Challenges and opportunities in training

Tracy K. Teal, PhD
Data Carpentry, Executive Director

Vision

Universal data literacy powering superior and reproducible research: towards a community of practice

Universal **data literacy** powering superior and reproducible research

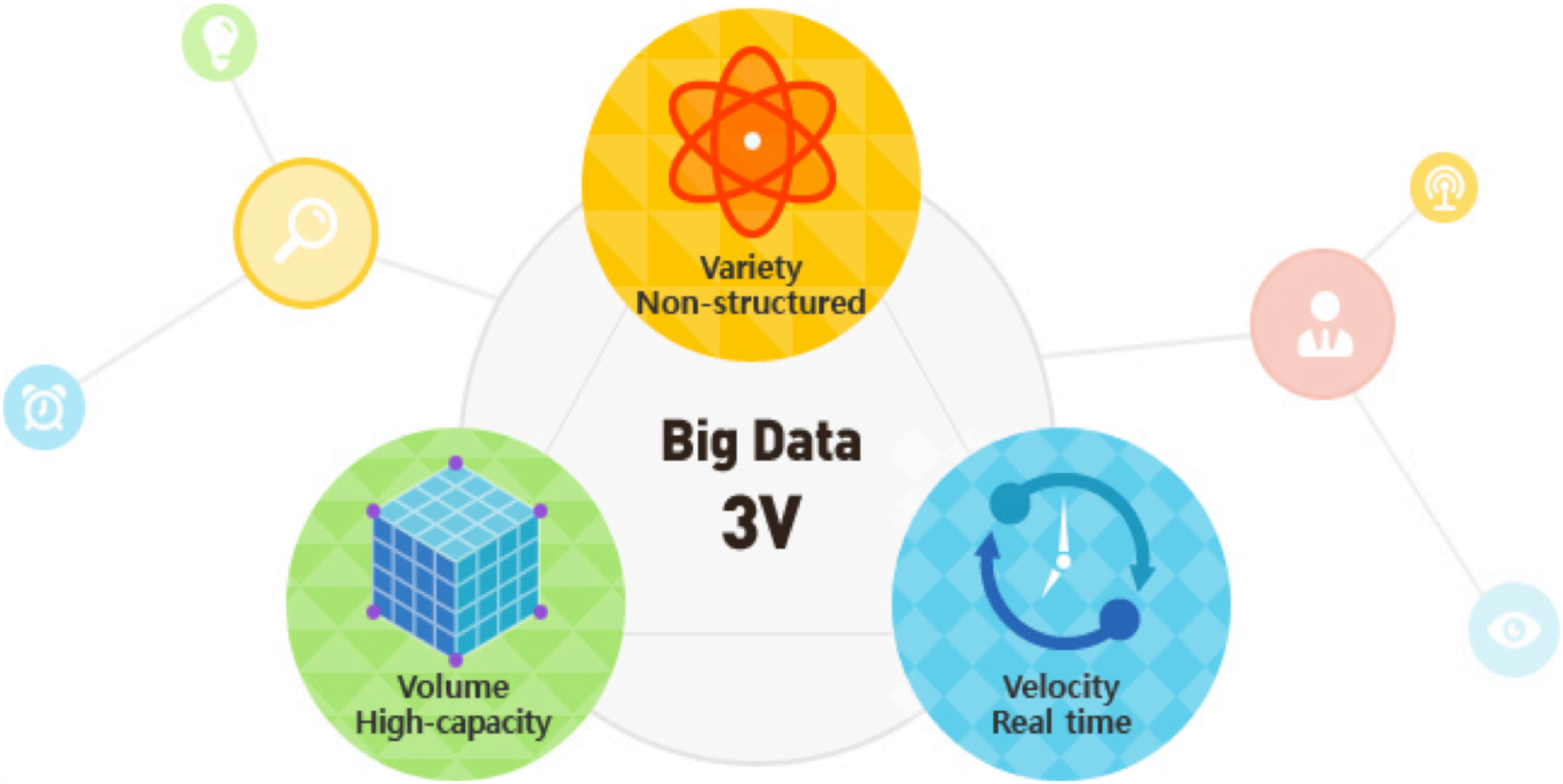
Universal data literacy powering **superior and reproducible research**

Universal data literacy powering superior and reproducible research

Let's Make Gender Diversity in Data Science a Priority Right from the Start

by Francine D. Berman & Philip E. Bourne

Data production is no longer the bottleneck

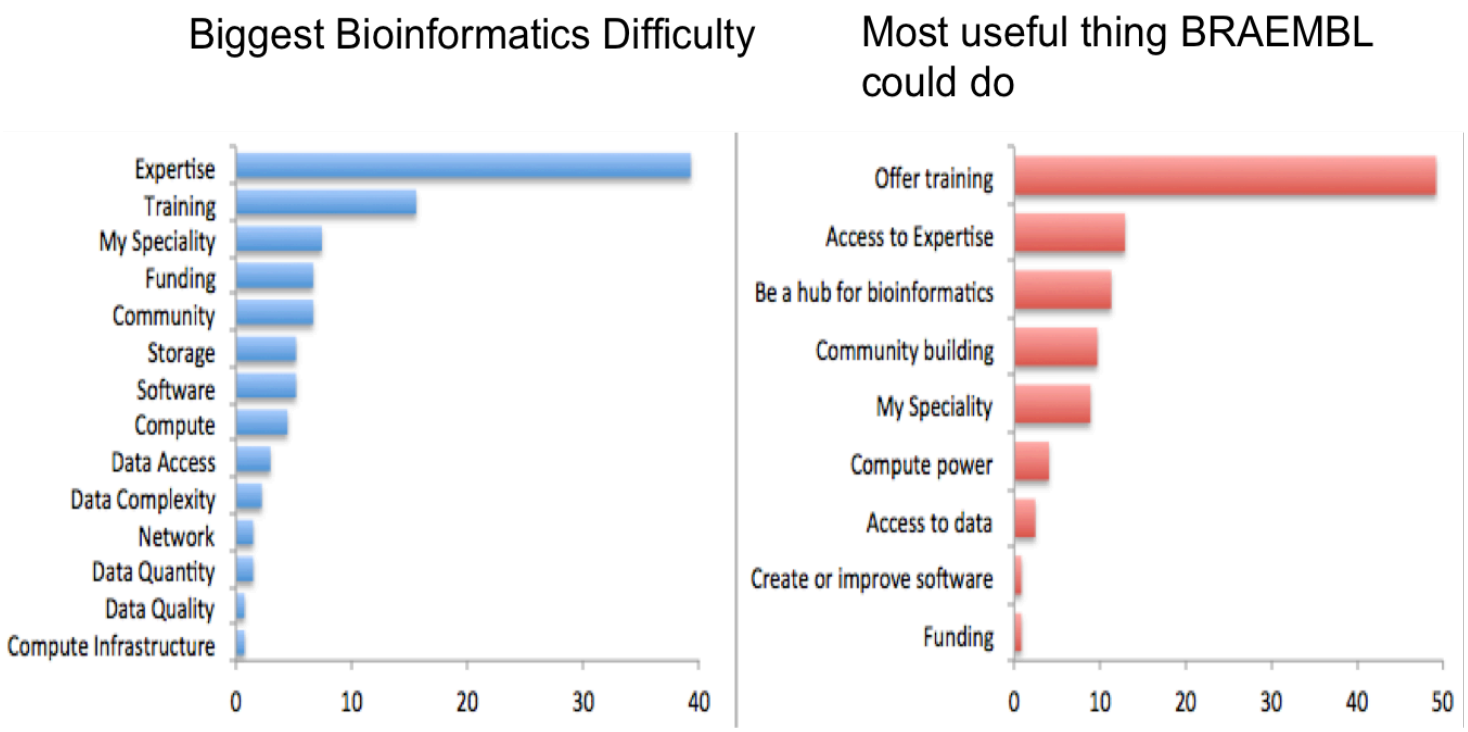

Variety
Non-structured
Big Data
3V
Volume
High-capacity
Velocity
Real time

How do we scale data literacy along with data production?

Training Opportunities

- K-12
- Undergraduate
- **Active Researchers**

Motivated researchers



BRAEMBL community survey report

Training in the Gaps

Active researchers are learning these skills "on the job"

Need to develop and deliver training that fits in their time and attention

Possibilities:

- online self-guided tutorials
- online self-guided short videos
- workshops
- short courses
- one-on-one or small group consulting

Process of developing & deploying training

(like a drug development process, but hopefully faster)

- Resarch & Development
- Phase 1 trials - small group
- Phase 2 trials - larger group, target audience
- Phase 3 trials? - larger group, with controls
- Approval!
- Phase 4 - broad deployment

Case study: Software and Data Carpentry

Research & Development

- Learning objectives: Best practices in software development and data analysis & management
- Delivery approach: intensive, hands-on teaching strategies in 2-day workshops, informed by educational pedagogy
- Delivery mechanism: collaboratively developed lessons available on-line for free use, instructors trained to teach
- Materials development - development of lessons and exercises
- Assessment - assessment surveys for both assessing learning outcomes and improving content

Case study: Software and Data Carpentry

Phase 1 and Phase 2

- Teach workshops to larger numbers of learners, assess outcomes and improve
- Development of an instructor training program to scale delivery
- Development of strategies for lesson improvement and maintenance

Case study: Software and Data Carpentry

Phase 3 and Phase 4

- Teaching a large number of workshops
- Engaging in assessment, but improving
- Continued scaling challenges, increased need for operations support
- New development
- Community!

Case study: Software and Data Carpentry

Community!

An active and engaged community of instructors and learners, both using and advocating for best practices in effective and reproducible research

How best to support, sustain and grow this community?

Acknowledgements

- Alfred P. Sloan Foundation
- Gordon and Betty Moore Foundation
- National Science Foundation

Summary

- Active researchers are very motivated, but need 'on the job' training
- Training for active researchers needs to teach in the gaps
- Different challenges & opportunities in the training process that potentially require different strategies for support and different groups to develop and deliver