

REDUCING BARRIERS IN PUBLISHING WATER QUALITY DATA

AN INTRODUCTION

DWANE YOUNG

U.S. EPA, OFFICE OF WATER

THE CHARGE!

Propose 1 to 2 concrete actions that we, as a community, can take to address these two challenges:

1. The learning curve on understanding WQX is too high
2. How can we make data more usable?

DATA PUBLISHING THROUGH WQX

- The Water Quality Exchange (WQX) is the standard for sharing/publishing water quality data
- Think of it as the ‘superhighway’ for sharing water quality monitoring data, which includes the Water Quality Portal as the means for accessing all the data
- Some terms:
 - “Data of Documented Quality”
 - “Metadata”



CHALLENGE #1: THE LEARNING CURVE ON UNDERSTANDING WQX IS TOO HIGH

- WQX can handle just about any kind of sample data
 - This leads to it being very complex
- A Question for us to Answer:
 - How do we make it easy to publish data of known quality?
 - Balance 'simplicity' with 'data reusability'



A COMPARISON



Consider a simple photo sharing app. It's easy to use:

1. Simply take a photo
2. Review that photo
3. Add some metadata
4. Share the photo for the world to see



<Lat/Long> White Pine Lake on a beautiful
<Time> September day!

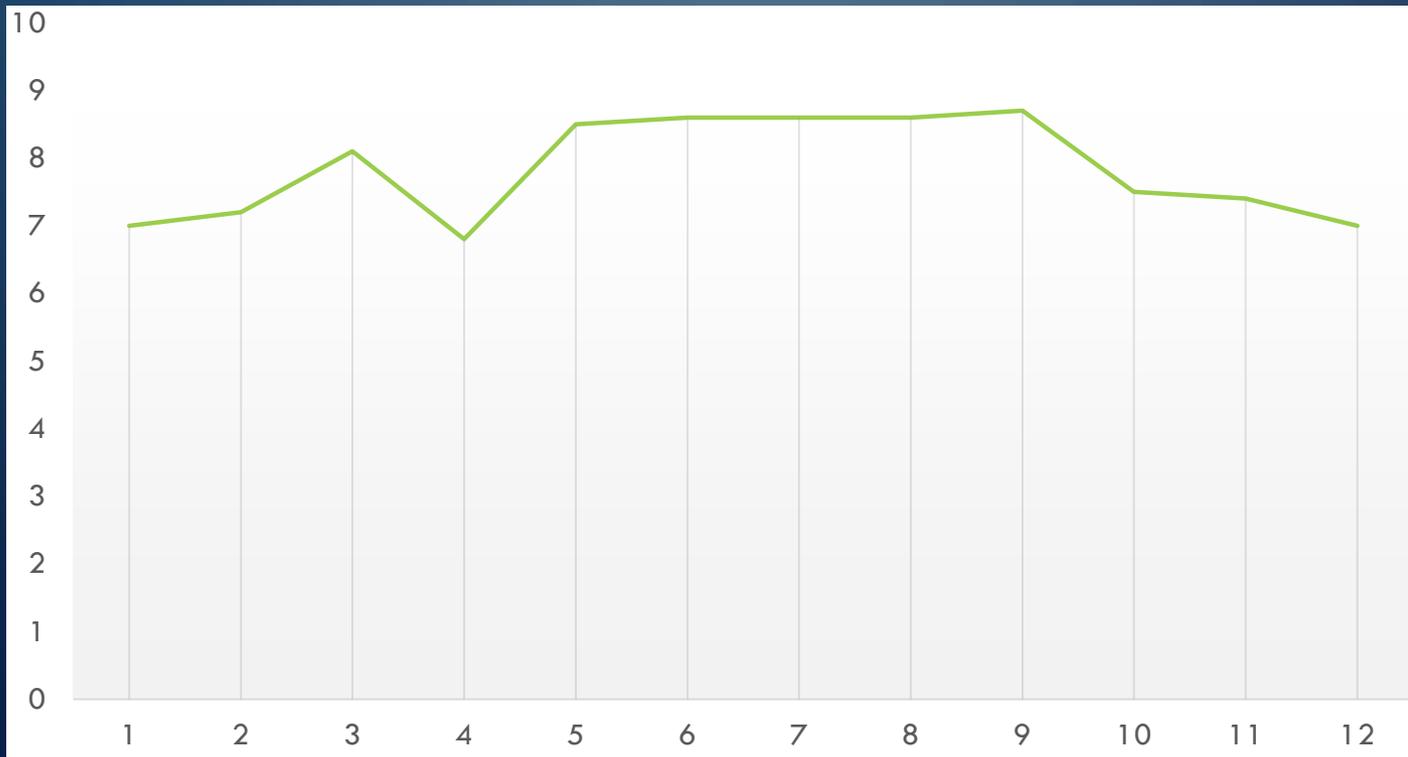


Provides a very specific use, but there are limits on how this information can be used

HOW ABOUT A WATER QUALITY SAMPLE?

- In order to really describe the sample you need:
 - Location (including metadata about the location)
 - Sample and Field Information
 - How was it collected
 - When was it collected
 - What Parameters were analyzed (pH, Dissolved Oxygen, etc.)
 - How was each parameter analyzed
 - Any data quality flags related to the data
 - Also in order to make it usable by other, you need unique identifiers for all of this

IT ALL COMES DOWN TO METADATA: SO, WHAT DO WE MEAN?



Is it enough to just have concentration over time?

CHALLENGE #2: HOW CAN WE MAKE DATA MORE USABLE?

- Clear communication on data quality
- Data discoverability
- Data analysis

How do we stimulate a positive feedback loop between data publishing and data use?



THE CHARGE!

Propose 1 to 2 concrete actions that we, as a community, can take to address these two challenges:

1. The learning curve on understanding WQX is too high
2. How can we make data more usable?

Ohio EPA Perspective

- Ohio law restricts data Ohio EPA can use for regulatory purposes
- 1st Priority is getting data with highest level of credibility into WQX/STORET
- Need bulk uploads and linking to existing databases

Ohio EPA Perspective

The Learning Curve on Understanding WQX is too high

- Success connecting USEPA assistance with Ohio DNR program to map their data to WQX
- Plan to provide technical support to small users
- Consider collaborating with labs and/or vendors that develop laboratory Information Management Systems

Ohio EPA Perspective

How Can We Make Data More Usable?

- Create fields and staging area in WQX to allow reviewers to correct/add information and certify level of quality/credibility where appropriate
- Develop apps/forms to automatically include location, weather with field notes and upload with sample information
- Use interactive maps to choose and download data
 - tie to instant download to Excel or R or option to display basic statistical results

Colorado Water Quality Monitoring Council



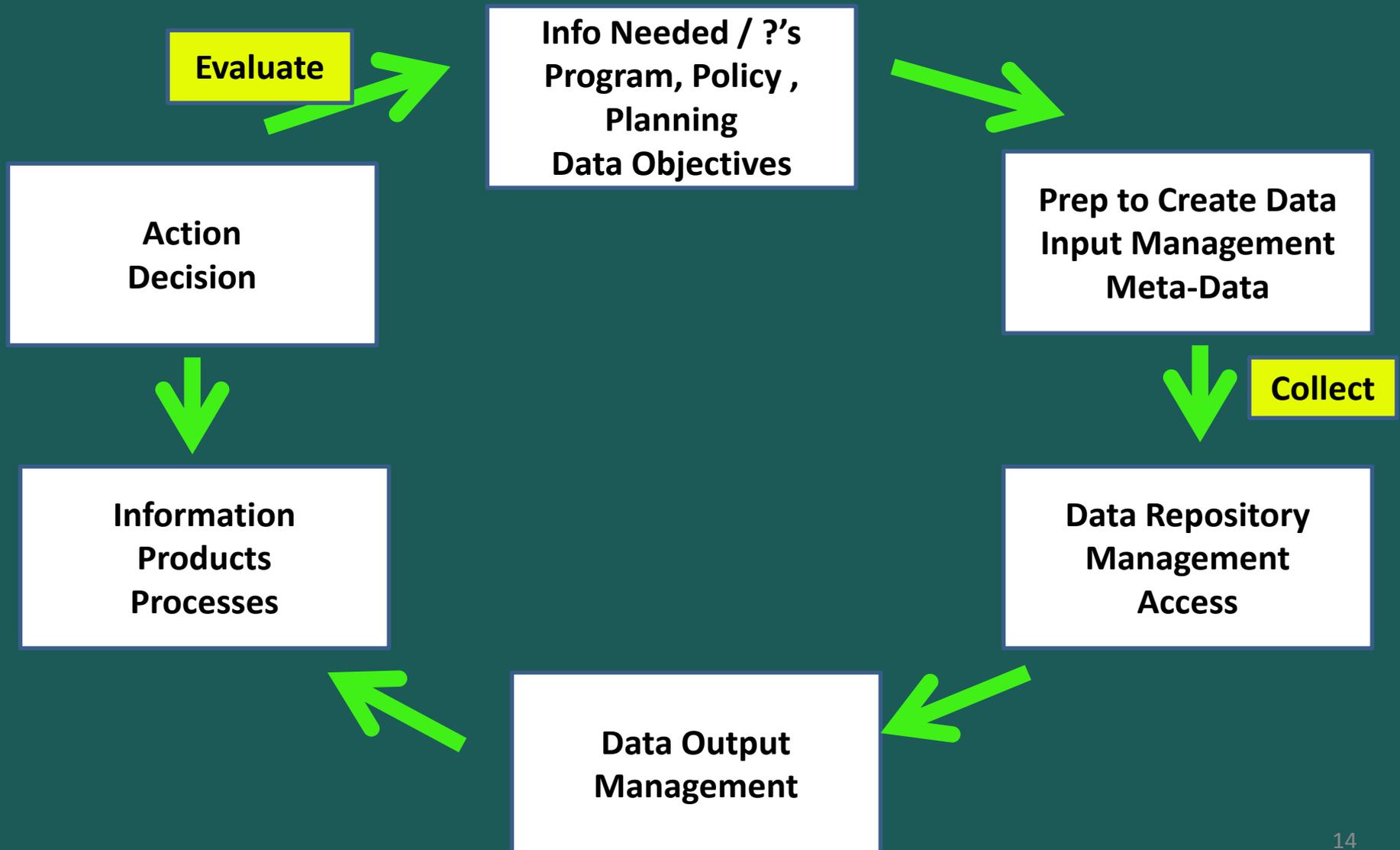
Colorado Data Sharing Network



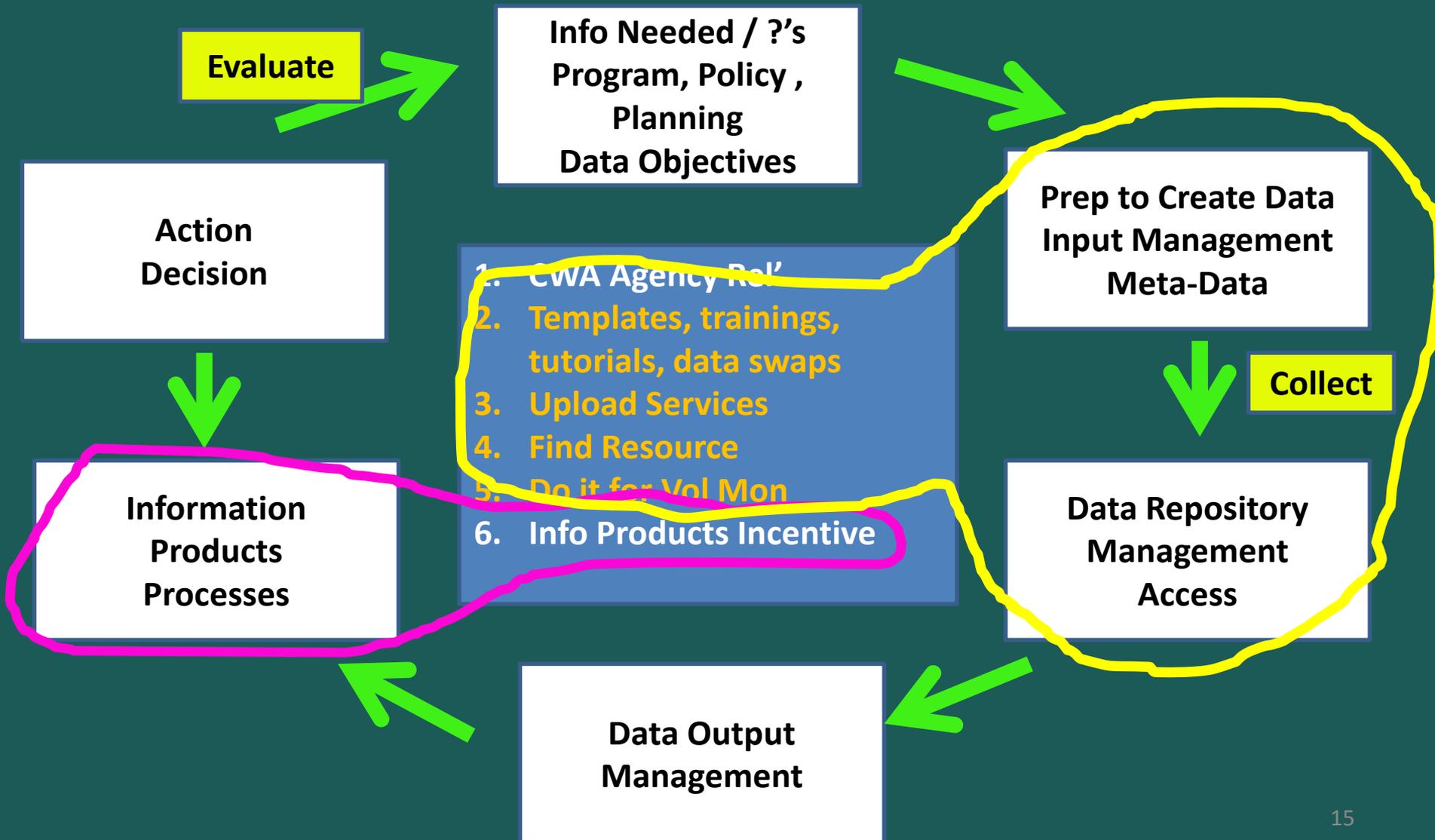
www.coloradowaterquality.org
www.coloradowaterdata.org

Data / Information Life Cycle

(for this context)



Learning Curve Understanding WQX Is Too High



How do we make data more usable?





Education/
Awareness



Problem ID,
Assess
Impairment,
Local
Decisions



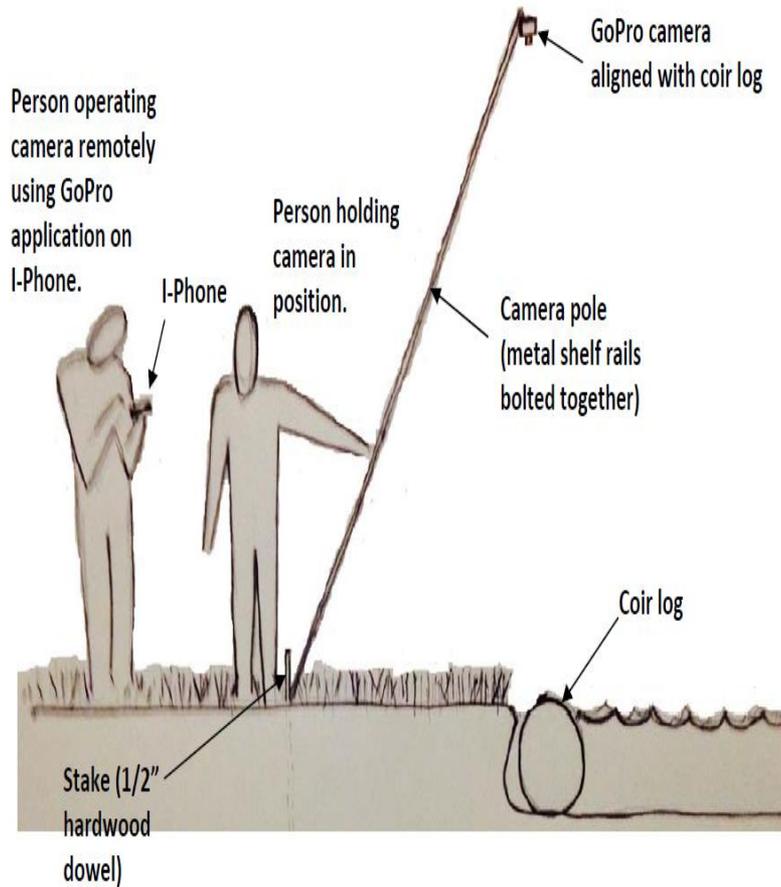
Legal &
Regulatory

Increasing Time - Rigor - QA - Expense \$\$

Graded Approach to Data Quality

- Provides flexibility for project plans
- Enables project partners to tailor the projects as needed
- Leads to an appropriate level of QA/QC

Regulatory Data Use



Big data is like teenage sex:
everyone talks about it,
nobody really knows how to do it,
everyone thinks everyone else is
doing it, so everyone claims they
are doing it...

(Dan Ariely)

FACILITATED DISCUSSION

1. The learning curve on understanding WQX is too high
2. How can we make data more usable?

Ground Rules:

- Use easel pads with table group as needed
- Avoid jargon
- Think of ways ideas *will* work before ways they *won't*

Agenda	
5 mins	Brainstorm, write ideas on post-its
5 mins	Talk to neighbor, narrow to top 3 ideas
10 mins	Discuss question with table group
5 mins	Report out by table on Q1, discussion
5 mins	Report out by table on Q2, discussion
5 mins	Vote on top ideas

Vote online: <http://bit.ly/1W6ot5T>
Comments? creswell.joel@epa.gov