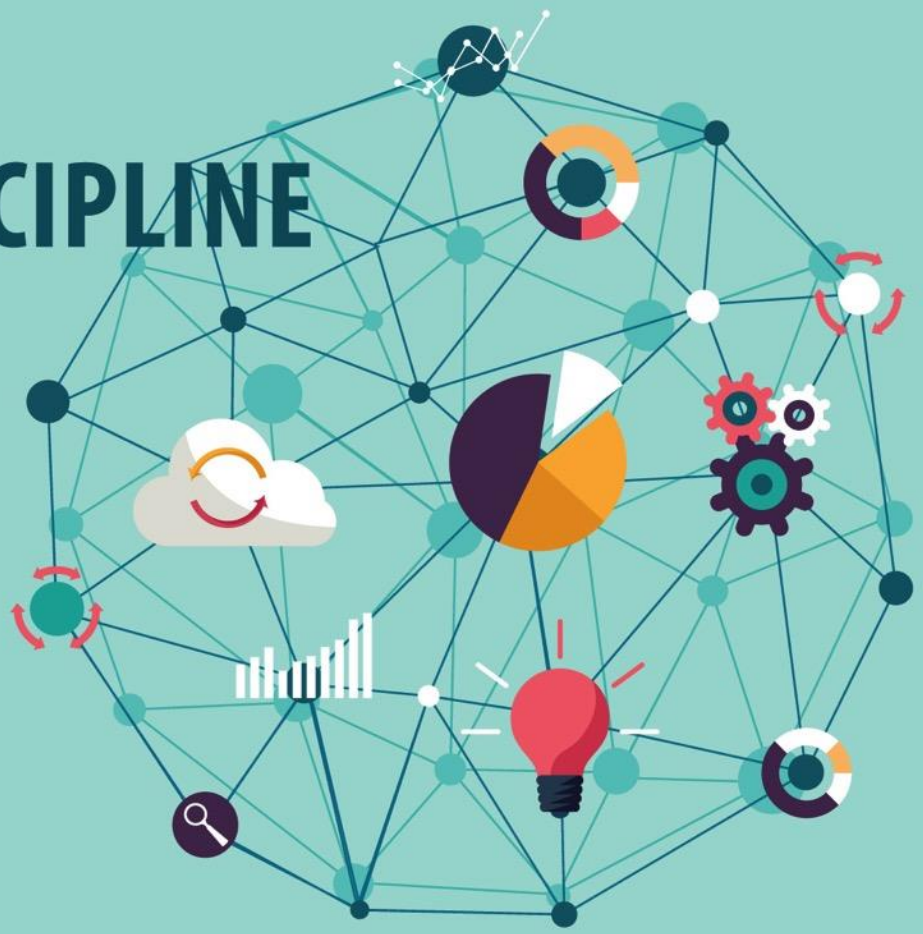
Envisioning the

DATA SCIENCE DISCIPLINE

The Undergraduate Perspective



Webinar Series
Fall 2017

The National Academies of | SCIENCES ENGINEERING MEDICINE

nas.edu/EnvisioningDS

Envisioning the DATA SCIENCE DISCIPLINE

The Undergraduate Perspective

9/12/17 – Building Data Acumen *(recording posted)*

9/19/17 – Incorporating Real-World Applications *(recording posted)*

9/26/17 – Faculty Training and Curriculum Development *(recording posted)*

10/3/17 – Communication Skills and Teamwork *(recording posted)*

10/10/17 – Inter-Departmental Collaboration and Institutional Organization *(recording posted)*

10/17/17 – Ethics *(recording posted)*

10/24/17 – Assessment and Evaluation for Data Science Programs

11/7/17 – Diversity, Inclusion, and Increasing Participation

11/14/17 – Two-Year Colleges and Institutional Partnerships

Provide input, download the interim report, and learn more about the study at www.nas.edu/EnvisioningDS

Envisioning the DATA SCIENCE DISCIPLINE

The Undergraduate Perspective

Assessment and Evaluation

Pamela Bishop, Univ. of Tennessee, Knoxville
Director, National Institute for STEM Evaluation and Research (NISER)
NIMBioS Associate Director for STEM Evaluation

Kari Jordan, Data Carpentry
Deputy Director for Assessment

Provide input and learn more about the study at www.nas.edu/EnvisioningDS

Envisioning the
DATA SCIENCE DISCIPLINE
The Undergraduate Perspective

Assessment and Evaluation

Evaluation of Data Science Programs

Pamela Bishop, Univ. of Tennessee, Knoxville
Director, National Institute for STEM Evaluation and Research (NISER)
NIMBioS Associate Director for STEM Evaluation

What is **PROGRAM EVALUATION?**

PROGRAM EVALUATION IS:

What is **PROGRAM EVALUATION?**

PROGRAM EVALUATION IS:

Systematic collection of **data** about the activities, characteristics, and results of programs to (1) to **make judgments** about the program, (2) **improve** or further develop program effectiveness, (3) **inform** decisions, and/or (4) **increase understanding**.

Michael Quinn Patton

4 Elements of evaluation

1. Systematic process
2. Data collection
3. Enhances knowledge
4. Decision making

Why is evaluation **IMPORTANT**?

PROGRAM EVALUATION

Why is evaluation **IMPORTANT**?

PROGRAM EVALUATION

- ✓ Enhances your project design

Why is evaluation **IMPORTANT**?

PROGRAM EVALUATION

- ✓ Enhances your project design
- ✓ Assists in allocation of your resources and timeline for deliverables

Why is evaluation **IMPORTANT**?

PROGRAM EVALUATION

- ✓ Enhances your project design
- ✓ Assists in allocation of your resources and timeline for deliverables
- ✓ Improves the implementation and effectiveness of projects

Why is evaluation **IMPORTANT**?

PROGRAM EVALUATION

- ✓ Enhances your project design
- ✓ Assists in allocation of your resources and timeline for deliverables
- ✓ Improves the implementation and effectiveness of projects
- ✓ Supports plans for sustainability

Why is evaluation **IMPORTANT**?

PROGRAM EVALUATION

- ✓ Enhances your project design
- ✓ Assists in allocation of your resources and timeline for deliverables
- ✓ Improves the implementation and effectiveness of projects
- ✓ Supports plans for sustainability
- ✓ Provides evidence to support future planning and funding

TYPICAL EVALUATION PROCESS

Provide input and learn more about the study at www.nas.edu/EnvisioningDS

TYPICAL EVALUATION PROCESS

Provide input and learn more about the study at www.nas.edu/EnvisioningDS

TYPICAL EVALUATION PROCESS

Provide input and learn more about the study at www.nas.edu/EnvisioningDS

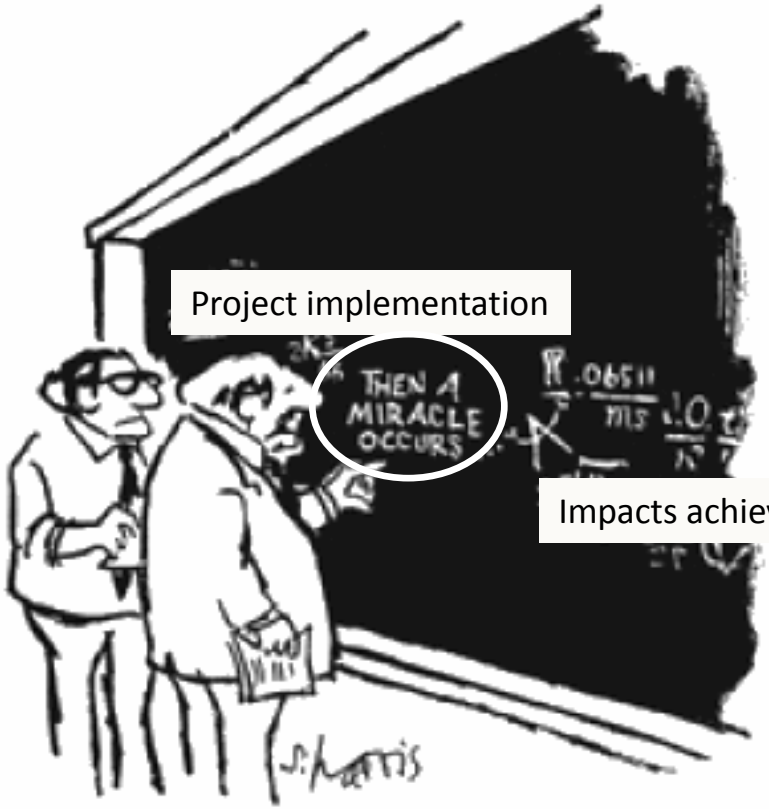
TYPICAL EVALUATION PROCESS

Provide input and learn more about the study at www.nas.edu/EnvisioningDS

TYPICAL EVALUATION PROCESS

Provide input and learn more about the study at www.nas.edu/EnvisioningDS

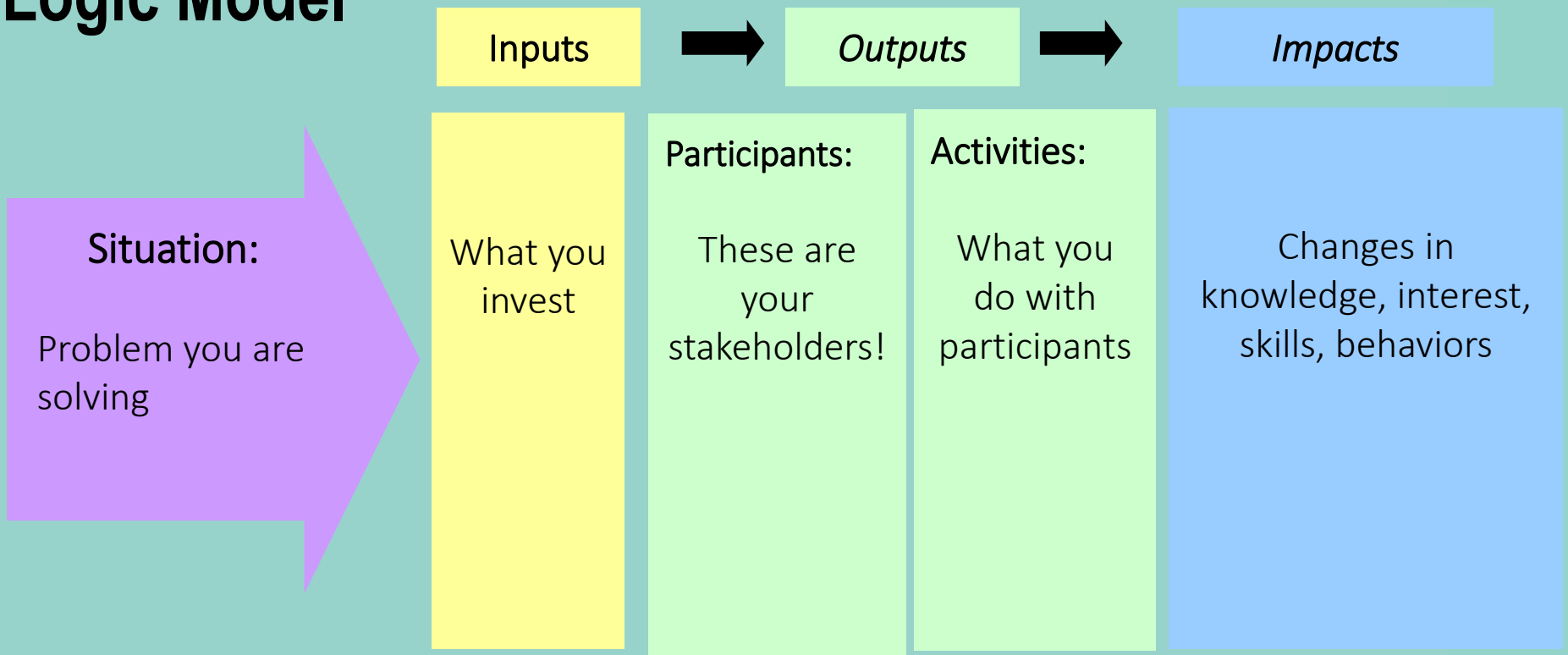
MAPPING your project



"I think you should be more explicit here in step two."

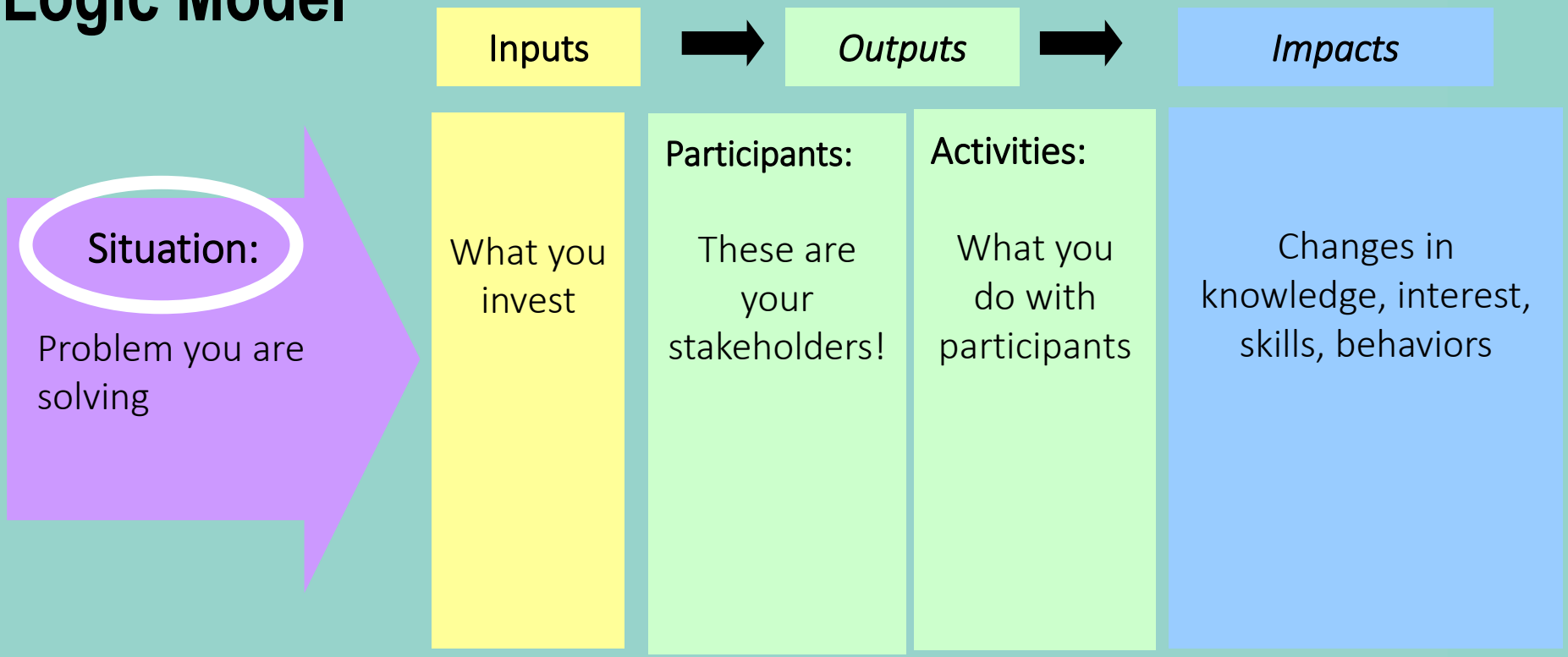
MAPPING your project

Logic Model



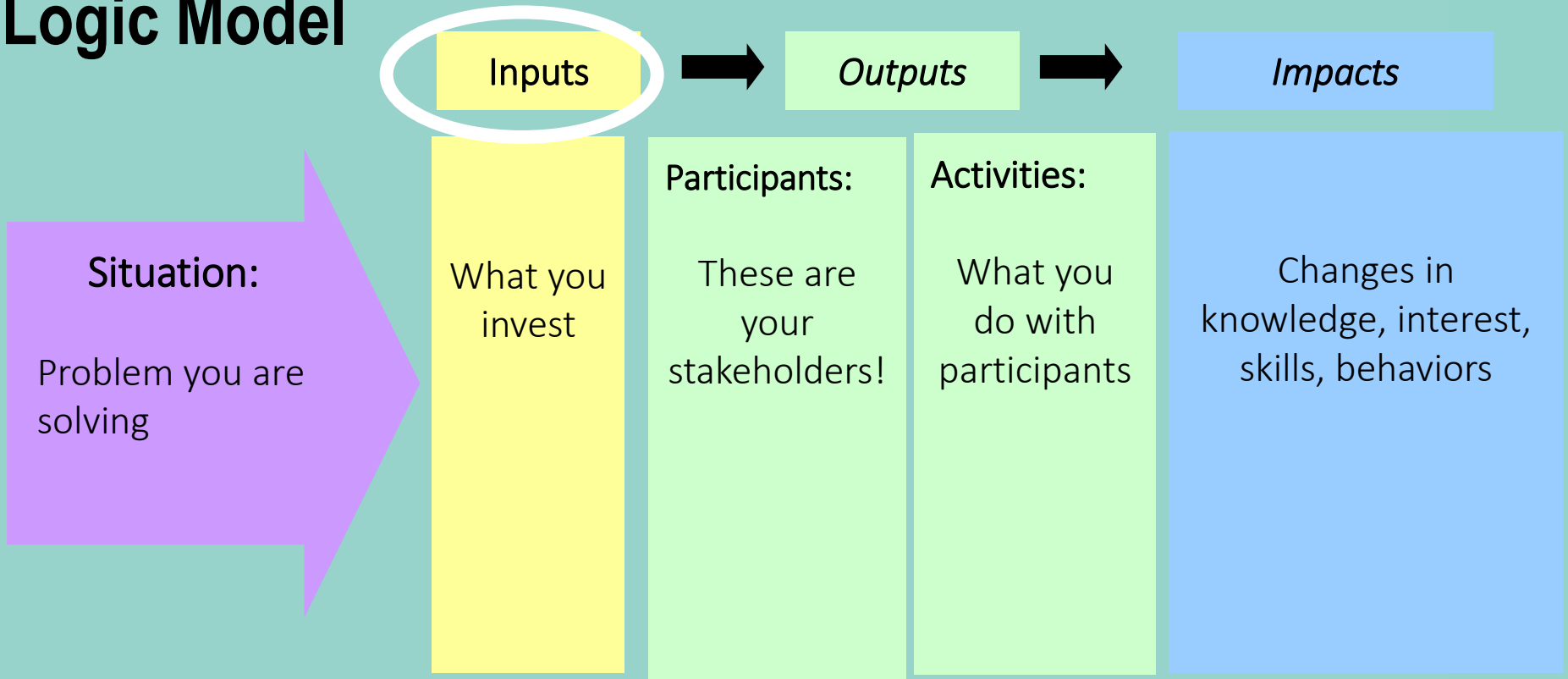
MAPPING your project

Logic Model



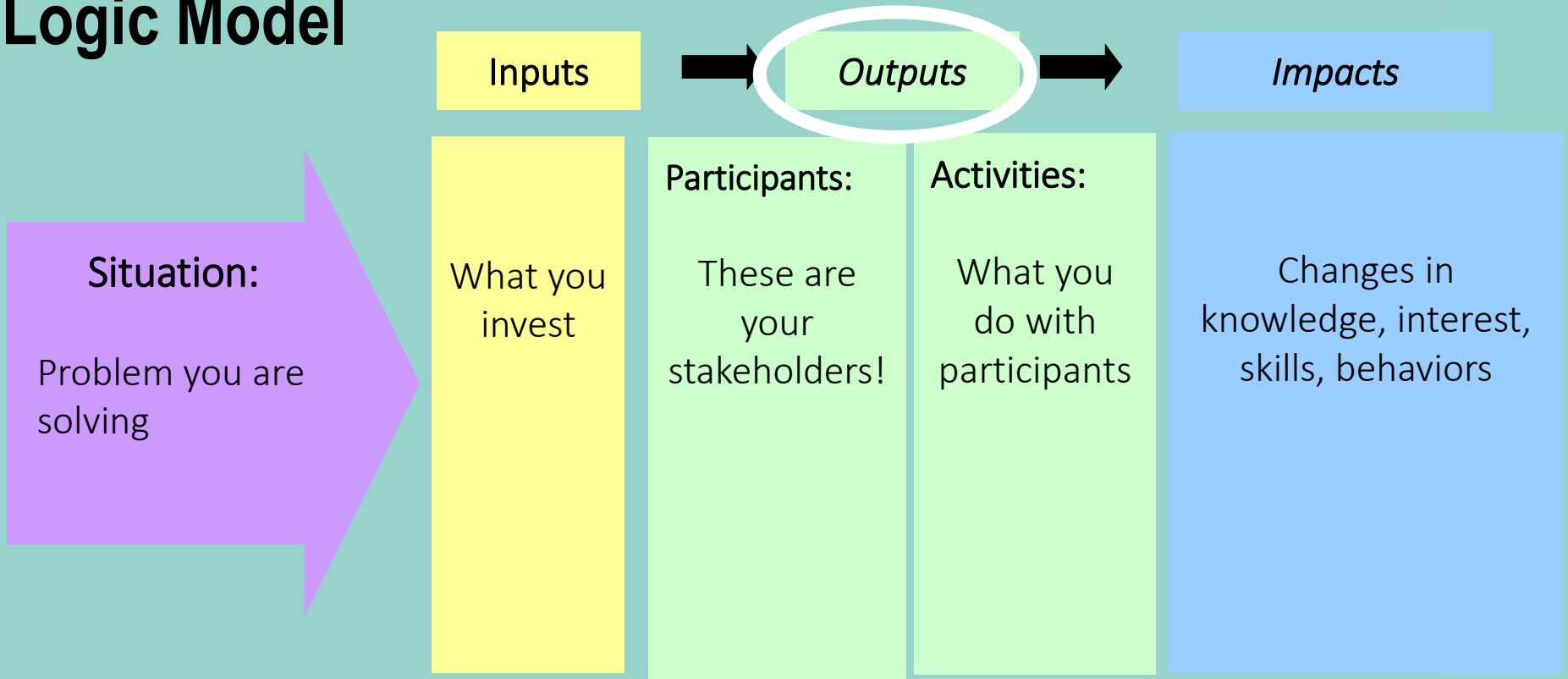
MAPPING your project

Logic Model



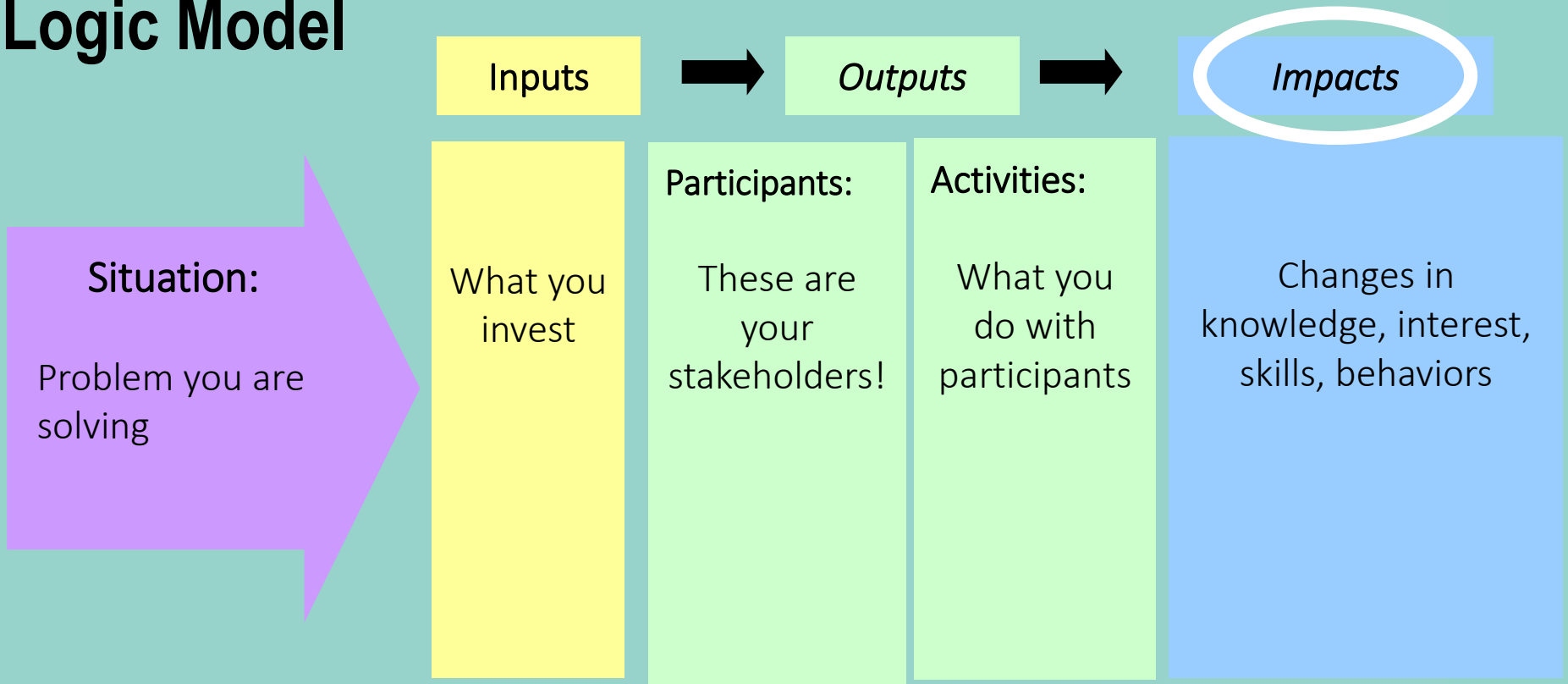
MAPPING your project

Logic Model



MAPPING your project

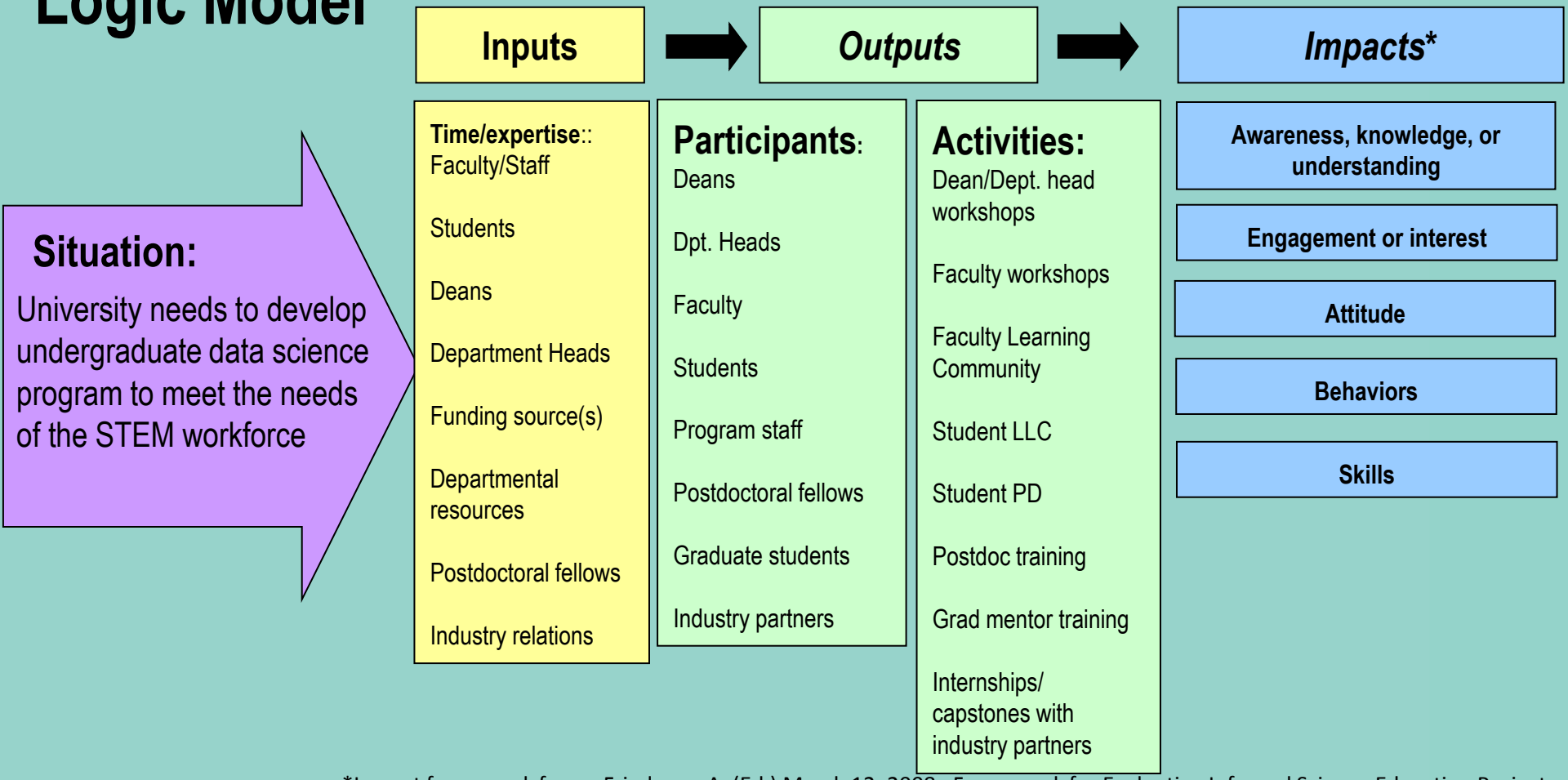
Logic Model



MAPPING your project

Example: Developing an Undergraduate Data Science Program at your University

Logic Model

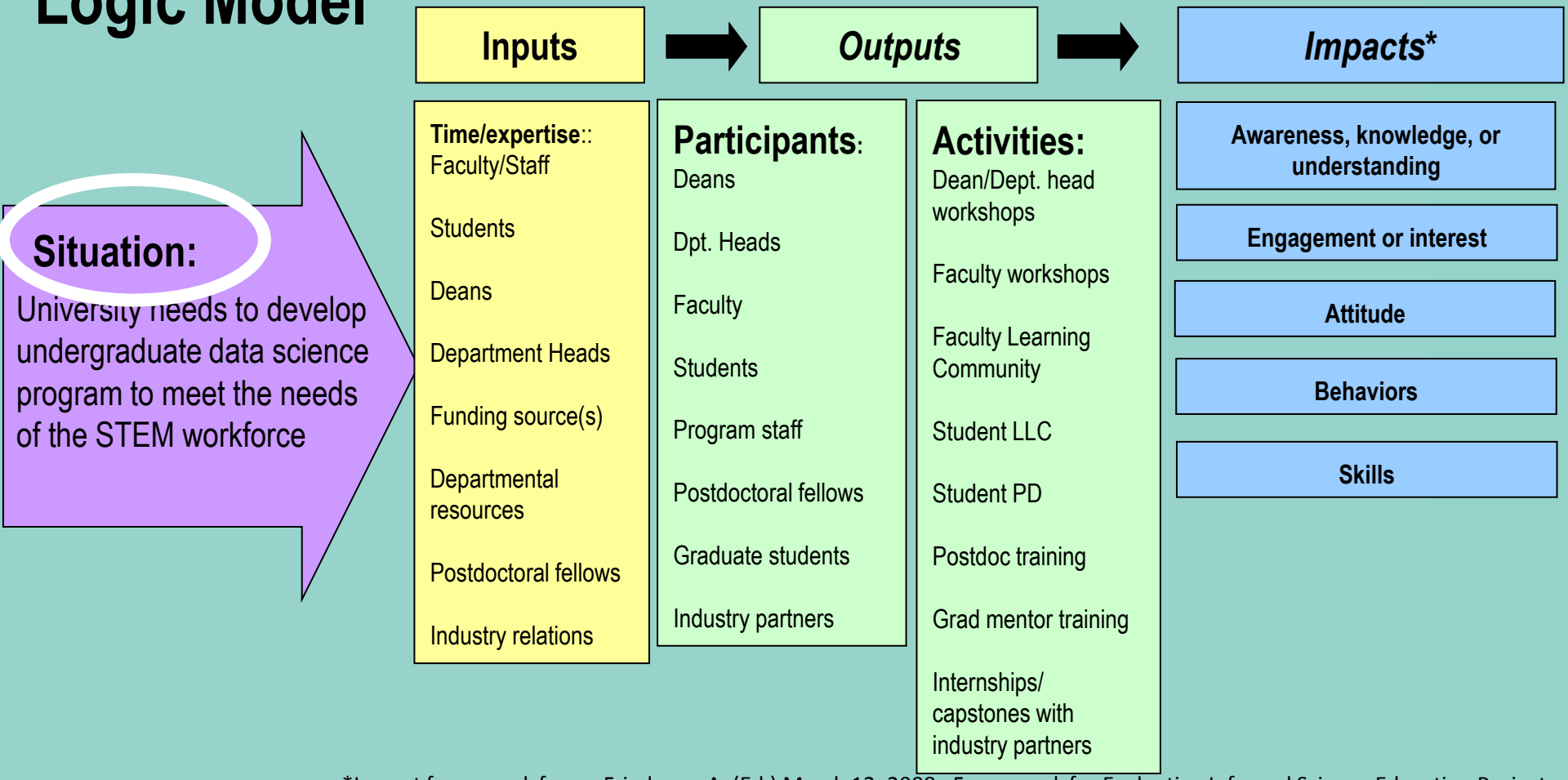


*Impact framework from: Friedman, A. (Ed.) March 12, 2008. Framework for Evaluating Informal Science Education Projects. Online: http://www.informalscience.org/framework-evaluating-impacts-informal-science-education-projects

MAPPING your project

Example: Developing an Undergraduate Data Science Program at your University

Logic Model

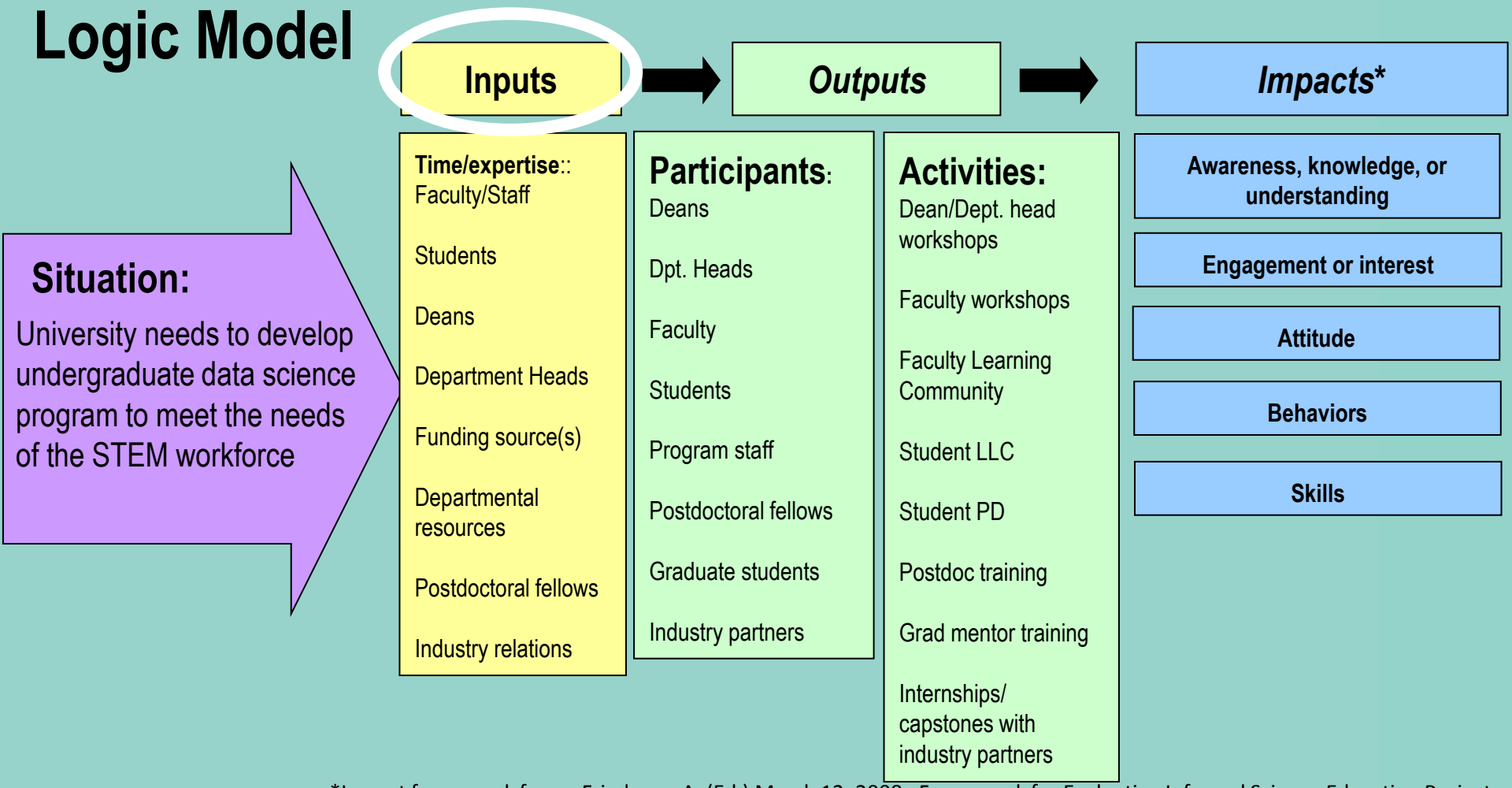


*Impact framework from: Friedman, A. (Ed.) March 12, 2008. Framework for Evaluating Informal Science Education Projects. Online: http://www.informalscience.org/framework-evaluating-impacts-informal-science-education-projects

MAPPING your project

Example: Developing an Undergraduate Data Science Program at your University

Logic Model

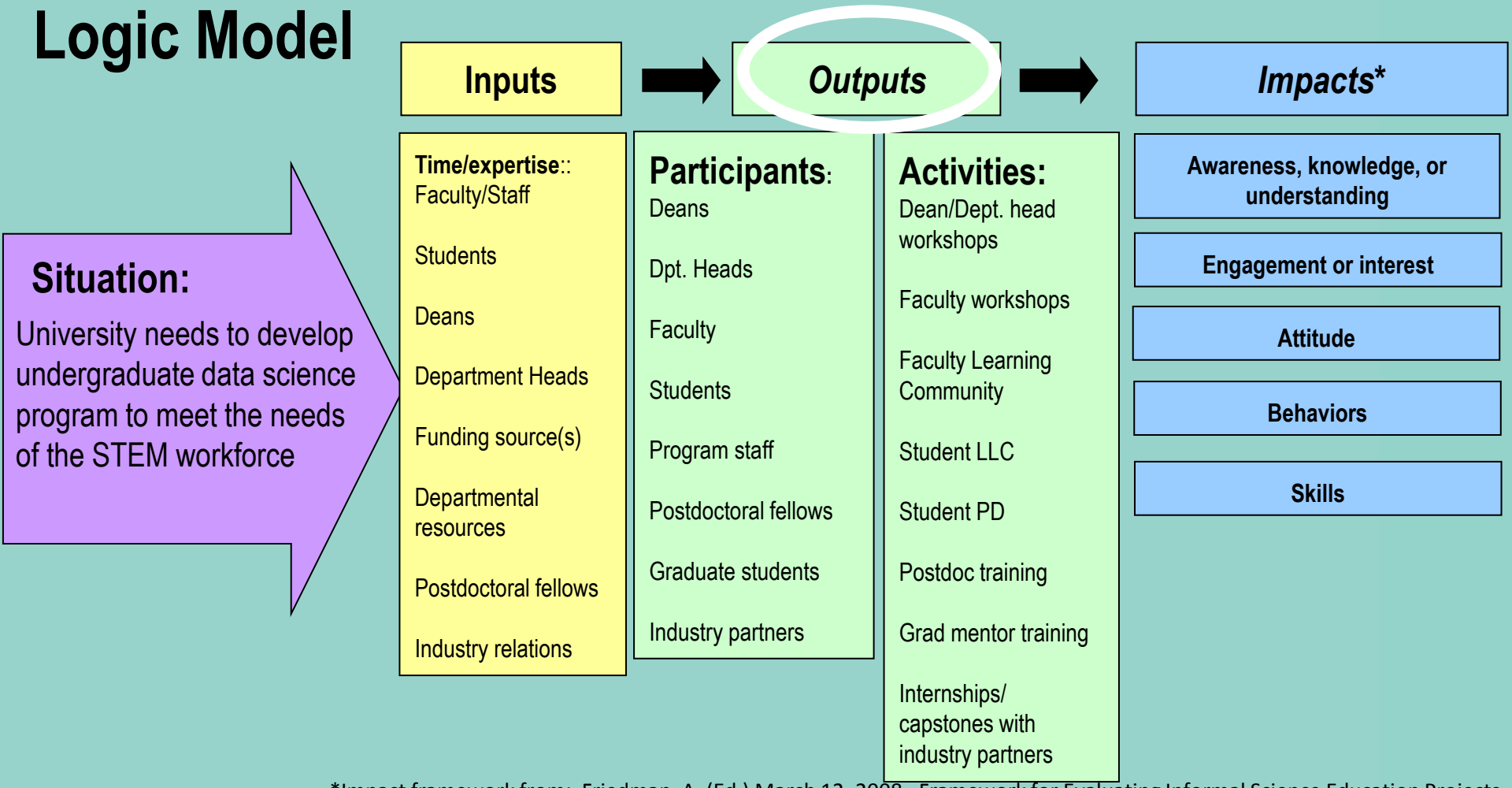


*Impact framework from: Friedman, A. (Ed.) March 12, 2008. Framework for Evaluating Informal Science Education Projects. Online: http://www.informalscience.org/framework-evaluating-impacts-informal-science-education-projects

MAPPING your project

Example: Developing an Undergraduate Data Science Program at your University

Logic Model

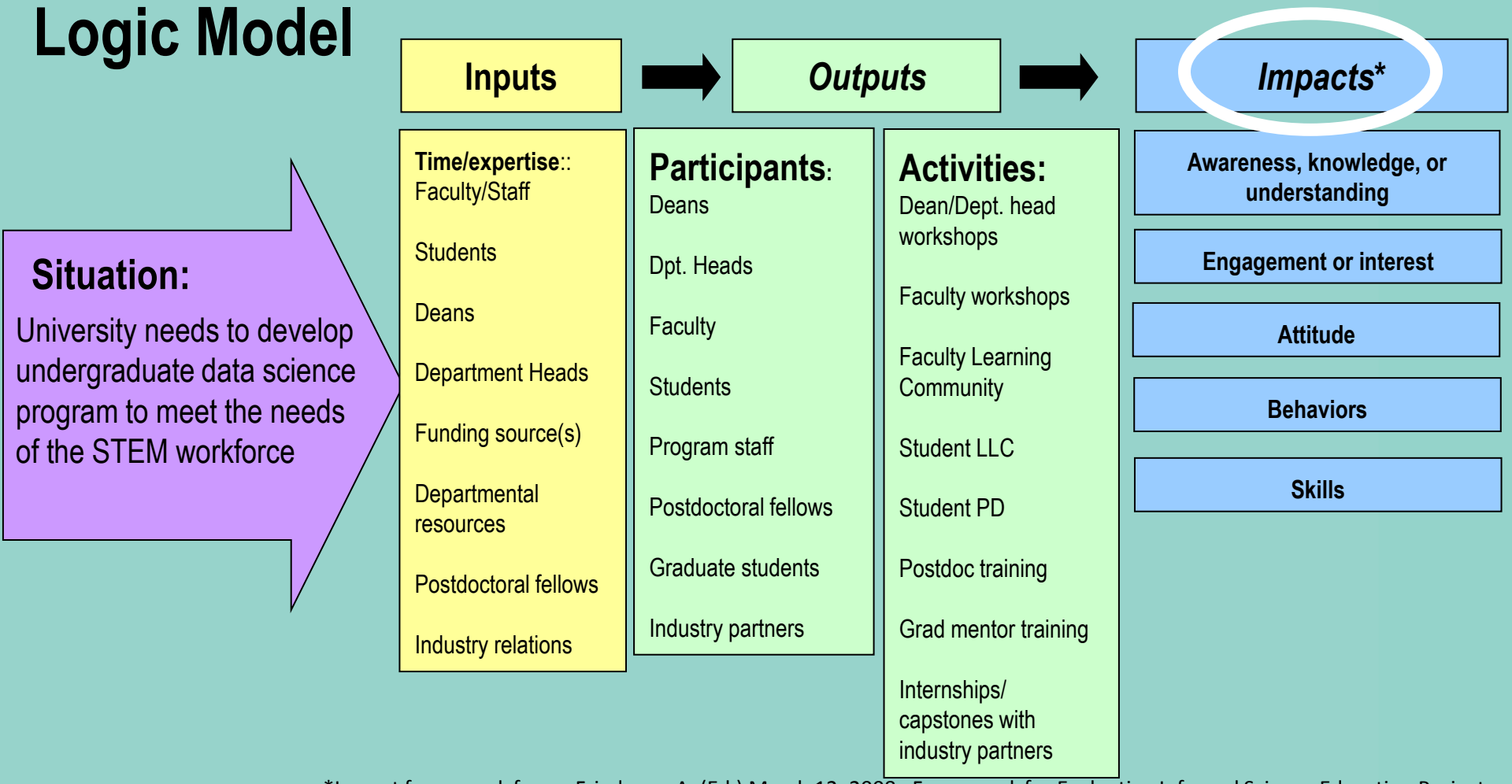


*Impact framework from: Friedman, A. (Ed.) March 12, 2008. Framework for Evaluating Informal Science Education Projects. Online: http://www.informalscience.org/framework-evaluating-impacts-informal-science-education-projects

MAPPING your project

Example: Developing an Undergraduate Data Science Program at your University

Logic Model



*Impact framework from: Friedman, A. (Ed.) March 12, 2008. Framework for Evaluating Informal Science Education Projects. Online: http://www.informalscience.org/framework-evaluating-impacts-informal-science-education-projects

Developing your EVALUATION FRAMEWORK

Impacts	Evaluation questions
STAKEHOLDER*: Administrators** ***IMPACT AREA: Increase awareness, knowledge, or understanding	
STAKEHOLDER*: Industry Partners** ***IMPACT AREA: Interest or engagement	
Faculty ***IMPACT AREA:*** Behaviors	
Students ***IMPACT AREA:*** Attitude	

Developing your EVALUATION FRAMEWORK

Impacts	Evaluation questions
STAKEHOLDER*: Administrators** ***IMPACT AREA: Increase awareness, knowledge, or understanding	To what extent administrators become aware of the importance of collaborating across departments? Do administrators understand ways in which they can remove barriers to faculty participation in the program?
STAKEHOLDER*: Industry Partners** ***IMPACT AREA: Interest or engagement	
Faculty ***IMPACT AREA:*** Behaviors	
Students ***IMPACT AREA:*** Attitude	

Developing your EVALUATION FRAMEWORK

Impacts	Evaluation questions
STAKEHOLDER*: Administrators** ***IMPACT AREA: Increase awareness, knowledge, or understanding	To what extent administrators become aware of the importance of collaborating across departments? Do administrators understand ways in which they can remove barriers to faculty participation in the program?
STAKEHOLDER*: Industry Partners** ***IMPACT AREA: Interest or engagement	What motivates industry partners to be involved in the program? What aspects of the internship and capstone activities do industry partners find useful?
Faculty ***IMPACT AREA:*** Behaviors	To what extend are faculty participating in workshops and FLCs? In what ways are faculty implementing the curricular and pedagogical changes learned and developed through the program?
Students ***IMPACT AREA:*** Attitude	Do students (especially URMS) have a positive attitude about participating in a data science field? What elements of the program help students feel they belong in data science?

Developing your EVALUATION FRAMEWORK

Impacts	Evaluation questions	Data collection (Timing)
Admins. ***IMPACT AREA:*** Inc. awareness, knowledge, or understanding	• To what extent administrators become aware of the importance of collaborating across departments? • Do administrators understand ways in which they can remove barriers to faculty participation in the program?	
Industry ***IMPACT AREA:*** Interest or engagement	• What motivates industry partners to be involved in the program? • What aspects of the internship and capstone activities do industry partners find useful?	
Faculty ***IMPACT AREA:*** Behaviors	• To what extend are faculty participating in workshops and FLCs? • In what ways are faculty implementing the curricular and pedagogical changes learned and developed through the program?	
Students ***IMPACT AREA:*** Attitude	• Do students (especially URMS) have a positive attitude about participating in a data science field? • What elements of the program help students feel they belong in data science?	

Provide input and learn more about the study at www.nas.edu/EnvisioningDS

Developing your EVALUATION FRAMEWORK

Impacts	Evaluation question	Data collection (Timing)
Admins. ***IMPACT AREA:*** Inc. awareness, knowledge, or understanding	• To what extent administrators become aware of the importance of collaborating across departments? • Do administrators understand ways in which they can remove barriers to faculty participation in the program?	• Administrator interviews (annual) • Departmental-level survey (annual) • Observation of workshops (as needed) • Document review (monthly)
Industry ***IMPACT AREA:*** Interest or engagement	• What motivates industry partners to be involved in the program? • What aspects of the internship and capstone activities do industry partners find useful?	
Faculty ***IMPACT AREA:*** Behaviors	• To what extend are faculty participating in workshops and FLCs? • In what ways are faculty implementing the curricular and pedagogical changes learned and developed through the program?	
Students ***IMPACT AREA:*** Attitude	• Do students (especially URMS) have a positive attitude about participating in a data science field? • What elements of the program help students feel they belong in data science?	

Provide input and learn more about the study at www.nas.edu/EnvisioningDS

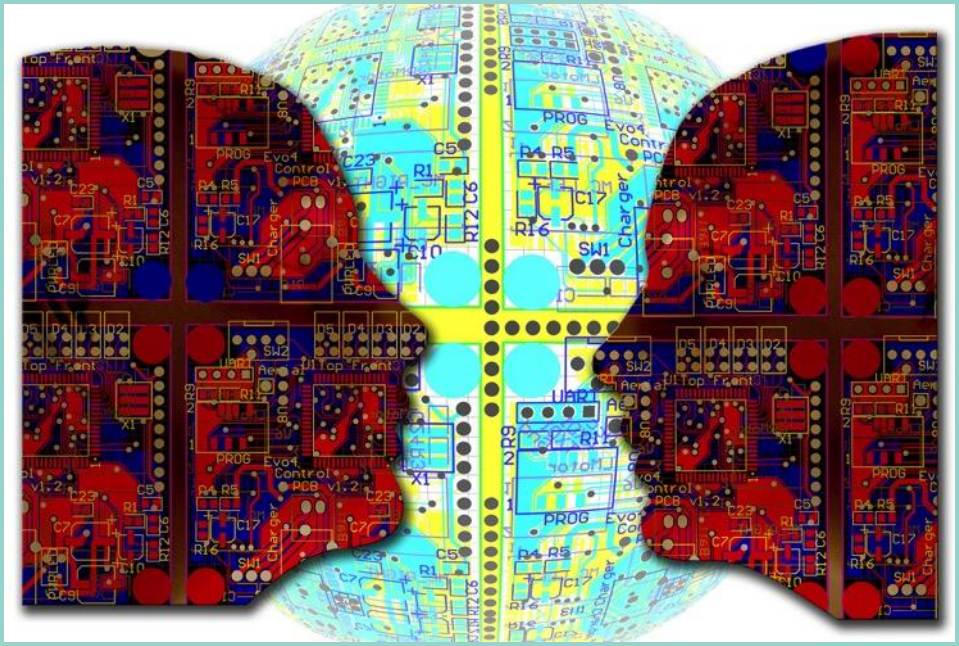
Developing your EVALUATION FRAMEWORK

Impacts	Evaluation question	Data collection (Timing)
Admins. ***IMPACT AREA:*** Inc. awareness, knowledge, or understanding	• To what extent administrators become aware of the importance of collaborating across departments? • Do administrators understand ways in which they can remove barriers to faculty participation in the program?	• Administrator interviews (annual) • Departmental-level survey (annual) • Observation of workshops (as needed) • Document review (monthly)
Industry ***IMPACT AREA:*** Interest or engagement	• What motivates industry partners to be involved in the program? • What aspects of the internship and capstone activities do industry partners find useful?	• Interviews with Industry liaison (biannual) • Interviews with LT (annual) • Survey of students (biannual)
Faculty ***IMPACT AREA:*** Behaviors	• To what extend are faculty participating in workshops and FLCs? • In what ways are faculty implementing the curricular and pedagogical changes learned and developed through the program?	• Document review (monthly) • Classroom observation • (as needed) • Syllabus review (biannually) • Faculty interviews (biannually)
Students ***IMPACT AREA:*** Attitude	• Do students (especially URMS) have a positive attitude about participating in a data science field? • What elements of the program help students feel they belong in data science?	• Survey of students (biannual) • Student focus group (annual) • Student interviews (annual)

Provide input and learn more about the study at www.nas.edu/EnvisioningDS

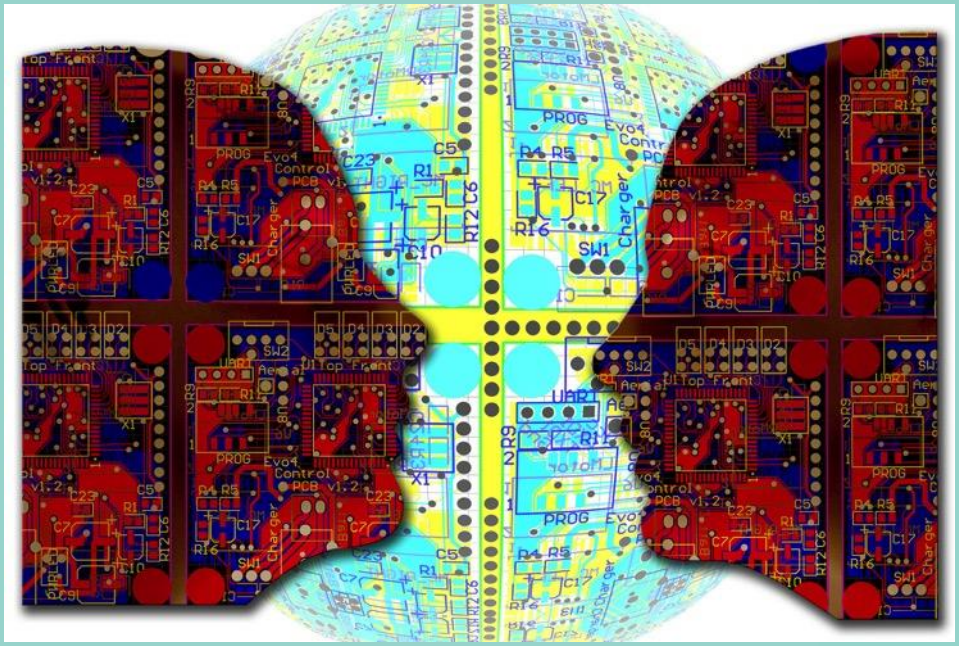
Tips for **DATA SCIENCE EVALUATIONS**

- ✓ Be mindful of the "data acumen" impact goals—not just data skills



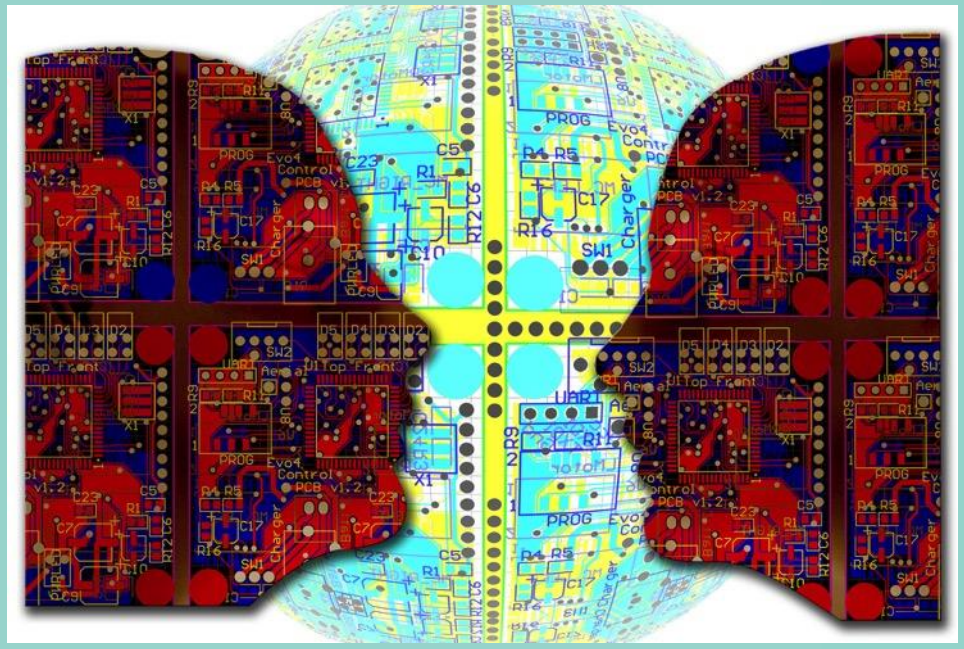
Tips for **DATA SCIENCE EVALUATIONS**

- ✓ Be mindful of the "data acumen" impact goals—not just data skills
- ✓ Consider overall data plan at the start of the program



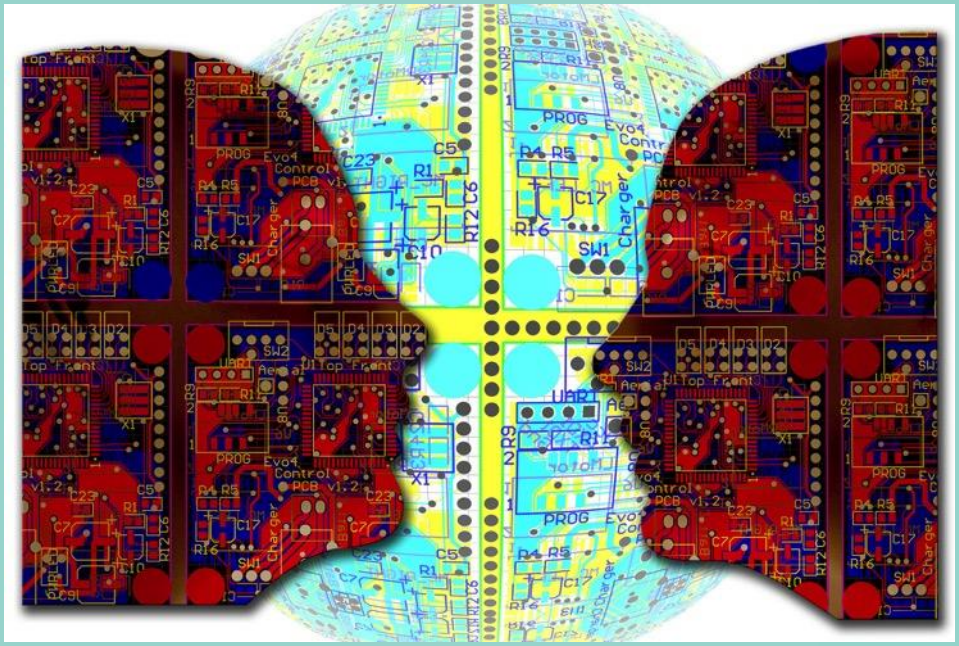
Tips for **DATA SCIENCE EVALUATIONS**

- ✓ Be mindful of the "data acumen" impact goals—not just data skills
- ✓ Consider overall data plan at the start of the program
- ✓ Consider data sources that are readily available for comparison to other relevant programs



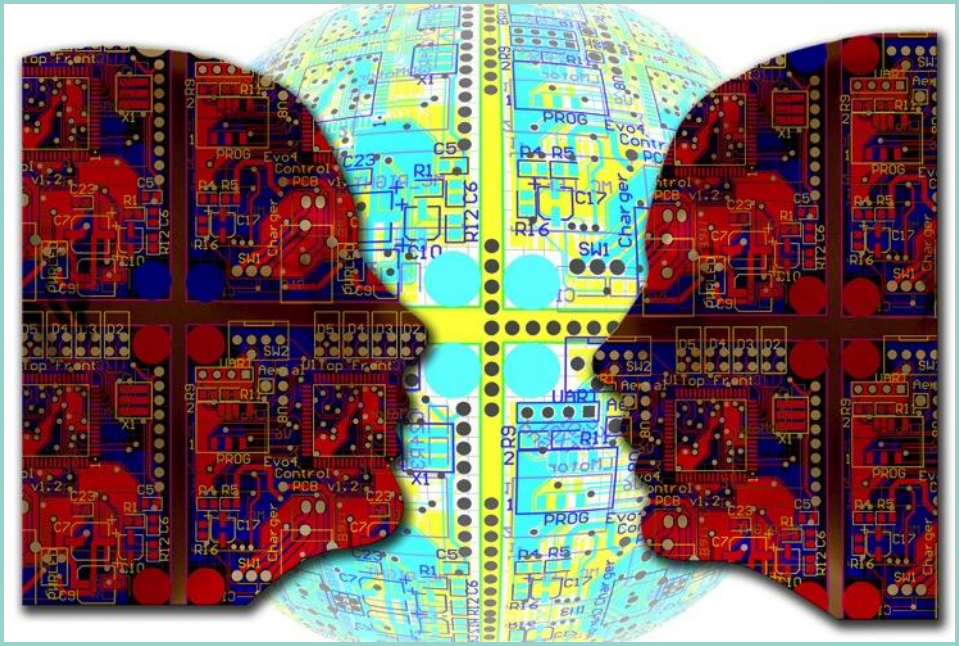
Tips for **DATA SCIENCE EVALUATIONS**

- ✓ Be mindful of the "data acumen" impact goals—not just data skills
- ✓ Consider overall data plan at the start of the program
- ✓ Consider data sources that are readily available for comparison to other relevant programs
- ✓ Address barriers to cross-disciplinary collaboration early and often



Tips for **DATA SCIENCE EVALUATIONS**

- ✓ Be mindful of the "data acumen" impact goals—not just data skills
- ✓ Consider overall data plan at the start of the program
- ✓ Consider data sources that are readily available for comparison to other relevant programs
- ✓ Address barriers to cross-disciplinary collaboration early and often
- ✓ Be aware of cultural differences



ANALYZE AND REPORT data

Quantitative data

Qualitative data

Analyze Data

↓

Interpret Results

↓

Report and Use Results

TYPICAL EVALUATION PROCESS

Provide input and learn more about the study at www.nas.edu/EnvisioningDS

REVISIT your project map

HOW TO LEARN MORE

NISER
NATIONAL INSTITUTE FOR STEM EVALUATION AND RESEARCH

https://www.stemeval.org

NISER Evaluation Channel
http://tiny.utk.edu/NISERvideos

Program Evaluation 101
NISER/NIMBioS NSF INCLUDES Conference Webinar
https://youtu.be/ZGsNJ1jIJD0

EvaluATE
http://www.evalu-ate.org/

BetterEvaluation
Sharing information to improve evaluatio
http://www.betterevaluation.org/

Provide input and learn more about the study at www.nas.edu/EnvisioningDS

Envisioning the
DATA SCIENCE DISCIPLINE
The Undergraduate Perspective

Assessment and Evaluation

Assessing Data Science Learning Outcomes

Kari L. Jordan, PhD
Data Carpentry
Deputy Director for Assessment

Data Science: Core Skills

1 Collect

- Structure data for effective use
- Data wrangling/munging
- Export for downstream applications

Tools: OpenRefine, etc.

2 Analyze

- Mathematical/Statistical analysis
- Open-source tools
- Collaboration

Tools: R, Python, etc.

3 Visualize

- Comparative charts/grids
- Storytelling (graphics/maps)
- Data driven documents

Tools: RMarkdown, Tableau, etc.

4 Share

- Context/customer
- GitHub/Zenodo
- Reproducibility

Tools: Git, Mercurial, etc.

Data Science Learning Community

Approach

Intensive trainings, workshops, and/or short courses on data literacy.

Learn best practices for (teaching) computational skills and how they can be used in the classroom.

Provide input and learn more about the study at www.nas.edu/EnvisioningDS


-Question-

Consider

Workshop Impact

- Learner profiles
- Positive learning environment
- Working memory and cognitive load
- Motivation and demotivation
- Feedback to improve teaching
- Challenge questions/exercises
- Pre/post surveys
- Long-term follow-up
- Ongoing community building

Test Case: Two-day workshop

Aim: Introduce best-practices for data organization and data cleaning.

Learning Objectives

Assessment Instruments

Sample Learning Outcomes

Data Organization (Spreadsheets)

Apply best practices to arrange variables and observations in a spreadsheet.

Apply quality control techniques to identify errors in spreadsheets and limit incorrect data entry.

Store spreadsheet data in universal file formats.

Export data from a spreadsheet to a .csv file.

Data Cleaning (OpenRefine)

Differentiate *data cleaning* from *data organization*.

Recall what facets are and how they are used to sort and summarize data.

Recall what clustering is and how it is applied to group and edit typos.

Export cleaned data from an OpenRefine project.

Day 1

Schedule

	Setup	Download files required for the lesson
00:00	1. Introduction	What are basic principles for using spreadsheets for good data organization?
00:18	2. Formatting data tables in Spreadsheets	How do we format data in spreadsheets for effective data use?
00:53	3. Formatting problems	What are some common challenges with formatting data in spreadsheets and how can we avoid them?
01:13	4. Dates as data	What are good approaches for handling dates in spreadsheets?
01:26	5. Quality control	How can we carry out basic quality control and quality assurance in spreadsheets?
01:46	6. Exporting data	How can we export data from spreadsheets in a way that is useful for downstream applications?
01:56	Finish	

Day 2

Schedule

	Setup	Download files required for the lesson
00:00	1. Introduction	What is OpenRefine useful for?
00:10	2. Working with OpenRefine	How can we bring our data into OpenRefine? How can we sort and summarize our data? How can we find and correct errors in our raw data?
00:45	3. Filtering and Sorting with OpenRefine	How can we select only a subset of our data to work with? How can we sort our data?
01:05	4. Examining Numbers in OpenRefine	How can we convert a column from one data type to another? How can we visualize relationships among columns?
01:25	5. Scripts from OpenRefine	How can we document the data-cleaning steps we've applied to our data? How can we apply these steps to additional data sets?
01:40	6. Exporting and Saving Data from OpenRefine	How can we save and export our cleaned data from OpenRefine?
01:55	7. Other Resources in OpenRefine	What other resources are available for working with OpenRefine?
02:05	Finish	

Assessing Learning Outcomes

Provide input and learn more about the study at www.nas.edu/EnvisioningDS

Assessment Questions

- **Ability**: Skills-based questions
- **Confidence**: Work with data, find answers to questions
- **Motivation**: Continuous learning, community involvement
- **Perception:** Usefulness tools, workshop impact, instructors/facilitators

Provide input and learn more about the study at www.nas.edu/EnvisioningDS

Sample Assessment Questions

7. Which of the following are fundamental rules for producing well formatted spreadsheet tables? Check all that apply.

- [ ] Put each variable (e.g. 'weight' or 'temperature') in its own column.
- [ ] Put each observation in its own row.
- [ ] Combine related pieces of information in one cell.
- [ ] Leave the raw data raw and make edits to a copy of the data.
- [ ] Place comments alongside data values within a single cell, so they don't get separated.

Sample Assessment Questions

14. Please rate your level of agreement with the following statements:

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Having access to the original, raw data is important to be able to repeat an analysis.	○	○	○	○	○
I can write a small program/script/macro to solve a problem in my own work.	○	○	○	○	○
I know how to search for answers to my technical questions online.	○	○	○	○	○
While working on a programming project, if I get stuck, I can find ways of overcoming the problem.	○	○	○	○	○
I am confident in my ability to make use of programming software to work with data.	○	○	○	○	○

Sample Assessment Questions

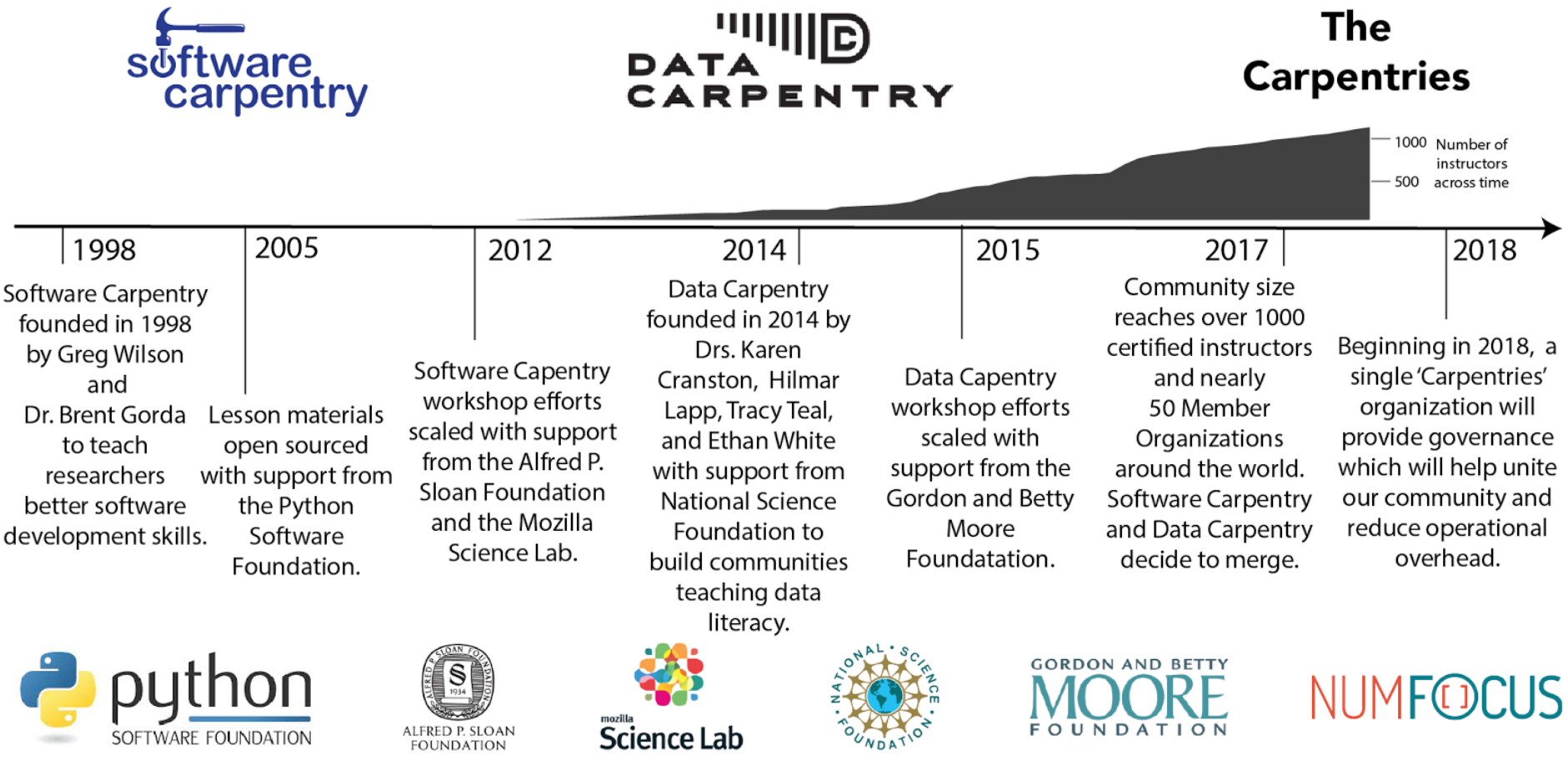
18. The statements below reflect ways in which completing a Carpentries workshop may have impacted you. Please indicate your level of agreement with these statements.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I have improved my coding practices as a result of completing the workshop.	○	○	○	○	○
My research productivity has improved as a result of completing the workshop.	○	○	○	○	○
I have gained confidence in working with data as a result of completing the workshop.	○	○	○	○	○

Summary

- Identify goals/learning outcomes.
- Design your intervention.
- Assess learning outcomes.
- Build/grow your community (meetups, etc.)
- Long-term follow-up.

We can help!



Provide input and learn more about the study at www.nas.edu/EnvisioningDS

Envisioning the DATA SCIENCE DISCIPLINE

The Undergraduate Perspective

Assessment and Evaluation – Q&A

Kari L. Jordan, PhD

Data Carpentry

Deputy Director for Assessment

kariljordan@carpentries.org

www.linkedin.com/in/kariljordan

www.github.com/kariljordan

@drkariljordan on Twitter

Envisioning the DATA SCIENCE DISCIPLINE

The Undergraduate Perspective

Assessment and Evaluation – Q&A

Pamela Bishop, Univ. of Tennessee, Knoxville
Director, National Institute for STEM Evaluation and Research (NISER)
NIMBioS Associate Director for STEM Evaluation

Kari Jordan, Data Carpentry
Deputy Director for Assessment

Provide input and learn more about the study at www.nas.edu/EnvisioningDS

Envisioning the DATA SCIENCE DISCIPLINE

The Undergraduate Perspective

9/12/17 – Building Data Acumen *(recording posted)*

9/19/17 – Incorporating Real-World Applications *(recording posted)*

9/26/17 – Faculty Training and Curriculum Development *(recording posted)*

10/3/17 – Communication Skills and Teamwork *(recording posted)*

10/10/17 – Inter-Departmental Collaboration and Institutional Organization *(recording posted)*

10/17/17 – Ethics *(recording posted)*

10/24/17 – Assessment and Evaluation for Data Science Programs

11/7/17 – Diversity, Inclusion, and Increasing Participation

11/14/17 – Two-Year Colleges and Institutional Partnerships

Provide input, download the interim report, and learn more about the study at www.nas.edu/EnvisioningDS