How to motivate scientists to publish data online

Mark J. Costello,
University of Auckland,
New Zealand

SCOR/IODE workshop on data publication,
Oostende, 17-19 June 2008
Data is important

• Suggests and supports theories and models
• Provides facts and figures for resource management
• Its open access has economic benefits (McMahon 1996, Weiss 2002)
• Restricted and conditional access impedes use
Benefits data publication to individual scientists

Visibility brings:
recognition, invitations to collaborate, greater citation rate

Fulfils obligation to funding agencies

Satisfaction of potential wider benefit to science and society ("doing the right thing"), including repatriation to source countries

(For fuller list see Costello and Vanden Berghe 2006 open access publication at http://www.int-res.com/abstracts/meps/v316/)
But most data is not available

Despite national and international pleas and policies:

- OECD, ICSU, UNESCO, IOC, GBIF, NSF, EC, Antarctic Treaty

Exceptions?

- Molecular data in GenBank, ProteinBank
- Type specimens in museums and herbaria
Why do most scientists not publish data?

Journals will not publish primary raw data
It is my data, why should I make it available?
The data I used was not my own and I did not get permission to publish it.
If I release data then I may be scooped by somebody else producing papers from it.
I have not finished analysing the data. I may do further analysis on the data.
Somebody will use my data and benefit from such use, and worse still, they may be a commercial organisation or consultant.
The publisher may profit.

I fear that the data will be used for an incorrect purpose.
I do not have the skills to publish data on the internet
IPR related to data and databases differs between countries
I will not get due recognition for creating the data.
I have reasons I’d rather not mention
I will only publish data under certain conditions
But all reasons answered for print media!

Problem =
scientific community does not see online data as a “publication”

Need change of mindset and culture – we can help!
How? What is needed?

1. Use word data “publication” (not “sharing” or “availability”)
2. Citation of datasets must be promoted by authors, editors, data managers
3. Data publication required by referees and editors
4. Citation metrics included in online data publications
5. Government support for online data publication
Outcome?

Online data publication treated like print publications
Authors plan for data publication
Authors submit data in correct format (schema) with metadata
Authors pay charges for data publication (including quality control, archiving and integration)
Online data publication services as essential as print media
We can start this process today!

- change the way you refer to and think about online data publication

- contact journal editors and make them aware of options for data publication

Thank you

m.costello@auckland.ac.nz