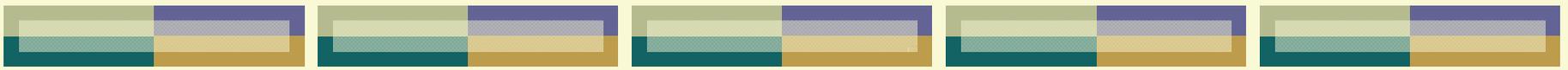


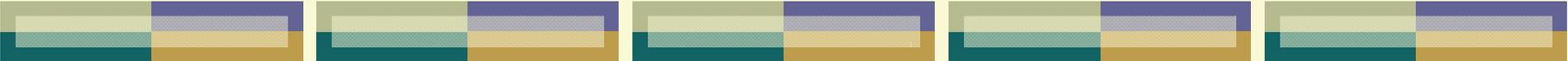
Whither (or wither) Statistical Compendia in the Age of the Web?

Carol A. Hert
Syracuse University
IASSIST, 2004



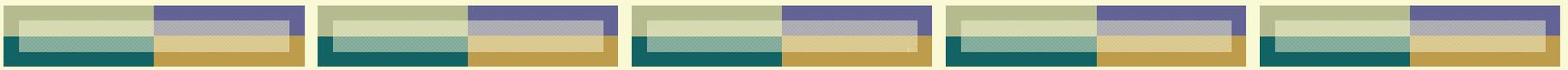
Background

- **Based on results of study funded by BJS in spring 2003**
 - **Impetus**
 - Maturing of web technologies related to information dissemination
 - Mandates for increased access to government information
 - Questions of role of a statistical abstract in this environment
- 



Goals of Project

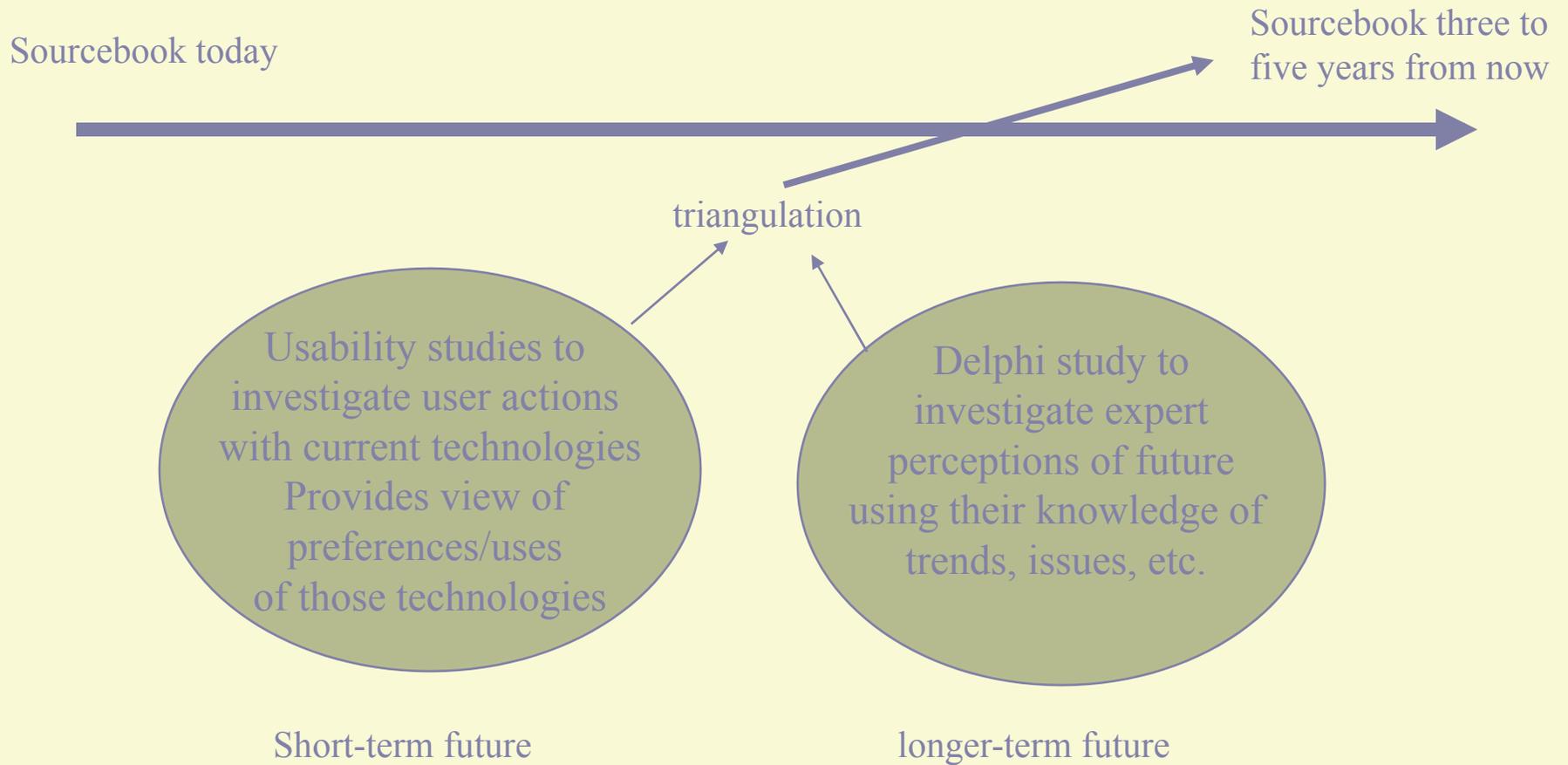
- An integrated look at *Sourcebook of Criminal Justice Statistics*: its users, mission, its future in light of current and emergent technologies
 - *Sourcebook of Criminal Justice Statistics*
 - Published by BJS, produced by University of Albany
 - 30 year history as critical resource in criminal justice statistics
 - Stats from range of sources, over 600 tables
- 

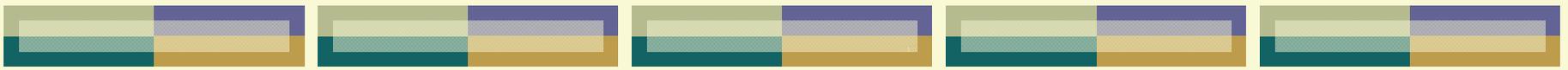


Multi-method Approach

- Delphi study to determine points of consensus on the future mission of *Sourcebook* and on the related requirements for achieving that mission.
 - Document analysis, site visits, focus groups to determine existing knowledge about users
 - Usability studies
- 

Investigating the Future





The Delphi Methodology

- **Used to develop consensus among a group while preserving individuals' anonymity**
 - Structured communication
 - Individual contributions solicited
 - Group judgment/view assessed (by researcher)
 - Opportunities to individuals to revise views (4 rounds in this study)
 - **123 potential participants invited , 43 participated**
 - Experts in criminal justice (practitioners and academics), statistical agency personnel, information scientists, etc.
- 

Round 1 Questions

1. What do you see as the role of statistical almanacs such as *Sourcebook* in the future?

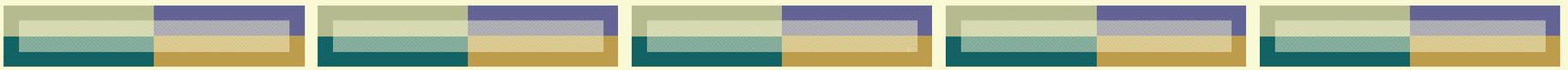
Set of statements generated by researchers, presented in Rounds 2-3 for assessment by participants

Draft mission statement based on agreements

Requirement statements assessed in relationship to draft mission in Round 3

2. What will be needed to accomplish this role?

3. What are the potential opportunities and constraints to achieving this repositioning?



Mission Statement

Sourcebook

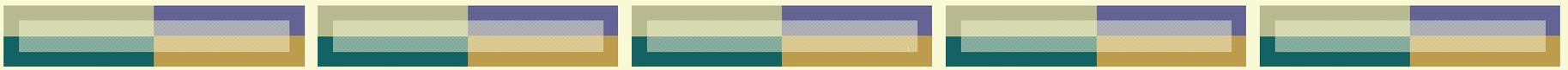
marketed as the definitive source of CJ statistical information at the federal, state, and local level for practitioners, researchers, policy makers, and statisticians.

Sourcebook

promotes/champions the statistics and their producers.

provides data in a variety of ways (i.e., tables, chart, graphics) with associated metadata and explanations, and links to source data.

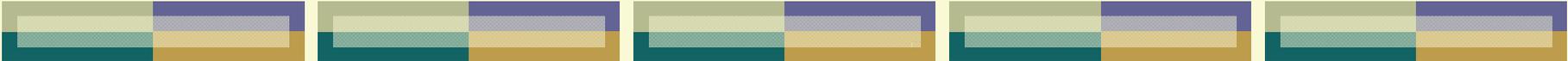




Associated Scenario

Sourcebook:

- Links to criminal justice sites, commentary, and analysis.
 - Provides access to downloadable data sets.
 - Supports historical and other comparative analyses
 - Provides a snapshot of CJS information at a moment in time, as well as remaining current.
 - *Sourcebook* should not be privatized.
- 



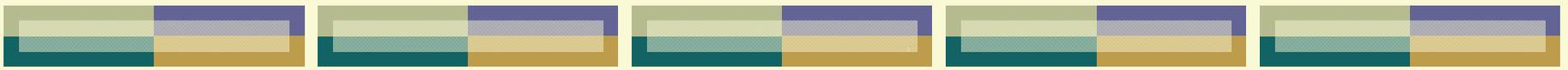
Essential Requirements to Attain Mission

Content Aspects

- Attention to terminology and vocabulary issues
- Indexing
- Understanding the key information requirements in the criminal justice field

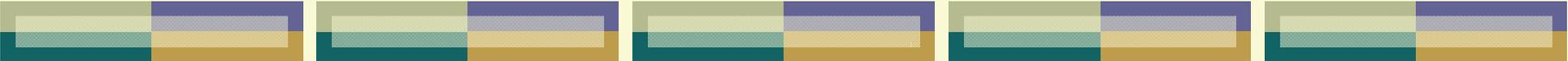
Delivery Aspects

- Focus on usability aspects
 - Internet accessibility
 - Robust search engine
 - Archiving capabilities
- 



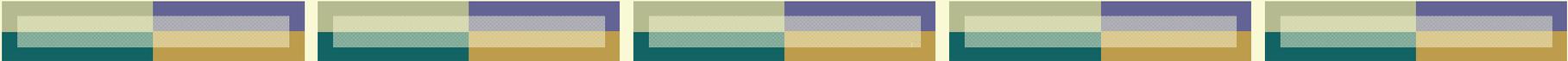
User Studies in the Project

- User studies provide “on-the-ground” view
 - What we didn’t know about users (that would inform future directions)
 - How people make choices among sources of information
 - What tools, presentations, information, etc. help and hinder them as they search for and use criminal justice statistical information
 - In essence, how does *Sourcebook* fit into the larger landscape of criminal justice information?
- 



User Studies

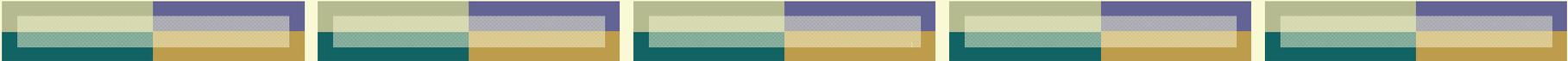
- **Users performed 2-3 tasks using several selected sources for each task**
 - Enabling comparisons across sources
 - **27 users**
 - 16 Washington DC, Olympia WA: expert users from statistical agencies, researchers
 - 11 undergraduate students (Albany, NY)
- 



Preliminary Insights

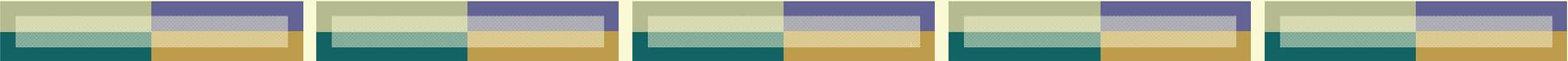
- **Users generally performing matching activity**
 - they look for words on screen (or input into search engine) that match their need-if they don't see them, can't make connection





Preliminary Insights (2)

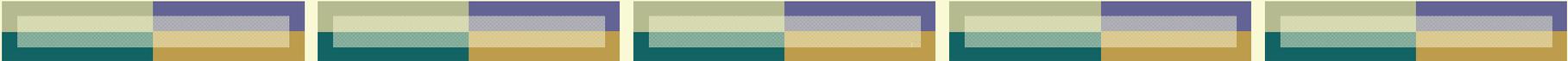
- **Retaining structure from paper version often not helpful**
 - Content structure by process of criminal justice system rarely used
 - Having table in only one place limits retrieval
 - Different uses of .pdf files to retain printability, or “mimic” paper copy sometimes helped, sometimes hindered.
 - Navigational challenges, ease of printing
- 



Preliminary Insights (3)

- **Permeability of boundaries between sources (or parts of website) presented problems.**
 - FBI search searches entire FBI site, not UCR component
 - Sourcebook online search limits to particular components (does the user know which component he/she is searching)

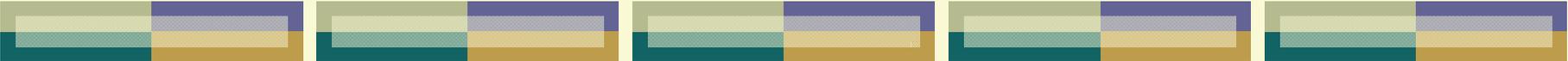




Preliminary Insights (4)

- **Geographic searches challenging because you generally can't search by location name**

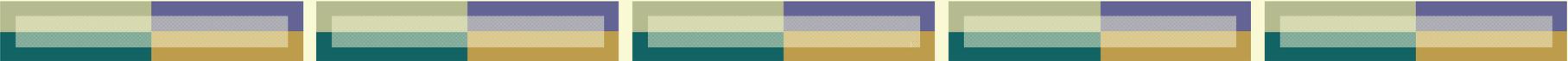




Wither Statistical Compendia?

- **No! Important role for vetted, quality information presented in a coherent package**
 - Provides most commonly used statistics in one spot
 - Provides useful starting point for understanding a domain's statistics
 - Points to other sources with more detailed data





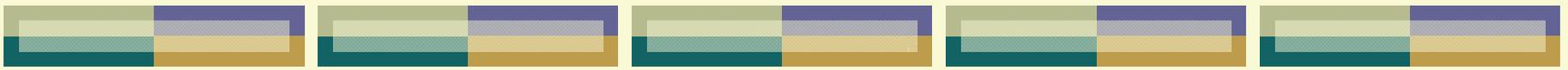
Whither Statistical Compendia?

● Unlink from paper versions

- Provide multiple entry points to a table
- Investigate how people use in an online environment to better understand appropriate organizational/content structures

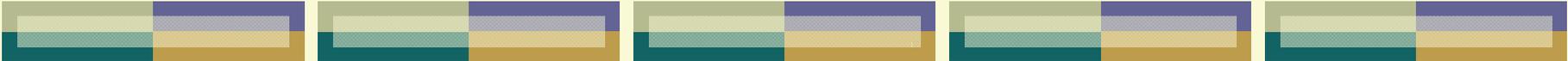
● Enhance Terminological Support





Whither Statistical Compendia?

- **Maintain the coherent package of a Statistical compendium**
 - Search only within the compendium, compendium index
 - **Exploit technological capabilities**
 - Link to sources, technical material linked to tables, perhaps a “Reference shelf” within a compendium
 - Enhanced indexing at table level
- 



Exploit technological capabilities

- Explore technological available for transforming statistical presentations or presenting multiple presentations “manually”
 - Assess technologies for enabling users direct access to data to perform unique manipulations
- 