

# NISS

## Microdata Tool Kit: Prospectus

Alan Karr  
March 3, 2005

---

---

---

---

---

---

---

---

### Background: Current NISS research on SDL for numerical microdata

- 8 methods, each with “settable parameters”
  - Rank swapping
  - Resampling
  - Addition of noise
  - 5 forms of microaggregation
- 3 utility measures
  - KL distance between original and masked data
  - Confidence ellipse for a single regression
  - Confidence intervals for a single regression
- 2 disclosure risk measures
  - % of correctly linked records
  - % of correctly linked records, weighted by distance

---

---

---

---

---

---

---

---

### What Would the MDTK Do?

- Computational experiments similar to—but more complicated than—those done by the DSTK
  - User selects
    - Methods and one or more parameter values for each
    - Utility measure(s)
    - Risk measure(s)
  - MDTK
    - Computes risk and utility measures
    - Calculates frontier
    - Visualizes results

---

---

---

---

---

---

---

---

## How Would it Be Structured

- Modular and extensible, to allow “plug in” of
  - New SDL methods
  - New risk measures
  - New utility measures
- Like DSTK, consisting of engines run by script (specs) files that can be generated
  - Manually
  - Via GUI

---

---

---

---

---

---

---

---

## What Would it Take?

- Good code in place for
  - SDL methods
  - Existing risk and utility methods
- Big issue is the extensibility (APIs, data structures, inter-component communication, ...)
- High-end approach
  - Right professional programmer working 4-6 months
- More modest approach
  - Professional software architect to create core
  - Student programmers to do GUIs, viz, ...

---

---

---

---

---

---

---

---

## The “Trailer”

- Some combinations of methods work better than either component method alone
  - Example: Micmul03 followed by noise\_70
    - ~3X reduction in KL data distortion compared to Micmul03 alone
    - >25X reduction in disclosure risk compared to Micmul03 alone
- So, MDTK might also allow combinations!

---

---

---

---

---

---

---

---