A CWS Data Security Model

Presentation to OFWIM
September 29, 2014

Presenter: Eric Woodsworth
Canadian Wildlife Service
Rationale

• Sensitive data: Data the release of which could reasonably be expected to cause injury to the interests of persons, institutions, or nations, or to environmental assets such as species and habitats.

• Organizations need a policy and mechanism to identify sensitive data and to design systems to control access, even if a small proportion of the institutional data is sensitive.

• Creates a way to consistently identify and protect our most sensitive environmental assets in the face of strong demand for access to information.
Rationale …

• The logjam clears: non-sensitive data are disseminated quickly, sensitive data are consistently and appropriately protected.

• Public trust is improved:
  – Access to tax-funded data
  – Responsible stewardship of sensitive species, consistent with the CWS mandate.
Why now?

• External Demand
  – Open Data
  – Open Government
  – Until recently, no obligation to share science data with the public, other than through Access to Information.

• Partnerships
  – New data sharing agreements

• Redesigning our data system to serve all clients

• It all requires overarching policy for the wildlife context
The Basics

• Rules are needed why?
  – Reflect organizational values/policy/legislation
  – Policy helps to establish working culture, provides a starting point for evolution of ideas.
  – Consistent application of policy reduces end-runs and one-offs
    ▪ No second chances
  – Improved ability to defend denials of access (Access to Information Act)

• Key questions:
  – What will we share?
  – With whom?
  – In what form?
  – Based on what facts?
Design Tasks

- Define components of sensitivity
- Define data groups having consistent sensitivity
  - Variations by human effects, geography, accessibility
- Define user classes
- Define modification/generalization methods
- Link to government security classification
- Define system permissions, at the record level
- Generation of metadata on access/use restrictions
Proposed Model
Generalization

• Reduced geographic precision (e.g. .1, .01, .001 degrees) – Chapman & Grafton 2008
• Reduced taxonomic precision, habitat info, observer’s name, traditional uses: i.e. limit data mining
• Guidance in CH ID Toolbox.
Consultations

• With SAR experts and IM/IT management committee…
• Are these the right decision factors?
• Parse each factor:
  – Restrictions
  – Data groups (sub-/supra-specific)
  – User classes
  – Generalization methods/levels
• Parameterize the table with contrasting datasets; reparse factors until stable.
• Legal review of restriction criteria
• Define a priority species list for implementation
• A “no ecological sensitivity” list → rapid access
Considerations

- **Species at Risk Act (SARA)**
  - Critical habitat is defined as “The habitat that is necessary for the survival or recovery of a listed wildlife species and that is identified as the species’ CH in the recovery strategy or in an action plan for the species. “

- **Advice from Legal Services, 2007**
  - Critical habitat is the habitat that is identified in a recovery strategy or action plan that is included in the Public Registry. Therefore, if the decision is taken to *not include Critical Habitat* in a recovery document using s. 124 of SARA because of threats to the species and/or its habitat, there is effectively no Critical Habitat for the purposes of SARA.
  - Hence no legislative protection under SARA
Balance of Risks

• Risk of intentional destruction based on released information
  – Retaliatory killing or habitat destruction by irate landowners
  – Data mining based on stenotypic habitat requirements
  – Market value & poaching

• Risk of unintentional destruction due to insufficient information
  – Exclosures
  – Routing of trails, roads, pipelines
  – Development easements

• Hence need an ongoing review cycle.
Q & A

• How do you make data sharing decisions in your organization?
• Can you recommend improvements to the model?