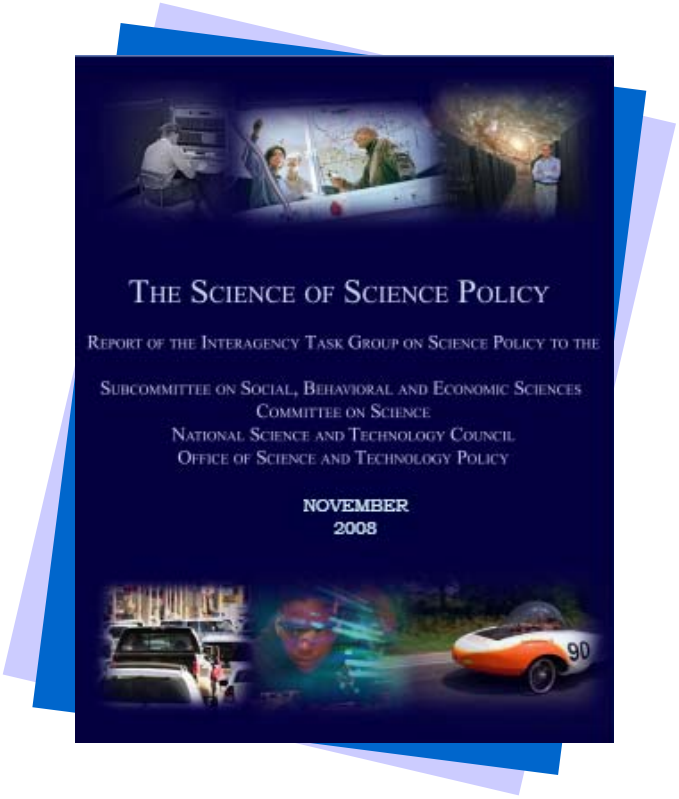
The STAR (Science and Technology in America's Recovery) Project And Beyond

Background: Roadmap

- Science of Science Policy Interagency Task Group
- The SoSP Roadmap
  - Published in November, 2008
  - Four guiding themes
  - Ten key questions
- December, 2008 Workshop
  - Engage the current community of practice
  - Interactive evaluation of Roadmap



Background: Workshop Results

- Finding: There currently is no data infrastructure that meets the Federal need for evidence based science policy decisions
- Recommendation: Create such a data infrastructure
- Action: Identification of set of investments for new administration

Building an Empirical Framework for Science Policy

- Start with correct unit of analysis
  - Science is done by scientists. Need to identify universe of individuals funded by federal agencies (PI, co-PI, RAs, graduate students etc.)
- Include full description of input measures
- Include full description of outcomes (economic, scientific and social)
- Combine inputs and outcomes
- Create appropriate metrics that capture all dimensions of science investments

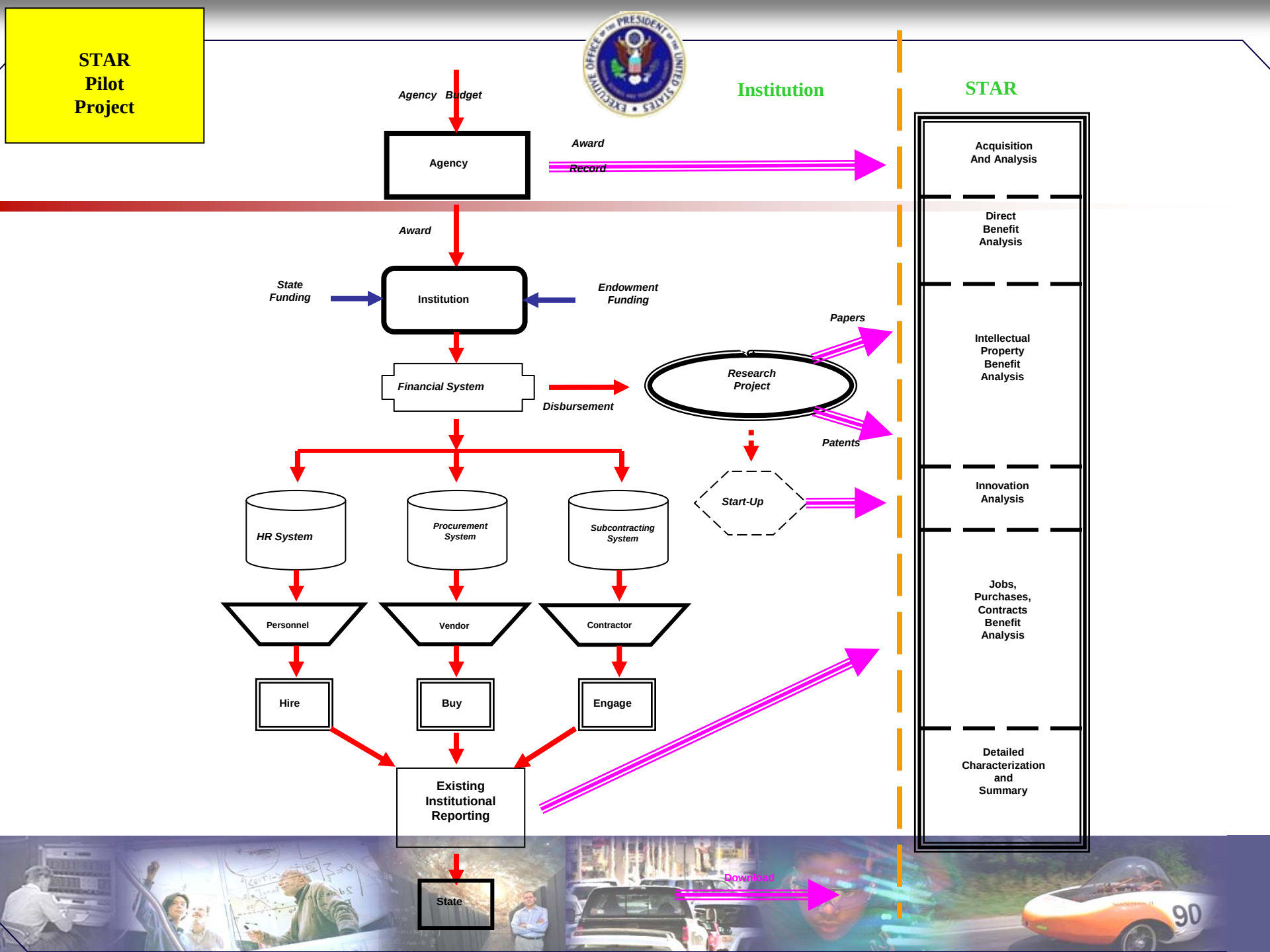
The Basic Approach: Organizing Principles

- Operational
  - Minimum Burden: *Leverage existing information at academic institutions and agencies; do not interfere with existing systems*
  - Minimum Cost: *Leverage existing investments in administrative records, patent and citation data*
  - Full confidentiality protections: *Best practices*
- Analytical
  - Utilize existing investments in data: *Statistics of Income*, *Webscraping; Visualization*
  - Create open and transparent approach to knowledge creation: *Collaboratory*

Initial Application: ARRA Reporting

- Automatically generate job creation measures
  - Create administrative tracking system
    - Existing payroll management systems
    - Unemployment insurance wage records
  - External validation and accountability
    - Credible researchers
    - External tagging
- Evidence based full impact metrics
  - Short term (initial direct and indirect job creation impact)
  - Match to other sources

STAR Pilot Project



Request from universities

- Transaction based file – every month
  - Personid
  - Internal Award number
  - Date
  - Amount Paid
  - Hours
  - Full/time part/time
  - Occupational classification (sometimes on separate file)
- Cross reference table
  - Internal award number
  - Agency
  - Agency Award number
  - Grant short title
  - Grant type
  - CFDA codes

Current Status

- Partner with FDP sites
  - Provide administrative records
  - Partner in development of metrics
- Seven have been visited:
  - Delaware, Pennsylvania, George Mason, Texas (Austin), Alabama (Tuscaloosa), CalTech, Massachusetts
  - Texas, GMU, Delaware, Caltech and UA data have been provided and analyzed – preliminary report available

The Following Is What Our Volunteer Grant Recipient Communities Data Has Revealed (Preliminary)

Initial results: Example of University XX

Table 1a: Overall Summaries			
	2009-04	2009-05	2009-06
# of Awards	3,807	3,802	3,655
Total Awards $	$67,297,858	$65,558,643	$59,677,284
# Employed by Awards	19,638	19,225	15,244
# of Science Awards (CFDA code = 47, 93, or 81)	514	509	583
Total Science Awards $	$2,294,936	$2,209,999	$4,575,101
# People Funded by Science Awards	1,018	1,002	1,338
# Est. FTE Equivalents by Science Awards	851	838	1,187

Table 2a: Employment Patterns

Employment Pattern	Total Employed April 2009	Total Employed May 2009	Total Employed June 2009	Note
April_May___June_	13,522	13,522	13,522	Continuous
April_May_____	5,237	5,237		
April_____June_	185		185	
April_____	694			
_____May___June_		291	291	Job "created" in May and June
_____May_____		175		Job "created" in May
_____June_			1,246	Job "created" in June
Grand Total	19,638	19,225	15,244	

Table 2b: Employment Patterns for Science Award Recipients

Employment Pattern for Award Recipients	Total # Recipients April 2009	Total # Recipients May 2009	Total # Recipients June 2009
April_May___June_	699	699	699
April_May_____	248	248	
April_____June_	20		20
April_____	51		
_____May___June_		37	37
_____May_____		18	
_____June_			582
Grand Total	1,018	1,002	1,338

Results so Far

- Have identified systemic and idiosyncratic data issues
  - Relatively clean data
  - Some anomalies that will need to be coded up in next stage
- Response of pilot universities
  - Doable
  - Very supportive and responsive
- Briefing of Ed Deseve and Frank DiGiammarino

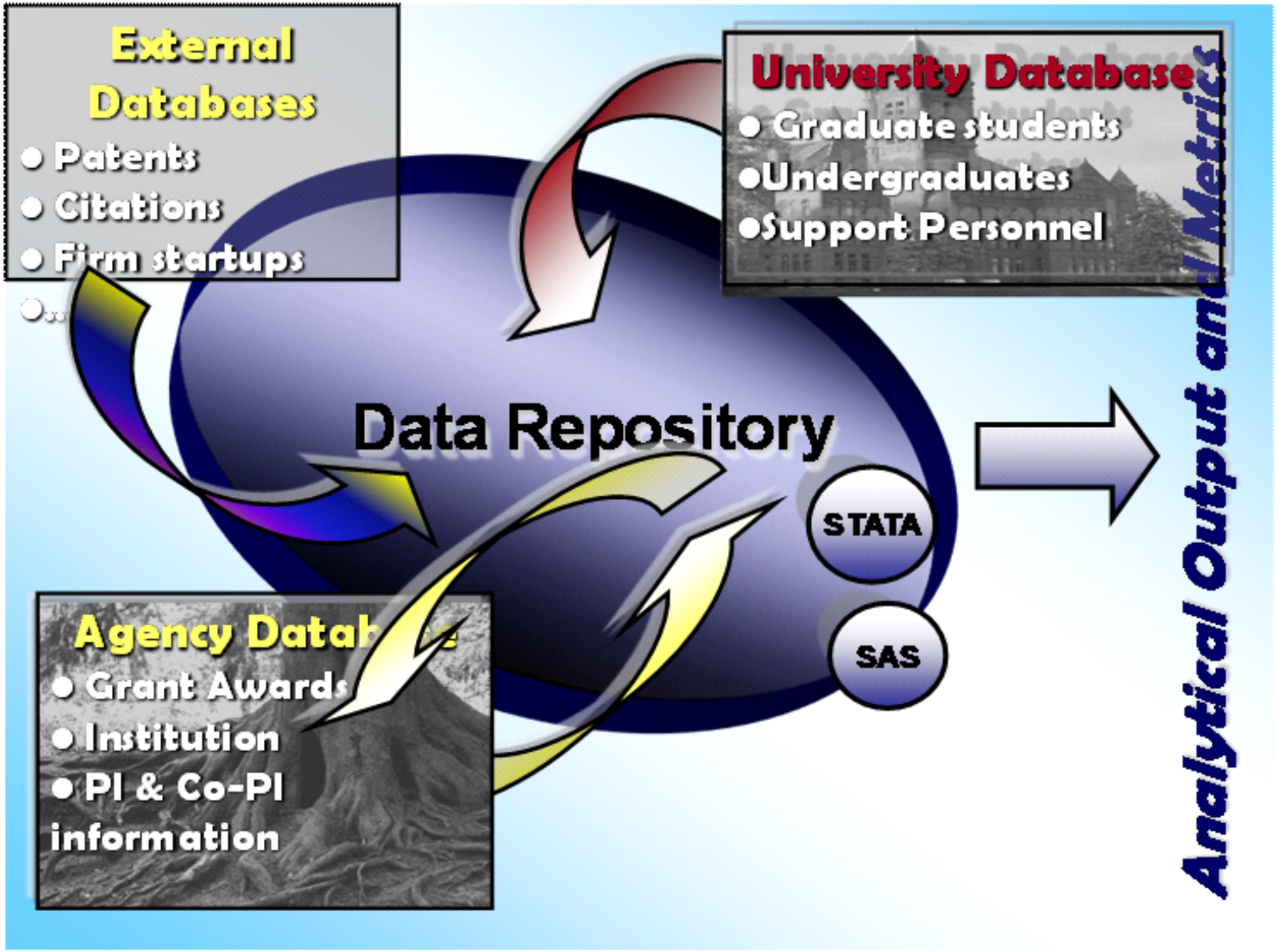
– Enthusiastic reception
– Noted that the approach is exactly in line with the ARRA Section 1512 guidance, which calls for recipients to use payroll and finance data to do their estimates
– Full agreement with pilot approach and timeline

Next set of requests

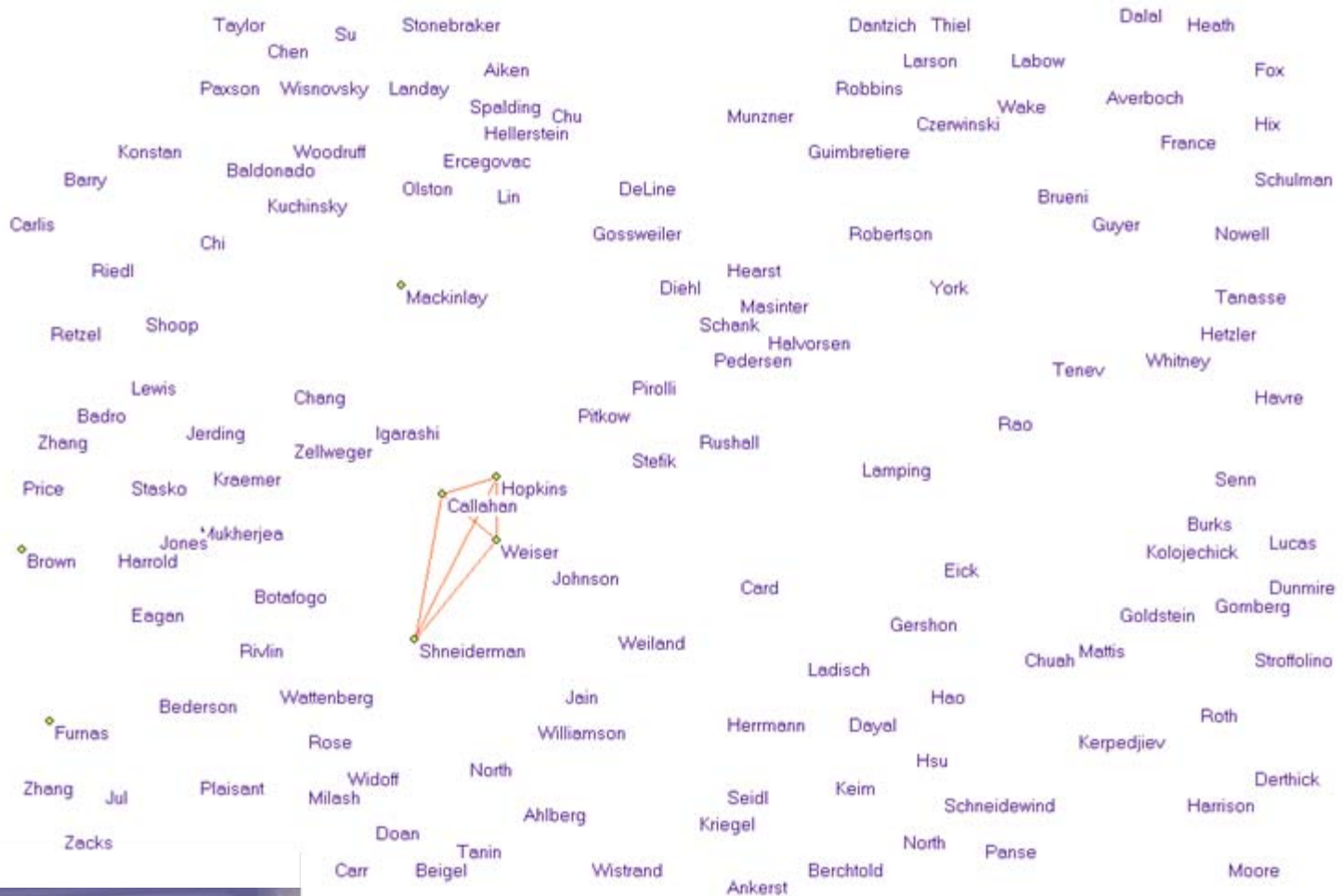
- Mapping to standardized occupational categories
  - Teaching faculty, research faculty; postdocs; grad students; undergrad students; tech support; administrative support
- Universities will send employment records for the twelve months previous to the current month.
- Report on indirect costs to cognizant agencies
  - so we can allocate overhead to labor costs
- Standardization of report format

Glimpse of the future: Building on the framework

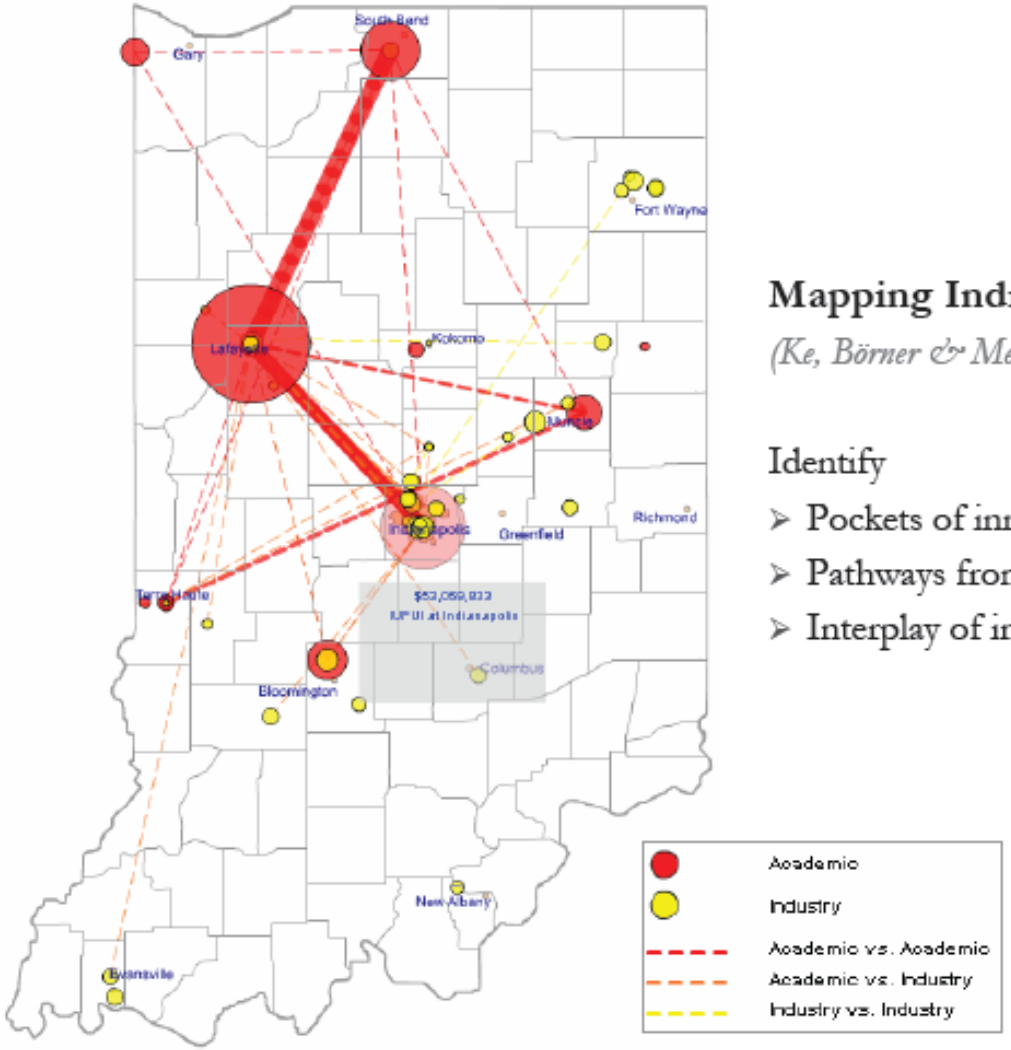
- Framework: a collection of integrated databases
  - Agency records transmitted on a flow basis
  - University records transmitted on a flow basis
- Reduce Burden on PI's and Universities
  - Automated webscraping and reporting of outcomes to agencies, state legislatures and other constituencies
  - Systematized, standardized and validated ongoing measurement of long term impact of science
    - Economic: Patents, patent applications, new businesses
    - Scientific Outcomes: Creation and uptake of ideas: e.g. citations, new fields
    - Social outcomes: Health, welfare, environment…


External Databases
Patents
Citations
Firm startups
...
University Database
Graduate students
Undergraduates
Support Personnel
Data Repository
STATA
SAS
Agency Datab
Grant Awards
Institution
PI & Co-PI information
Analytical Output and Metrics

Examples of Research Possibilities



Modeling Policy Outcomes: Mapping Innovation Pathways



Mapping Indiana's Intellectual Space

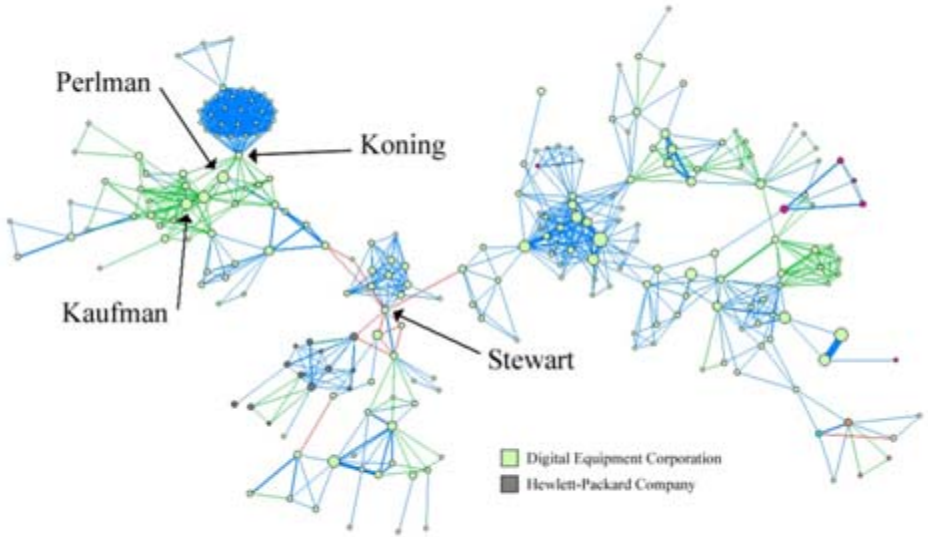
(Ke, Börner & Mei, 2005)

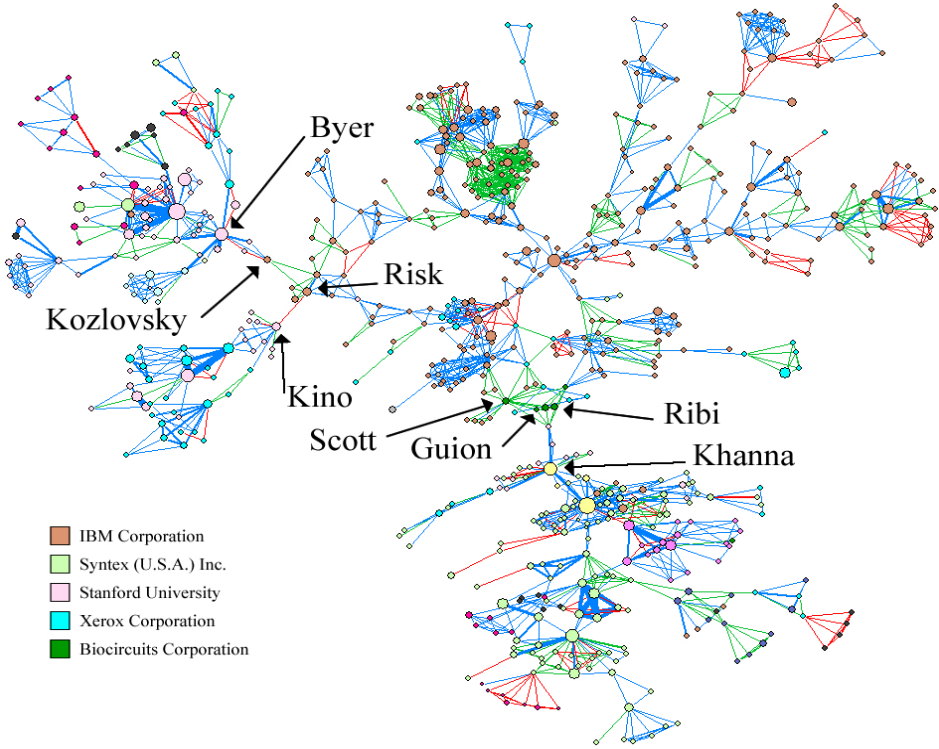
Identify

- Pockets of innovation
- Pathways from ideas to products
- Interplay of industry and academia

Motivation: regional advantage

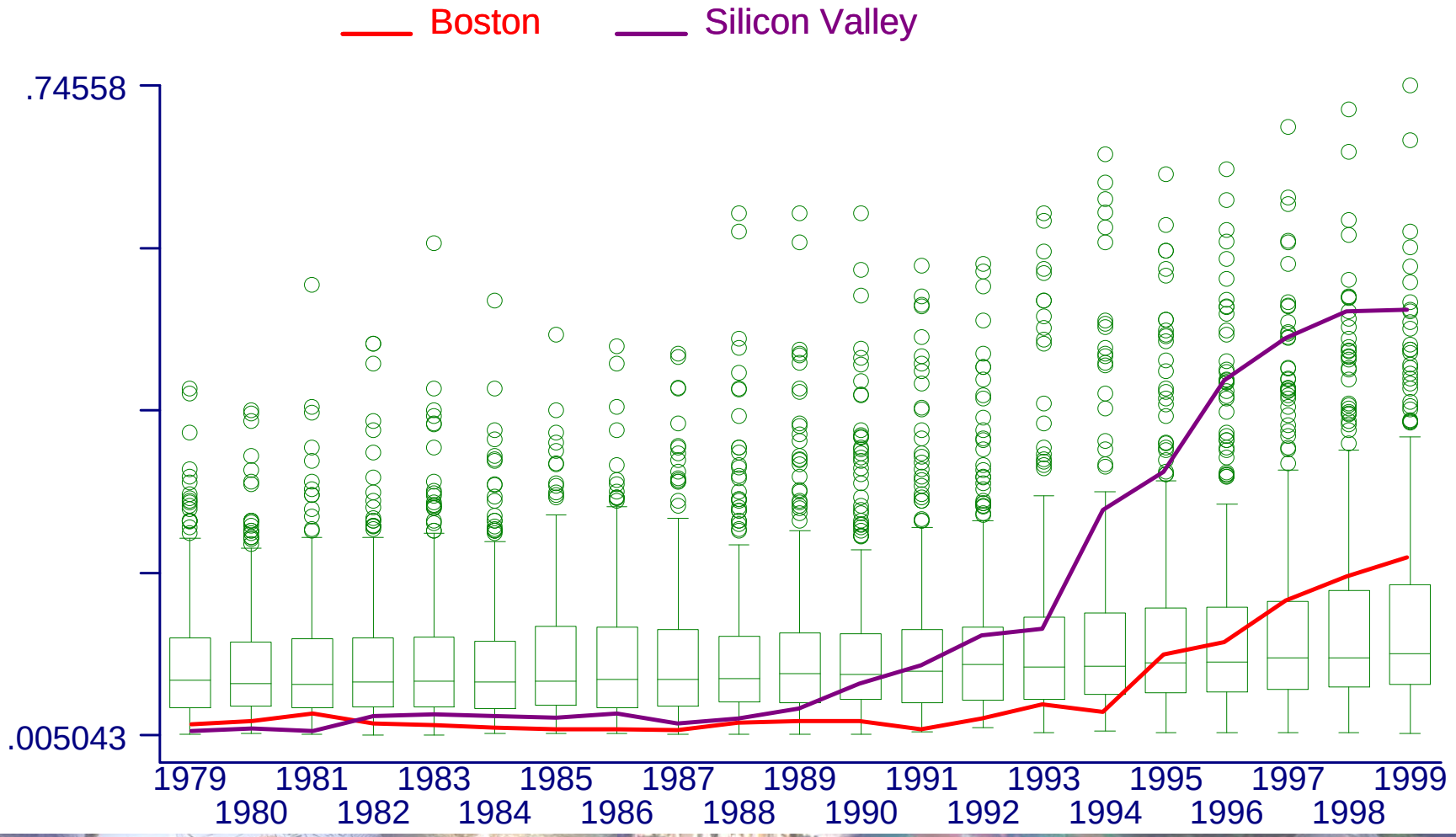
- Why is Silicon Valley successful?
  - What is the role of funding and ocial networks?
- Large sample data source: All U.S. Patents, 1975-present

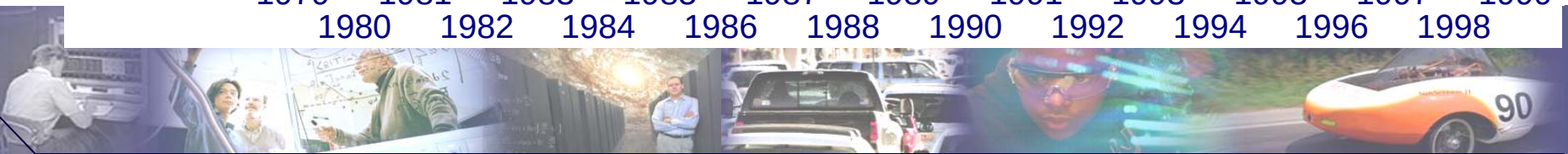




Despite similarities, Boston and Valley structures diverge in 1990s

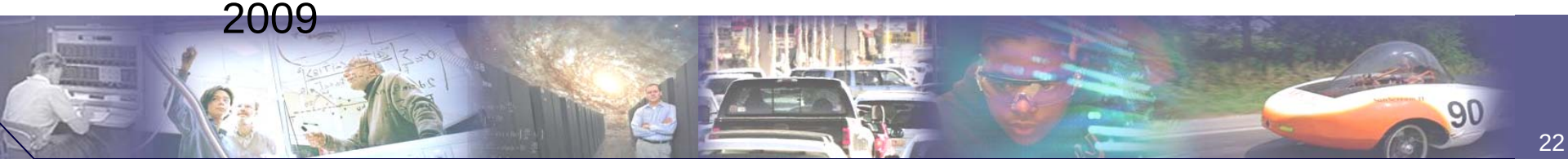
Proportion of MSA Inventors in Largest Component



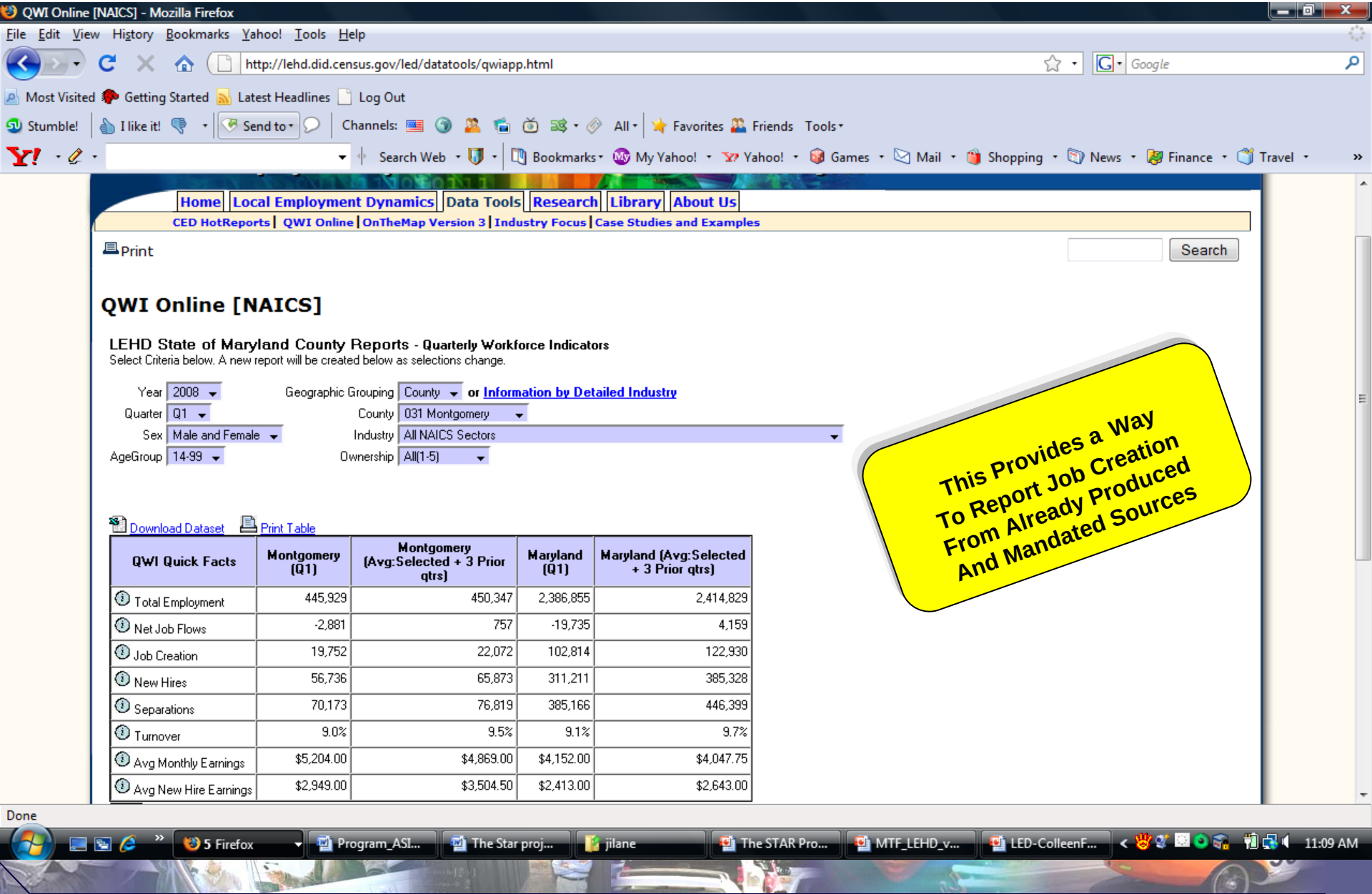


Summary and Next Steps

- Current:
  - Actual, administratively based, externally verifiable, measures of job creation for six universities
- Future:
  - Participating universities will send one full year of data, standardized occupations, in a standardized format to STAR team to start prototype production
  - Additional universities invited to participate in scaling up of job creation components
  - Pilot universities invited to participate in matching exercise with citations, patents and other economic/scientific outcome metrics
  - Report to OSTP and participating agencies scheduled for March 2009

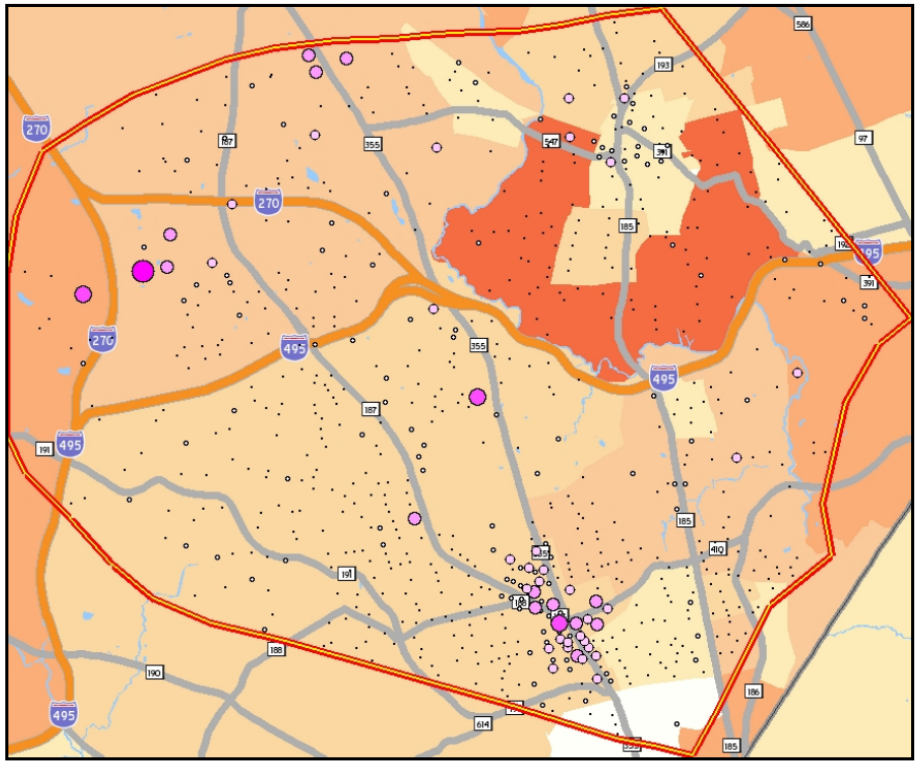


Prior Experience



QWI Quick Facts	Montgomery (Q1)	Montgomery (Avg:Selected + 3 Prior qtrs)	Maryland (Q1)	Maryland (Avg:Selected + 3 Prior qtrs)
Total Employment	445,929	450,347	2,386,855	2,414,829
Net Job Flows	-2,881	757	-19,735	4,159
Job Creation	19,752	22,072	102,814	122,930
New Hires	56,736	65,873	311,211	385,328
Separations	70,173	76,819	385,166	446,399
Turnover	9.0%	9.5%	9.1%	9.7%
Avg Monthly Earnings	$5,204.00	$4,869.00	$4,152.00	$4,047.75
Avg New Hire Earnings	$2,949.00	$3,504.50	$2,413.00	$2,643.00

Work Area Profile Report



This map is for demonstration purposes only. For a more detailed and customizable map ouput, please use the "Print Map" tool located above the Map Viewer.

2006
- 1 - 44 Workers
- 45 - 358 Workers
- 359 - 1210 Workers
- 1211 - 2870 Workers
- 2871 - 5605 Workers
- 5606 - 9687 Workers

Total Primary Jobs	2006	
	Count	Share
Total Primary Jobs	84,169	100.0%

Jobs by Worker Age	2006	
	Count	Share
Age 30 or younger	22,046	26.2%
Age 31 to 54	49,297	58.6%
Age 55 or older	12,826	15.2%

Jobs by Earnings Paid	2006	
	Count	Share
$1,200 per month or less	15,703	18.7%
$1,201 to $3,400 per month	27,189	32.3%
More than $3,400 per month	41,277	49.0%

Jobs by Industry Type (2-digit NAICS)	2006	
	Count	Share
Agriculture, Forestry, Fishing and Hunting	73	0.1%
Mining, Quarrying, and Oil and Gas Extraction	8	0.0%
Utilities	29	0.0%
Construction	5,127	6.1%
Manufacturing	2,698	3.2%
Wholesale Trade	1,705	2.0%
Retail Trade	8,386	10.0%
Transportation and Warehousing	952	1.1%
Information	2,551	3.0%
Finance and Insurance	7,052	8.4%
Real Estate and Rental and Leasing	3,714	4.4%
Professional, Scientific, and Technical Services	14,175	16.8%
Management of Companies and Enterprises	4,171	5.0%
Administration & Support, Waste Management and Remediation	6,300	7.5%
Educational Services	2,889	3.4%
Health Care and Social Assistance	9,269	11.0%
Arts, Entertainment, and Recreation	1,805	2.1%
Accommodation and Food Services	7,087	8.4%
Other Services (excluding Public Administration)	5,772	6.9%
Public Administration	406	0.5%