of designing the survey, we conducted in-person and virtual interviews with researchers experienced in lab automation and autonomy (Helge Stein, Shijing Sun, and Santosh Suram), researchers from labs outside the autonomy space interested in accelerating materials discovery (including a site visit to the Jaramillo lab at Stanford), and researchers from the Robotics Division at Toyota Research Institute (TRI). We gathered ideas from a brainstorming session with the Toyota materials researchers, including the Energy & Materials Division at TRI and materials informatics leadership at Toyota Research Institute of North America (TRINA).

The resulting survey is detailed in sections 2 and 3 below and presented a mix of multiple choice (questions with several options to choose from) and short answer questions.

Survey deployment and response rate

The survey was deployed on Qualtrics and was open to responses from March 23, 2023 to April 2, 2023. Links were sent to research groups actively collaborating with TRI, as well as advertised on LinkedIn, Twitter/X, and materials-related Slack workspaces. Given the methods of publicizing the survey, there may be some bias in the survey responses, with respondents more likely to be familiar with automation, autonomy, AI, and ML techniques for materials compared to the wider materials science and engineering community.

The survey received 100+ responses with response rate dropping for later questions and sections (see table below and questions described in SI Section 3). The respondents reported a variety of research roles and levels of experience.

	Experimental responses	Computational responses	Experimental + Computational
Through #15 “Assessing success”	65	37	102
Through #20 “Understanding the opportunity”	57	N/A	N/A
Through #26 Completed Part I	57	34	91
Through #44 Completed Part II	27	N/A	N/A

Anonymized responses are provided in TRI_lab_automation_survey_2023_responses.xlsx.

2. Survey introductory text

The purpose of this study is to collect information from individuals in the materials science community related to their use of automation in experimental research laboratories. The overarching goal of the project is to identify where automation opportunities exist in the materials discovery community and to recommend places where targeted infrastructure development related to automation may accelerate discovery.

This survey is split into two parts. The first part is designed to take approximately 10 minutes, and the second (optional) part an additional 5–10 minutes.

You are welcome to access the survey link on the same device to start, pause, or resume at any time.

In this survey, we use 'automation' to refer to robotic, experimental, and hardware automation; we are not focused on data automation.

Definitions:

Automation refers to the use of technology, machines, or systems to perform tasks or processes without the need for human intervention with the goal of increasing efficiency,

productivity, and/or accuracy by reducing the amount of manual labor required for a particular task.

Workflow is a series of steps or systematic tasks that are taken to complete a research project. A workflow may begin with project planning and hypothesis development and include all the tasks that lead up to a presentation and publication of research findings.

3. Survey questions

[Note that survey section headings and question numbering have been added for clarity in this SI and were not present in the survey itself. Question IDs in TRI_lab_automation_survey_2023_responses.xlsx are different and correspond to the IDs in the Qualtrics deployment. Responses from the "Future involvement" section (Q25-26) were decoupled from other survey responses by IF researchers before analysis by TRI researchers. Responses to questions in the "Future involvement" and "Concluding questions" sections (Q25-26 and Q45-47) are not included in this SI. Responses indicating the number of years since bachelors and number of years in the current research group have also been excluded, as have short responses clarifying the choice of "Other" for respondent role and facility.]

Demographics / respondent's role

1. What most closely describes your role at work?
    - Choose one from the options: undergraduate student, graduate student, postdoctoral researcher, non-PI scientist, PI, other
    - Respondents choosing "PI" received instructions *"For the rest of this survey, please answer work-related questions based on the work your entire research group does (rather than your personal day-to-day work)."*
2. How many years have you been in your current research group?
3. How many years have elapsed since you received your bachelor's degree?
4. What best describes your research work?
    - Choose one from the options: mostly experimental, mostly computational/theoretical, mix of computational/theoretical and experimental, other
5. Which best describes your facility?
    - Choose one from the options: university or academic research institution, national or international government-funded laboratory, private sector, other
6. Please check one or more:
    - I consider myself an expert in automation in the lab
    - I'm really interested in the possibilities generated using automation in the lab
    - I use automation (e.g., robotic arms or autosampler) that is already available in the lab
    - I don't have much experience with automation in the lab
    - I am not currently interested in using automation in the lab

Lab and research time (only shown to respondents whose research work is “mostly experimental” or “mix of computational/theoretical and experimental” in Q1)

7. How many hours a week do you typically spend in the lab?
8. How many hours a week do you typically spend working outside the lab?
9. What would be more helpful to you this month:
    - Automating one of your lab tasks
    - Automating one of your data analysis steps
10. Which is a bigger factor in limiting your research progress this month:
    - Getting access to the equipment you need (e.g., bench space, instrument time)
    - Time spent on set-up tasks, clean-up tasks, or troubleshooting
11. Think about your typical laboratory time this month. Please estimate what the relative contributions of each category are to your lab work. [presented as percentages]
    - Set-up and clean-up
    - Troubleshooting
    - Performing research

Assessing success

12. Which of the potential benefits would most motivate you to invest in automation? Please drag to rank order.
    - Day to day time savings/efficiency
    - Researcher happiness
    - Researcher safety
    - Enablement of new processes
    - Reproducibility (e.g., from lab to lab)
    - Generation of larger datasets
    - Other _____
13. **Flexibility** refers to the ability to modify an automation workflow to test various different research hypotheses. **Throughput** refers to the ability of an automation workflow to test one hypothesis over many samples and iterations. Which is more important for your research?
14. When thinking about automation, what is more valuable to you?
    - To do things that are completely impossible for a human to currently do (e.g., imagine you had four arms).
    - To get rid of dull, dirty, or dangerous lab tasks.
15. Can you provide an example?

Understanding the opportunity (only shown to respondents whose research work is “mostly experimental” or “mix of computational/theoretical and experimental” in Q1)

16. What is the rate-limiting or bottleneck step in your current workflow? *For example, sample preparation for TEM or substrate preparation for electrochemistry.*
17. If you could only automate (or better automate) one step in your current workflow, is it the rate limiting step? If not, what would it be and why? *For example, domain expertise is needed for quality control; data analysis.*
18. What is the primary bottleneck in your current research that cannot be automated? *For example, generating scientific hypotheses; I don't think robots can match my quality.*
19. How much more quickly could you complete your current research project if your workflow was automated?
20. What would you NOT want automated? Why? *For example, I enjoy cleaning glassware!*

Researcher sentiment

21. To what extent do you think increasing automation in your current workflow will change the nature of your work?
    - Choose one from the options: completely, a lot, somewhat, a little, not at all
    - Respondents whose research work is “mostly computational/theoretical” or “other” were shown an alternate wording: “To what extent do you think increasing automation in the field’s experimental workflows will change the nature of your work?”
22. What is your general outlook about the future use of automation in your research or field?
    - Choose one from the options: negative, somewhat negative, neutral, somewhat positive, positive
23. What are some reasons why you may have negative feelings? *Please check all that apply.*
    - I have no negative feelings
    - There are tasks I really enjoy doing that I don’t want taken over by an automated process
    - It would be technically challenging to implement and maintain effective automated processes
    - I wouldn’t trust an automated process to do the job effectively
    - Automation lessens the ability for scientific insight and/or research creativity
    - It may decrease the value of the research field
    - Other _____
24. How do you think automation is likely to change job satisfaction and general happiness in your day-to-day work?
    - Choose one from the options: significantly decrease, somewhat decrease, neither increase nor decrease, somewhat increase, significantly increase

Future involvement

25. Would you be willing to be involved in further study of the role of automation in your field? *For example, taking part in a 30 minute interview and/or keeping a research time diary.*
26. Would you be willing to name your lab group and institution? If yes, please write your research institution in the blank field. *Note that even if you reply yes, identifying information will only be shared internally with the survey research team and will not be shared publicly. Any information shared publicly will be aggregated and anonymized.*

Survey optional part 2 preamble

27. Reminder: you are welcome to access the survey link on the same device to start, pause, or resume at any time.
    In this section we ask for more details about your experimental workflow and or the procedures you use during lab work. Think about all of the laboratory work you do, from set up to data synthesis. You can also provide general feedback at the end of this section.
    Are you willing to answer additional questions? (estimated additional 10 mins)
    - Choose one from the options: yes, No (submit your responses and end the survey)
28. Is your research work primarily computational/theoretical?
    - Choose one from the options: yes, no

Experimental research (only shown if respondents reply “yes” for Q27 and are NOT “primarily computational/ theoretical” in Q28)

Safety and lab space

29. Which safety concerns are you exposed to over the course of a typical month? *Check all that apply:*
    - My role presents no safety concerns
    - Reactive
    - Corrosive
    - Flammable
    - Pyrophoric
    - Radioactive
    - Toxicity
    - Carcinogen
    - Physical: sharps
    - Physical: heavy items
    - Other _____
30. To what extent do safety concerns bottleneck your research progress?

- Choose one from the options: completely, a lot, somewhat, a little, not at all
- Question not shown if “My role presents no safety concerns” checked in previous question

31. If the safety element (*e.g., human exposure to a toxin*) was able to be isolated and fully automated, how much of a positive effect would it have on your ability to advance your research activities?
    - Choose one from the options: significantly decrease, somewhat decrease, neither increase or decrease, somewhat increase, significantly increase
32. How difficult would it be to change the existing space to enable increased automation?
    - Choose one from the options: A lot of effort (*e.g., major lab renovation or entirely new space needed*), Some effort (*e.g., minor lab renovation*), Little effort (*e.g., minimum renovation and equipment movement*), No effort (*automation can be installed without reconfiguration*)

Workflow categories

33. Which of the following categories fall into your regular (monthly basis) workflow? *Please choose all that apply.*
    - **General lab tasks** (*e.g., part of day-to-day work but not necessarily workflow*)
    - **Research planning** (*e.g., literature review, methodology research, hypothesis development*)
    - **Experimental planning and design** (*e.g., any preparatory steps, supplies and materials planning and acquisition*)
    - **Experimental set-up** (*e.g., preparing solutions, setting up equipment, organizing fume hood*)
    - **Materials synthesis** (*e.g., collecting and manipulating samples in glove box, performing chemical reactions, heating, cooling, stirring, furnace combustion)*
    - **Materials characterization** *(e.g., sample preparation, NMR, XRD, chromatography, spectroscopy, mass spectrometry, etc.*)
    - **Materials processing** (*e.g., treatment of materials into necessary form such as a film or powder*)
    - **Materials property measurements** (*e.g., understanding qualitative or quantitative properties of a material*)
    - **Device assembly** (*e.g., assembly or construction of a battery, probe, film or other device*)
    - **Performance testing of devices** (*e.g., device cycling tests*)
    - **Data analysis, interpretation, and management** (*e.g., downloading raw data from instrumentation, cleaning and collating data, converting data from signal intensity into units that can be interpreted by the researcher, running statistical analysis, analyzing spectra*)
    - **Other _____**
34. Which categories take up the most of your time? *Please choose up to three.*
    - Possible responses are the subset of categories chosen by the respondent in question Q33

35. Which categories do you spend the most time troubleshooting? *Please choose up to three.*
    - Possible responses are the subset of categories chosen by the respondent in question Q33
36. Which steps do other people execute (in other words, which, if any, of these steps involve some reliance on someone who is not you)? *Please choose all that apply.*
    - Possible responses are all categories in question Q33
37. Which categories of your workflow already contain some form of automation? *Please choose all that apply.*
    - Possible responses are the subset of categories chosen by the respondent in question Q33
38. Which categories of your workflow do you think have the most potential for automation? *Please choose all that apply.*
    - Possible responses are the subset of categories chosen by the respondent in question Q33

Workflow details

39. Which instruments do you primarily use? *For example, NMR, SEM, AFM, TEM, XRD.*
    - Question only shown to respondents who included “Materials characterization” as part of their regular workflow in Q33
40. What primary synthesis techniques do you use? *For example, solid-state, solution-processing.*
    - Question only shown to respondents who included “Materials synthesis” as part of their regular workflow in Q33
41. What primary processing techniques do you use? *For example, heat/gas treatment, thin-film fabrication.*
    - Question only shown to respondents who included “Materials processing” as part of their regular workflow in Q33
42. What primary property measurement techniques do you use? *For example, indentation, porosity measurements.*
    - Question only shown to respondents who included “Materials property measurement” as part of their regular workflow in Q33
43. What primary device assembly techniques do you use? *For example, battery fabrication.*
    - Question only shown to respondents who included “Device assembly” as part of their regular workflow in Q33
44. What primary device performance measurements do you use? *For example, battery cycling test.*
    - Question only shown to respondents who included “Performance testing of devices” as part of their regular workflow in Q33

Concluding questions (only shown if respondents reply “yes” for Q27)

45. Are there any questions that we should have asked you, but did not? If so, please write them below.
46. Are there any examples of people, labs, or organizations that represent an effective or innovative use of automation that you would recommend we research?
47. Do you have any other questions or comments you would like to share?

4. Visual summary of survey results

Demographics / respondent’s role

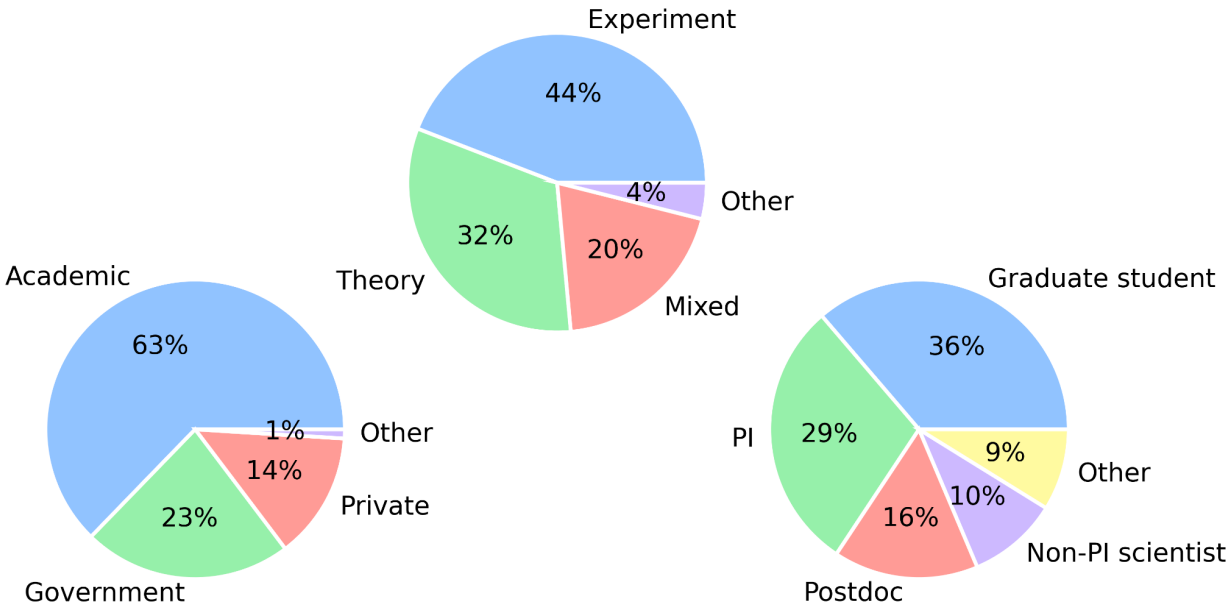

Figure 1: Respondents’ facility, role, and research work. 102 responses to these questions.

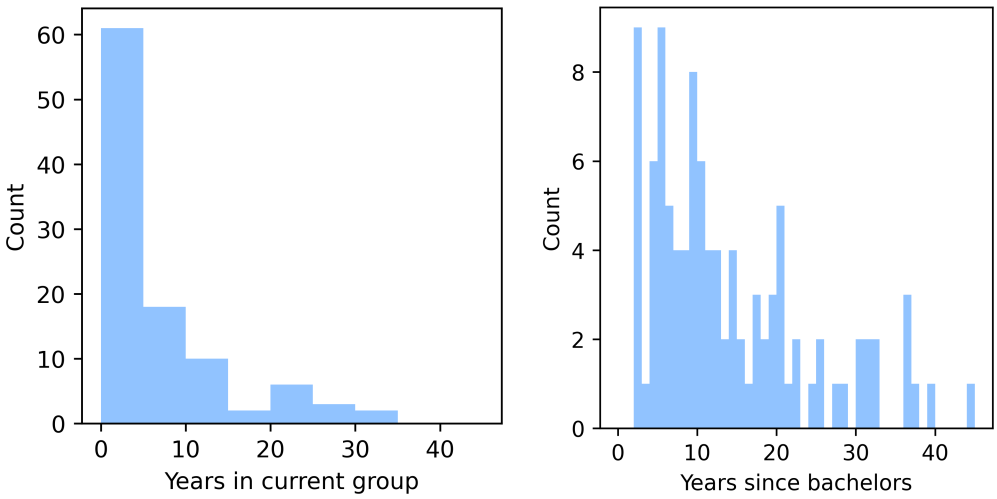

Figure 2: Years of experience. 102 responses to these questions.

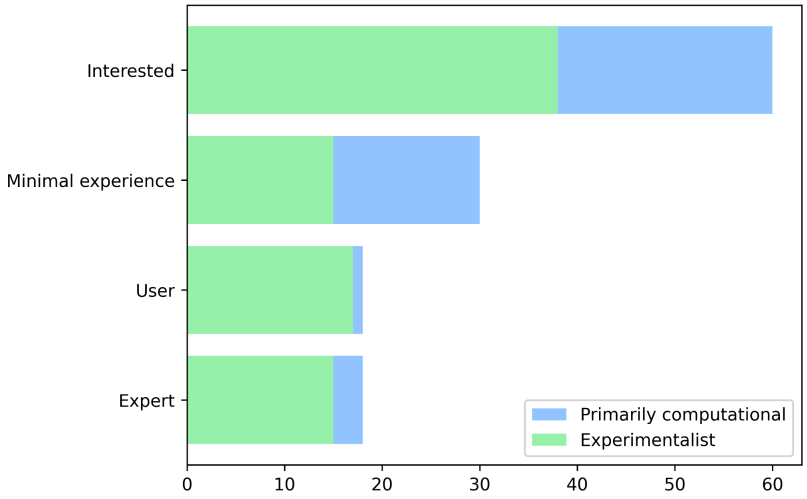

Figure 3: Automation experience and interest. Respondents could choose more than one option. Responses from 65 experimentalists (including primarily experimental and mix of experimental and computational) and 37 primarily computational.

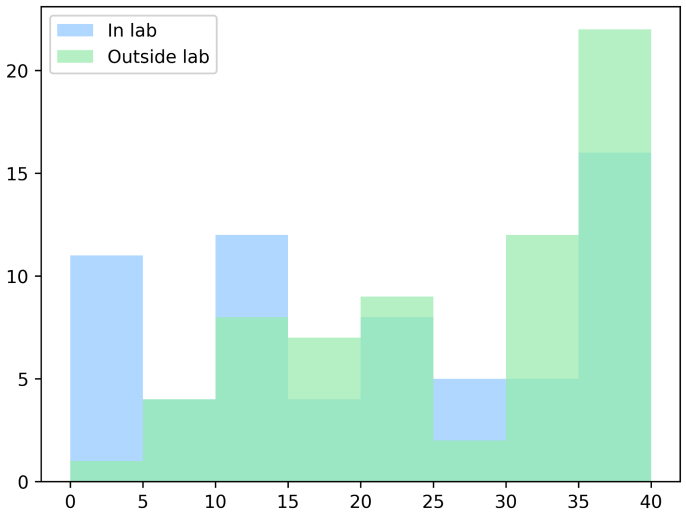

Figure 4: Number of hours per week spent inside or outside the lab. 65 responses to this question.

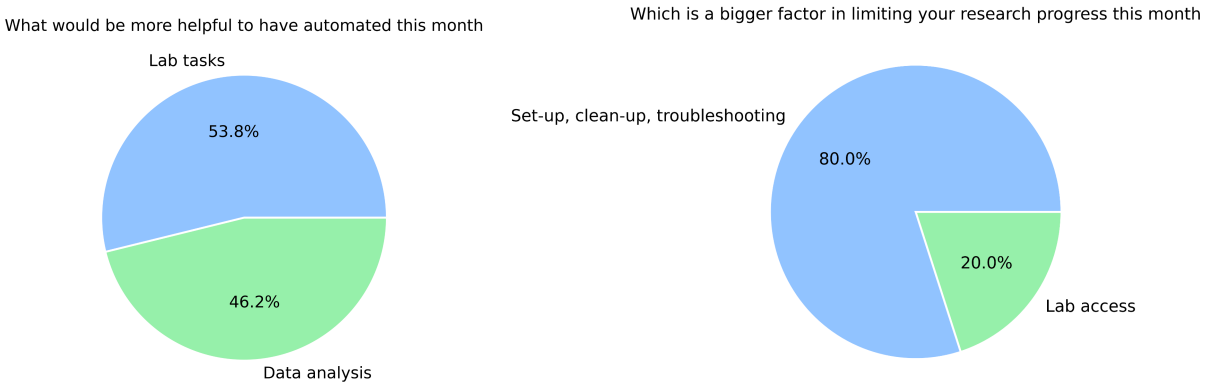

Figure 5: Bottlenecks. 65 responses to these questions.

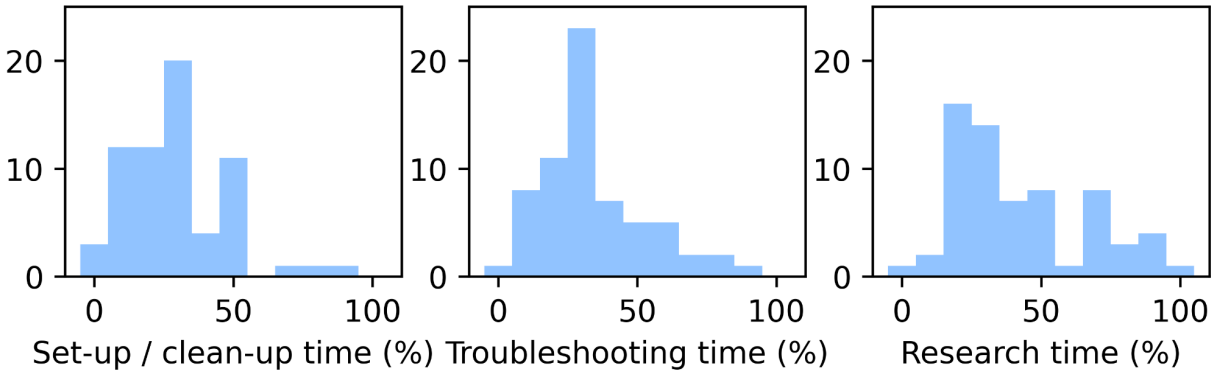

Figure 6: Time spent on different categories of lab work. 65 responses to this question.

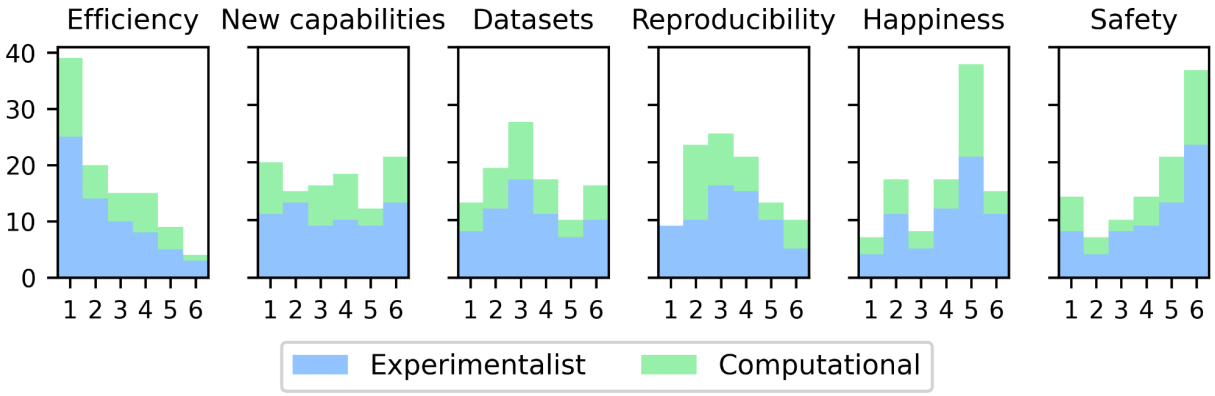

Figure 7: Motivations to automate, split between computational and experimental researchers. 102 responses to this question. Not shown on this figure: three respondents provided Other motivations of “more time to work on the understanding the theory”, “Accelerated R&D and discovery”, and “Generation of self-consistent datasets”.

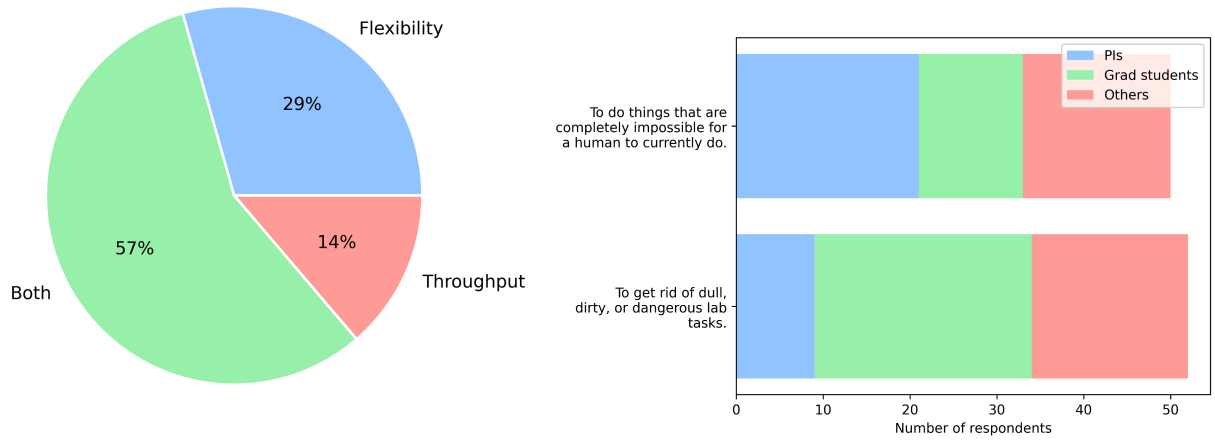

Figure 8: Considerations when automating. 102 responses to these questions.

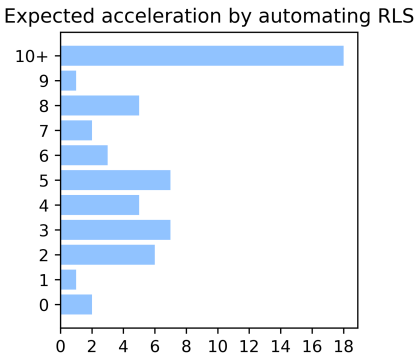

Figure 9: Amount of acceleration if rate limiting step is automated. 57 response to this question.

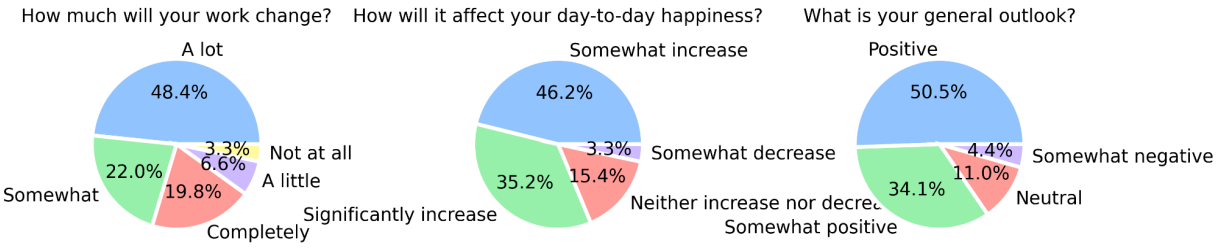

Figure 10: Materials researcher sentiment on automation / autonomy. 91 responses to these questions.

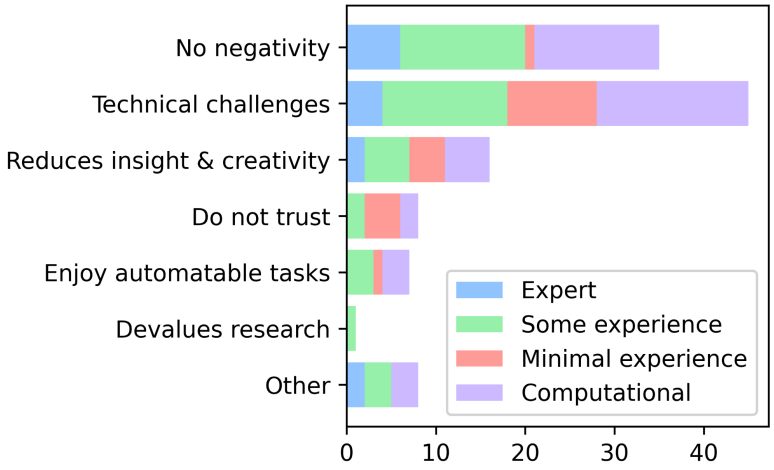

Figure 11: Reasons for negative sentiment. 91 total responses from 13 “Expert”, 31 “Some experience”, 13 “Minimal experience”, and 34 “Computational”.

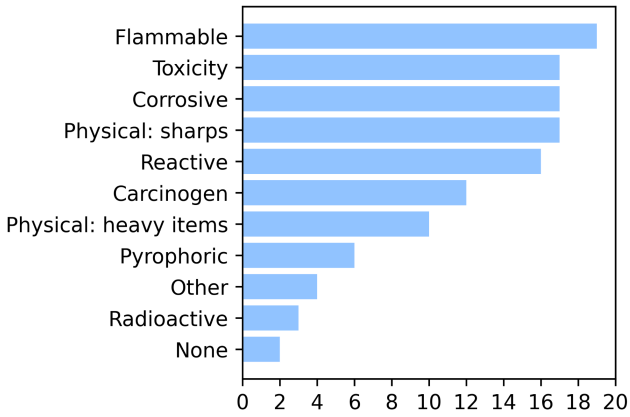

Figure 12: Hazards encountered in a typical month. 27 responses to this question.

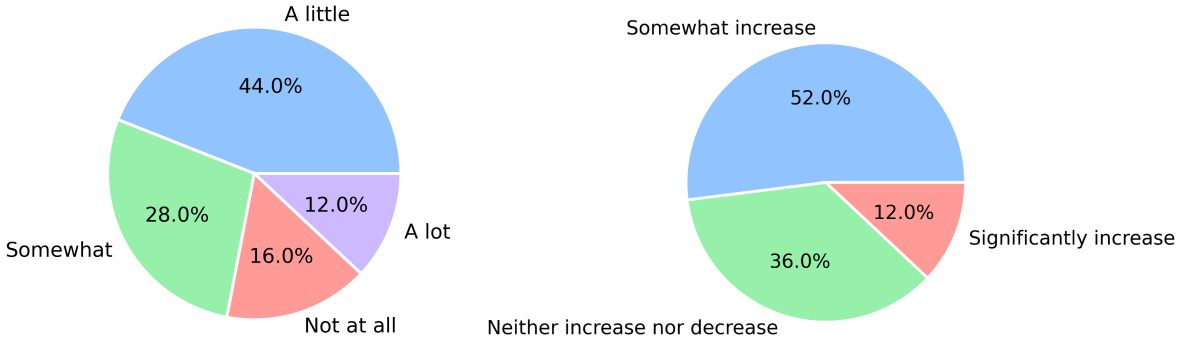

Figure 13: Whether hazards are bottlenecks, and does fully automating hazardous workflows positively advance research? 25 responses to these questions.

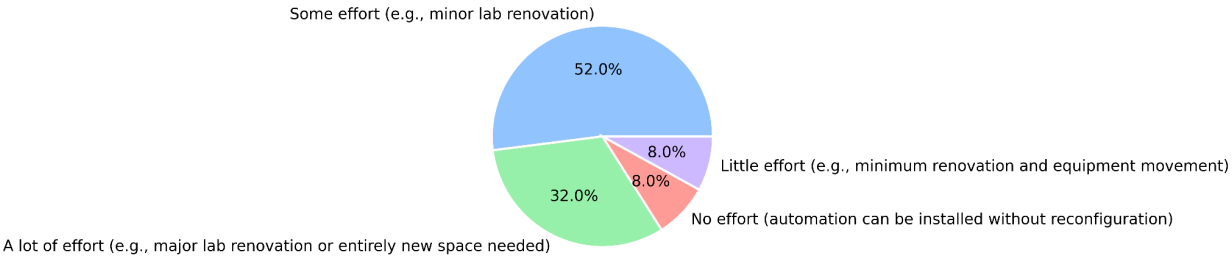

Figure 14: How difficult would it be to change the existing space to enable increased automation? 25 responses to this question.

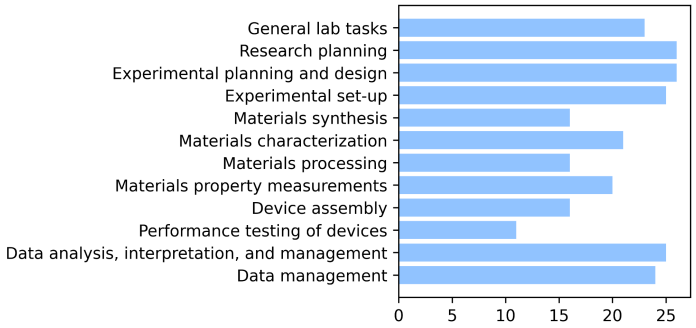

Figure 15: Workflow categories in the regular (monthly) workflow. 27 responses to this question.

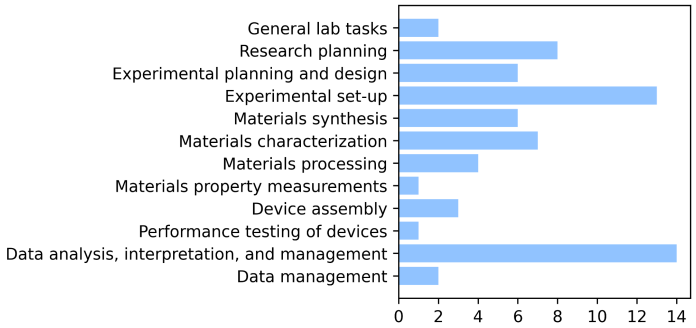

Figure 16: Workflow categories that take the most time. 27 responses to this question.

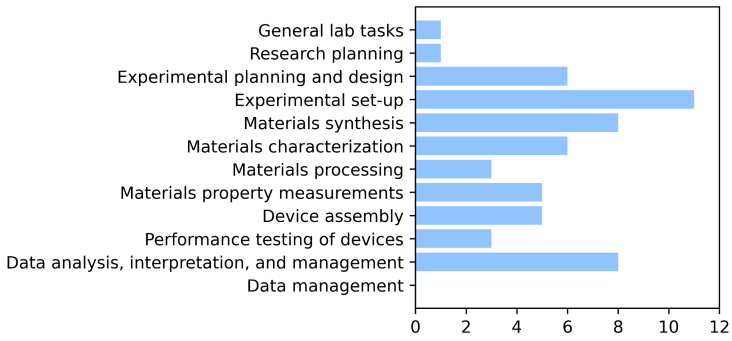

Figure 17: Categories that take the most time to troubleshoot. 27 responses to this question.

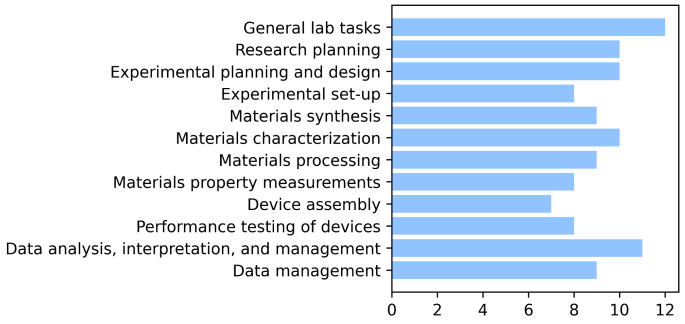

Figure 18: Categories that other people execute for you. 27 responses to this question.

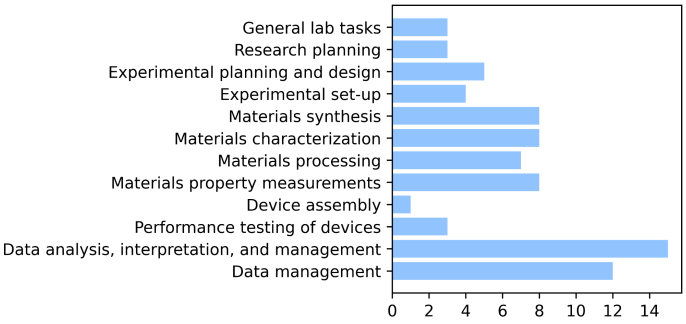

Figure 19: Categories that already have some automation. 27 responses to this question.

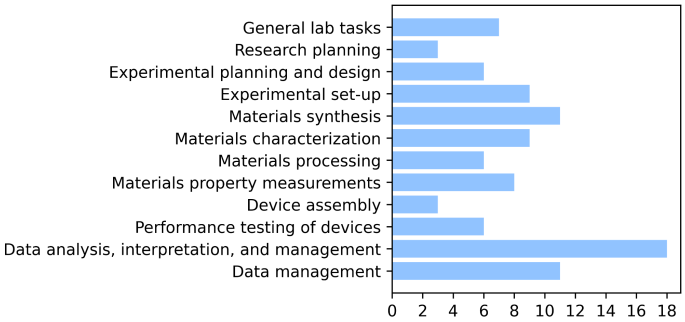

Figure 20: Categories with the most potential for automation. 27 responses to this question.

5. Levels of laboratory autonomy - minimally formatted table

	L0	L1	L2	L3	L4	L5
Process execution	Minimal	Partial, convenient processes	Partial, critical processes	Partial, most processes	Full, except occasional sample transfers	Full
Data analysis	Minimal	Partial	Full	Full	Full	Full
Data interpretation	None	None	Partial, convenient data streams	Partial, critical data streams	Full, simple models	Full, trusted by experts
Decision making	None	None	Partial, machine-assisted	Partial, machine-driven with human input as needed	Full, with human-on-the loop as supervisor	Full, human out-of-the-loop
Communication in workflows	None	None	Partial	Full, single workflow	Full, multiple workflows	Full, orchestrated workflows