



NSF Principal Investigator
Computational Needs Assessment
Codebook
DNA Learning Center
Cold Spring Harbor Laboratory

ID

Assigned respondent ID number

Division

- 1: Biological Infrastructure
- 2: Integrative Organismal Systems
- 3: Environmental Biology
- 4: Molecular and Cellular Biosciences
- 5: Emerging Frontiers

Q1 – Q9: With what type(s) of data do you currently work in your research? *Please select all that apply.*

- Q1: DNA/RNA/protein sequence
 - 0: unselected
 - 1: selected
- Q2: DNA/RNA/protein structure
 - 0: unselected
 - 2: selected
- Q3: Ecological
 - 0: unselected
 - 3: selected
- Q4: Images
 - 0: unselected
 - 4: selected
- Q5: Microscopic
 - 0: unselected
 - 5: selected
- Q6: Pathways/interactions/networks
 - 0: unselected
 - 6: selected
- Q7: Phenotype
 - 0: unselected
 - 7: selected
- Q8: Physiological/medical
 - 0: unselected
 - 8: selected
- Q9: Other
 - 0: unselected
 - Text: written-in

Q10: Do you currently work with large data sets (Big Data)?

1: Yes

2: No

Q11: Do you think you will be working with large data sets sometime in the next three years? *Note – only shown to individuals who answered “NO” in Q10.*

1: Yes

2: No

Q15 – Q26: Consider each of the following data analysis needs and respond in each column accordingly.

The table below represents the matrix of items in this question. Each need was considered three times. For each of the following questions, the coding is as follows:

1: Yes

2: No

3: I don't know

	Is this currently important in your research?	Does your institution meet this need for researchers?	Do you think this will be important to your research in three years?
Updated analysis software	Q15	Q16	Q17
Multi-step analysis workflows or pipelines	Q18	Q19	Q20
High-performance or cluster computing	Q21	Q22	Q23
Cloud computing (remote, configurable, on-demand computers)	Q24	Q25	Q26

Q27 – Q38: Consider each of the following data storage, discovery, and sharing needs and respond in each column accordingly.

The table below represents the matrix of items in this question. Each need was considered three times. For each of the following questions, the coding is as follows:

- 1: Yes
- 2: No
- 3: I don't know

	Is this currently important in your research?	Does your institution meet this need for researchers?	Do you think this will be important to your research in three years?
Sufficient data storage	Q27	Q28	Q29
Search for data and discover relevant data sets	Q30	Q31	Q32
Share data with colleagues	Q33	Q34	Q35
Publish data to the community and/or archives	Q36	Q37	Q38

Q39 – Q53: Consider each of the following computational support and training needs and respond in each column accordingly.

The table below represents the matrix of items in this question. Each need was considered three times. For each of the following questions, the coding is as follows:

- 1: Yes
- 2: No
- 3: I don't know

	Is this currently important in your research?	Does your institution meet this need for researchers?	Do you think this will be important to your research in three years?
Support for bioinformatics and analysis	Q39	Q40	Q41
Training on basic computing (Linux) and scripting (Python, R)	Q42	Q43	Q44
Training on data management and metadata	Q45	Q46	Q47
Training on integration of multiple data types	Q48	Q49	Q50
Training on scaling analysis to cloud or high performance computing	Q51	Q52	Q53

Q54 – Q86: In what academic discipline is your research? *Select all that apply.*

- Q54: Agriculture
 - 0: unselected
 - 1: selected
- Q55: Biochemistry
 - 0: unselected
 - 2: selected
- Q56: Bioengineering
 - 0: unselected
 - 3: selected
- Q57: Biogeography
 - 0: unselected
 - 4: selected
- Q58: Bioinformatics
 - 0: unselected
 - 5: selected
- Q59: Biomechanics
 - 0: unselected
 - 6: selected
- Q60: Biomedical research
 - 0: unselected
 - 7: selected

Q61: Biophysics
0: unselected
8: selected

Q62: Biotechnology
0: unselected
9: selected

Q63: Botany
0: unselected
10: selected

Q64: Cell biology
0: unselected
11: selected

Q65: Cognitive biology
0: unselected
12: selected

Q66: Conservation biology
0: unselected
13: selected

Q67: Developmental biology
0: unselected
14: selected

Q68: Ecology
0: unselected
15: selected

Q69: Environmental biology
0: unselected
16: selected

Q70: Epidemiology
0: unselected
17: selected

Q71: Evolutionary biology
0: unselected
18: selected

Q72: Genetics
0: unselected
19: selected

Q73: Integrative biology
0: unselected
20: selected

Q74: Marine biology
0: unselected
21: selected

Q75: Medicine
0: unselected
22: selected

Q76: Microbiology
0: unselected
23: selected

Q77: Molecular biology
0: unselected
24: selected
Q78: Neurobiology
0: unselected
25: selected
Q79: Population biology
0: unselected
26: selected
Q80: Paleontology
0: unselected
27: selected
Q81: Pathology
0: unselected
28: selected
Q82: Physiology
0: unselected
29: selected
Q83: Phytopathology
0: unselected
30: selected
Q84: Psychobiology
0: unselected
31: selected
Q85: Structural biology
0: unselected
32: selected
Q86: Zoology
0: unselected
33: selected
Q87: Other
0: unselected
Text: written in

Q88 – Q93: What position(s) do you currently hold? *Select all that apply.*

Q88: Researcher (PI)
0: unselected
1: selected
Q89: Researcher (post-doc)
0: unselected
2: selected
Q90: Industry
0: unselected
3: selected
Q91: Faculty/educator
0: unselected
4: selected
Q92: Student
0: unselected
5: selected

Q93: Other

0: unselected

Text: written in

Q94: How large is your research group?

- 1: Small research project (one PI)
- 2: Small laboratory (fewer than five people)
- 3: Large research project (multiple PIs)
- 4: Large laboratory (greater than five people)
- 5: Individual researcher
- 6: Do not work in a research capacity

Q95: Other (please specify)

Q96: What are your teaching responsibilities?

- 1: One course per term
- 2: Two courses per term
- 3: Three courses per term
- 4: Four courses per term
- 5: I do not have any teaching responsibilities

Q97: Other (please specify)