Declaration: a progressive culture for data in public health

A progressive culture for data in public health keeps up with fast-moving methods, tools, and technology for continuously learning things about the world and empowering choices informed by those learnings. This declaration frames the major elements for cultivating and sustaining this progressive culture and for leading public health into the modern era. Public health scientists who care about data can flourish in a culture that fosters technical skills, inspires and rewards intellectual drive, and supports learners in community.

- Technical skills include math and statistics, programming and data structures, communications and visualization, and domain knowledge in public health and allied fields.
- Nontechnical skills put the traits of a good learner or knower into action, motivating and enabling those who care about data to learn and to practice wisely.
- Support for empowering and doing good things with data comes from a community of peers, mentors, and advocates centered on learning.

A progressive culture for data in public health manifests the following principles:

1. Data as object: A progressive culture for data is **dedicated to learning *from* data**. In a progressive culture, data have value because they mediate how we learn things about the world. Those learnings allow us to make informed choices about how we interact with the world, for example, through public health interventions.
2. Data as subject: A progressive culture for data is **dedicated to learning *about* data**, because data come in many structures, sizes, shapes, and speeds, from small, flat data tables to massive, unstructured data streams. Data conform to a variety of standards, or no standards at all. The varied characteristics and complexity of data both enrich and constrain the ways that data reveal characteristics of the world.
3. Data as mediator: A progressive culture for data is **dedicated to learning *with* data** through its full life cycle[*], primarily through knowing how analytic methods allow us
    - 3.1. to pose rich questions about the world, amenable to rich methods;
    - 3.2. to guide how we generate, transmit, obtain, and prepare data;
    - 3.3. to probe data to answer questions about the world;
    - 3.4. to place answers from data in context, mindful of assumptions and alternatives;
    - 3.5. to present data-driven answers to audiences clearly and correctly;
    - 3.6. to preserve those answers and ensure that the entire life cycle is transparent, accessible, traceable and, to the extent possible, reproducible.
4. Community supports doing good things with data through 4 primary roles:
    - 4.1. **Learner-practitioners** with basic or intermediate data skills come from any discipline to do good things with data, mindful of the full life cycle of data. In a progressive culture for data, everyone who wants to do good things with data has the intellectual support

[*] Items 3.1-3.5 are adapted from *The Art of Data Science*.

to do so, accepting that they must proceed with rigor and stand behind their work. Practitioners learn continuously and show how modern tools and methods solve modern problems.

4.2. **Expert practitioners** go deep on data science methods. Experts help ensure that everyone who wants to do good things with data, can. They set norms for the practice of data science and for learning from, about, and with data. They enable, guide, correct, and empower practitioners to proceed with rigor and stand behind their work.

4.3. **Managers** supervise practitioners and experts, to ensure that they have the resources and direction that they need to achieve good things with data, now and in the future. Managers foster and reward curiosity, invest in learning (not just training), and encourage innovation and interesting mistakes. They advocate for the means to enable practitioners and experts to continue increasing their capability, efficiency, and effectiveness. Managers hold practitioners and experts to account for producing knowledge learned from, about, and with data.

4.4. **Lay advocates** work in community with practitioners, experts, and managers as persons literate in the value of data to help learn things about the world. Laypersons know how to assess, use, and advocate for learning from data. They help ensure supportive resources to enable learning.

5. Members of a progressive culture for data **respect the fundamentals, value innovation, and practice pragmatic, principled pluralism**. Members apply wisely and well all methods that can help them achieve technical excellence to learn from, about, and with data. These methods include the classical, conventional, and innovative; statistics and machine learning; correlational, causal, and predictive inference; analysis, synthesis, and forecasting. Principled pluralism allows honest disagreement about methods, results, and interpretation.

6. A progressive culture for data in public health **fosters public trust**, motivated by public service to conduct itself ethically, **protect privacy**, and ensure that data and methods are **radically open and transparent**.

7. In a progressive culture for data, **leadership is radically and intentionally inclusive**, continually shaping and sustaining the culture of good data practice. Practitioners, experts, managers, and laypersons lead from every level, regardless of their career stage, job title, credential, or job series, so that everyone who wants to do good things with data, can. In a progressive culture for data, governance enables effective leadership, but governance does not substitute for leadership.

The public health sector faces constant challenges to stretch modest resources, to anticipate and respond to threats, and to promote population health. A progressive culture remains rooted in history and continues to learn from old data in new ways, and it anticipates the future and handles evolving demands and to keep up with fast-moving methods, tools, and technology. Cultivating a progressive data culture in the present will best position the field of public health as ever ready to learn from and act on data.