

# The Brookings Institution

Metropolitan Policy Program



**National Infrastructure for Community Statistics: An Overview**

**dg.o2005**

**Birds-of-a-Feather Session**

**May 17, 2005**

# The Urban Markets Initiative

- Three-year effort supported by Living Cities
  - collaborative of 15 nonprofit, private and public investors committed to the revitalization of America's urban centers
- Improve quality and use of data and information available for urban areas, in order to:
  - Allow private and public decision-makers to identify untapped investment opportunities
  - Improve the quality of investment decisions
  - Increase the competitiveness of urban areas as places for investment

# National Infrastructure for Community Statistics: What

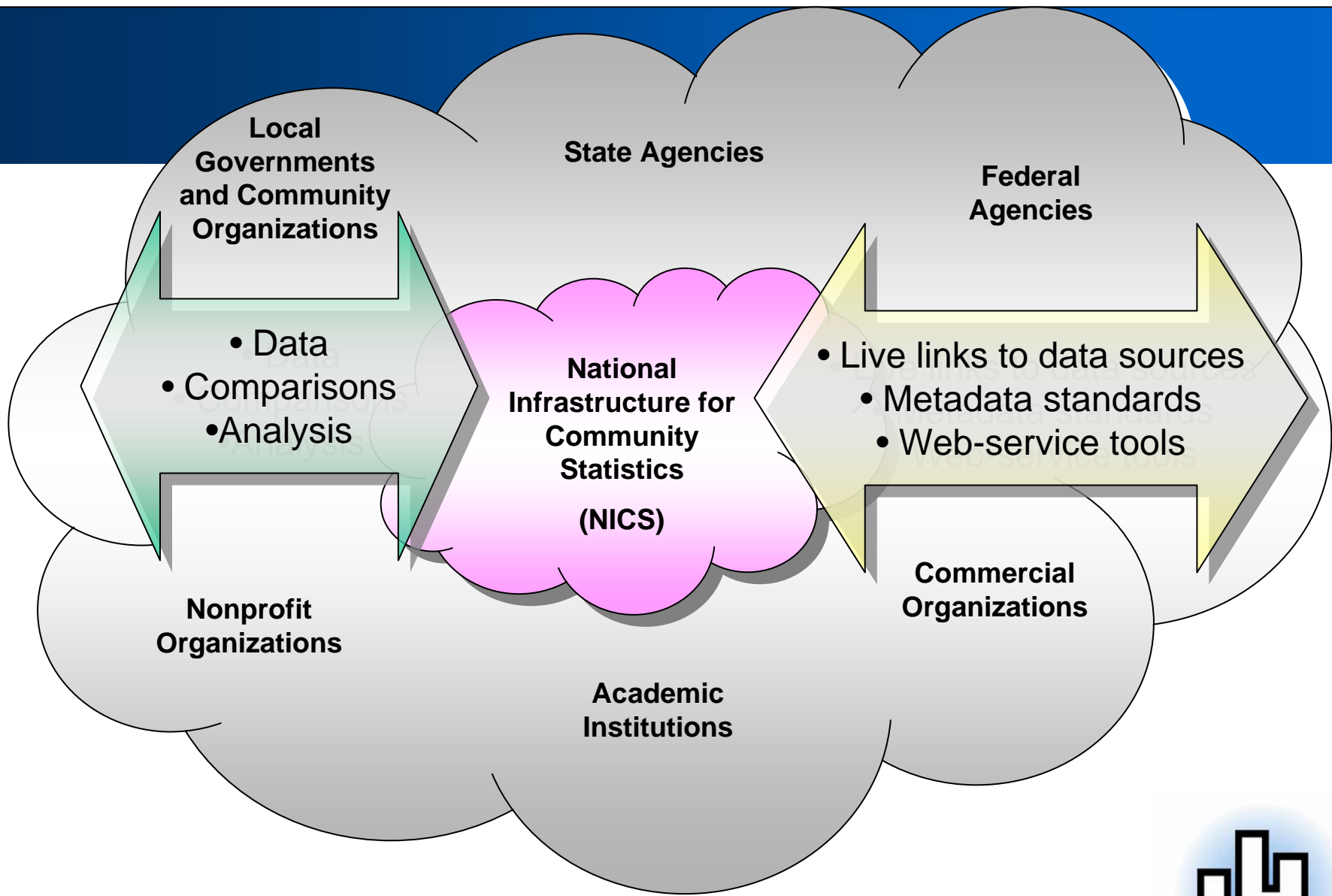
- Community-level data (metro, place, neighborhood, parcel):
  - from multiple sources (federal, national nonprofit, state, local, and commercial)
  - that can be easily combined, compared and analyzed
  - across domains and jurisdictions
  - for better decision-making

# NICS: How

- Web-based “marketplace” providing access to community-level data and tools with which to integrate, view and analyze these data
- Marketplace: a forum to exchange data, services and products. A place to match demand for and supply of:
  - Data: 1000s of community-level datasets
  - Tools and Services: to access, manage, combine, clean, compare, analyze, and present community statistics
- Service-oriented architecture

# NICS: Who

- NICS is intended as a “wholesale” resource for organizations involved in “retail” data-related activities serving broad audiences (e.g., data intermediary, federal program analysis).
- Participants – federal/state/local governments, research institutions and other nonprofits, commercial
  - Data providers and collectors
    - Including federal and state agencies with mandate to gather and analyze administrative data from local governments
  - Data intermediaries
  - Program agencies
  - Middleware and tool providers



### Community Data Users

- Municipalities
- Metro Planning Orgs.
- Community-based Orgs.
- Data Intermediaries
  - CSS Network
  - NNIP
  - Census Info Centers
- Indicators groups--poverty, sustainability, asset-building

### State Agencies

- NGA
- State CIOs (NASCIO)
- State DHS, Health, Jobs
- State Budget Offices (NASBO)
- State Data Centers (Census)
- State Archives

### Federal Agencies

- Federal Statistical Agencies (FedStats, Census, BLS, BEA, NCHS, etc.)
- Federal Program Agencies (e.g., EDA, ETA, FHWA)
- Federal Management Orgs. (OMB, GAO, CIO Council)

## National Infrastructure for Community Statistics (NICS)

### Foundation/ Investor Users

- Outcomes data
- Comparative data
- Performance measurement
- Success Measures
- SIA Models

### Academic

- University analysts
- University service providers

### Nonprofit

- DataPlex (Fannie Mae Foundation)
- KNII (National Academy of Sciences)

### Commercial

- Data Providers
- Value-added Data Intermediaries
- Market Research
- Analysts
- Services

# Potential Impacts of NICS

- Greater understanding of community socioeconomic and geophysical conditions, trends and opportunities
- More effective program and investment decisions
- Improved measurement of program and investment impacts
- Enhanced local and national indicators efforts



# Potential Impacts of NICS

- Expanded base of users of federal statistics
- Significant enhancement of federal ability to access locally produced data
- Greater linkage among multiple federal efforts regarding community statistics and mapping
- Enhanced federal and state investments in e-government initiatives

# NICS Community of Practice

- Mission – promote the development of NICS
- Involves over 100 representatives from
  - Federal agencies (e.g., OMB, Census, EPA, GSA, USGS/FGDC, FedStats, BJS, IRS, HUD, HHS, GAO, Federal Reserve, FDIC)
  - State agencies (e.g., Maryland Department of Planning)
  - Local organizations (e.g., City of Baltimore, Metro Chicago Information Center, National Association of Counties)
  - Academia (e.g., MIT, University of Florida, University of Memphis)
  - National nonprofits (e.g., Fannie Mae Foundation, Joint Center for Political and Economic Studies, Urban Institute)
  - Commercial firms (e.g., Claritas, Maya Design)

# NICS Process

- Concept Phase (May-August 2004): Develop NICS concept, create CoP
- Learning Phase (September 2004-February 2005)
  - Four workshops to ascertain desirability and feasibility of NICS from perspective of local, state, federal, private, nonprofit organizations
  - \$250,000 raised for Initiation Phase
- Initiation Phase (March 2005-June 2006)
  - Business plan
  - Services development
  - Organizational development (surveys, fundraising)
  - Prototype web gateway
- Implementation Phase (June 2006 - )

# Participant Needs: Motivating Examples

Local policy and program analysts want access to data at small areas of geography, for uses such as determining:

- **Neighborhood crime/health-risk rates:** number of incidents / population at risk
- **Program eligibility:** what percent of households meet income cutoff given size, local rent levels, etc.
- **Neighborhood (dis)investment:** what parcels are ripe for reinvestment, what federal \$ are going into a neighborhood
- **Employment and resident patterns:** where do local residents work (by skill level); what jobs are accessible; how has this changed over time

# Participant Needs: Problems

Explosion of fine-grained, spatially detailed data is great, **but:**

- Hard to **get/keep/maintain local copy** of all the needed/useful data
- Too hard to **get/understand/use metadata**
- Cross-domain **geographies are different** so small-area data are hard to overlay well
- Hard to add **local knowledge to reinterpret** government datasets
- **Privacy and confidentiality** rules limit access to small-area data

# Participant Needs: Data

- Data for small areas + metadata
  - Recognize much data exists, particularly administrative, but not now accessible
  - Recognize importance of data standards, good metadata, tools to facilitate access
- Data topics, highest priority for state/local
  - Economic development, community development/real estate, program eligibility determinations, health, public safety, environment, infrastructure
  - Reliable denominators for small area indicators

# Participant Needs: Tools

- Small area estimation models/tools
- Tools and methods to address confidentiality
  - For example, synthetic estimates for small areas
- Mapping tools to bridge different geographies and permit comparisons over time
  - For example, police precincts, planning districts, census tracts
- Data entry/publishing tools to allow input of data into system with quality controls

# Participant Needs: Services

- Means to develop and maintain metadata
- Managing and operating distributed data systems
- Managing organizational and legal issues (e.g., data sharing, intellectual property)



# Promising Developments

- Metadata
  - Improved metadata standards, tools, best-practices
  - Progress with semantic web, ontologies, RDF, etc.
- Service-oriented architecture
  - Shared data services instead of exchanged data sets
  - Loosely coupled, distributed components (not tight coupling with 'broadcast' model)
  - Growth of web services with XML messaging and interoperable geospatial services
- Analytic assistance and user interfaces
  - Codifying 'business rules,' data mining, and analytic processes
  - Synthetic data, custom 'rollups,' and trusted intermediaries
  - Improved user interfaces for statistical analysis

# Participants Bring Resources

- State/Local
  - Data from local systems
    - More multi-source systems have been developed although need to develop intermediaries many more places
  - Networks to widen NICS participation
- Federal
  - Data from national files
  - Existing state/local partnerships networks
    - e.g., Labor Market Information agencies, State Data Centers, Census Information Centers
  - Guidance on standards
    - e.g., meeting requirements for GASB 24, OMB's Data Reference Model

# Sobering Realities

- Hard to switch focus from data sets to data services
  - New paradigm, immature technologies
- Community statistics involves analysis
  - Appropriate re-use of data is hard (not just data access)
  - NICS audience wants time series (as well as cross-sectional studies)
- Data integration for small-area geographies is hard
  - Serious privacy/confidentiality limits
  - Stored data <> delivered data – lots of on-the-fly processing
- Separate GIS and statistics communities

# To Address These Needs

- NICS will provide a marketplace providing access to community-level data and tools with which to integrate, view, and analyze these data
- Utilizing a service-oriented architecture format

# Service Categories: Data and Tools

- NICS-ready datasets
  - Services such as FedStats, DataPlace, federal statistical agencies, local data intermediaries, and more
  - Directories of administrative records available for public use
  - Services to meet minimum metadata standards
  - Provide information on data quality
- Data transformation tools
  - Cross-referencing tools
  - Façade tools
  - Standardization tools
  - Geographic roll-up tools
  - Synthetic data tools to ease analysis of sensitive or confidential datasets

# More Tools

- User interface tools
  - Display software to ease comparisons
  - Usability tools
  - Visualization tools
  - Aesthetics tools
- Analytic tools
  - Tools to determine eligibility for federal funds
  - Small area estimation tools
  - Automated computation tools to calculate margins of error
  - Microdata analytic tools
  - Indicator development tools, such as profile templates that generate standard indicators
  - Comparability analysis tools
  - Statistical literacy tools

# And More Tools

- Resources for building and operating distributed systems
  - Management
  - Operation
  - Service-chaining
- Resources for managing organizational and legal issues
  - Data-sharing or exchange MOUs
  - Data access protocols
  - Intellectual property management
  - Privacy and confidentiality

# And Finally

- Metadata tools and services
  - Model components of metadata
  - Hints for metadata development
  - Library of metadata standards
  - Metadata applications
  - Automated tools for attaching metadata



# Initiation Phase: Key Tasks

- Prepare business plan
- Complete market surveys that assess:
  - Broad range of information users
  - Decision-makers who use information
- Stimulate the creation of NICS services, and demonstrate NICS feasibility, through “use cases”
- Develop approach to creating metadata
- Develop initial web presence developed and prototype “gateway”
- Continue to educate/update/involve/expand CoP
- Raise implementation phase funds and secure institutional sponsorships as required to move into implementation

# Use Case Examples

- LEHD
- Pittsburgh-KNII prototype
- EPA Region 4 Technology Demonstration
- Local Virtual Data Warehouse efforts
  - Boston, Memphis, Indianapolis, Chicago
- Statistical Knowledge Network
- CDC Environmental Public Health Tracking Network
- Fannie Mae Foundation's DataPlace

# For Further Information

For NICS CoP events agendas, presentations, and participants:

<http://colab.cim3.net/cgi-bin/wiki.pl?NationalInfrastructureforCommunityStatistics>

Comments, questions:

Andrew Reamer, Deputy Director  
Urban Markets Initiative  
Metropolitan Policy Program  
The Brookings Institution  
1775 Massachusetts Avenue, NW  
Washington, DC 20036  
(202) 797-4398  
areamer@brookings.edu