Data Management & Archiving for IP3 and WC2N

Michael Allchin
IP3 / WC2N Data & Information Manager
Oh…?
Oh NO!!!
NOT
data-flaming-management?!?
Somebody shoot me.
Seriously, shoot me now.
Please...?
Huh –
this is one time when
a Blackberry really *would* be useful
Hang on – where’s my iPod?
Ho hum –
nearly coffee-time
BUT WAIT…
But wait…

Three good reasons for taking data management seriously
Three good reasons for taking data management seriously

1. We have to do it!
Three good reasons for taking data management seriously

1. We have to do it!
2. Scientific transparency
Three good reasons for taking data management seriously

1. We have to do it!
2. Scientific transparency
3. Making a contribution to the ‘big picture’
The `Bigger Picture’

US National Academy of Sciences - study of data issues:

3 pillars of ‘data husbandry’

*Nature, Vol 460, Issue 7255, July 30 2009*
The `Bigger Picture`

US National Academy of Sciences:
3 pillars of ‘data husbandry’

1. Integrity
The `Bigger Picture`

US National Academy of Sciences: 3 pillars of ‘data husbandry’

1. Integrity
2. Access
The `Bigger Picture`

US National Academy of Sciences: 3 pillars of ‘data husbandry’

1. Integrity
2. Access
3. Stewardship
One really good reason for taking data management seriously

It will make your life easier when developing / testing / running models
So –

What’s The Plan?
The Plan: IP3

1. Establish primary archive
The Plan: IP3

1. Establish primary archive
2. Load validated data
The Plan: IP3

1. Establish primary archive
2. Load validated data
3. Provide search / retrieval tools
The Plan: IP3

1. Establish primary archive
2. Load validated data
3. Provide search / retrieval tools
4. Export to models
The Plan: WC2N
The Plan: WC2N

errmmm...
The Plan: WC2N

- Different types of dataset
The Plan: WC2N

- Different types of dataset
- Some in very large volumes
The Plan: WC2N

- Different types of dataset
- Some in very large volumes
- Permanent archive destinations already established for some
The Plan: WC2N

Options...
The Plan: WC2N

Options:

- Primary archive
The Plan: WC2N

Options:
- Primary archive
- Metadata catalogue
The Plan: WC2N

Options:
- Primary archive
- Metadata catalogue
- ‘Hybrid’
So What Should I Do Now?
So What Should I Do Now?

Go home and prepare your data for archiving!
Data-Preparation: Good
Data-Preparation: Not Quite So Good
“Data professionals should be recognized for their crucial role in stewardship: certainly they deserve more respect and support than researchers sometimes give them.”

US National Academy of Sciences - Report on Scientific Data Management
June 2009
Amateurs talk theories -
Professionals talk data!