

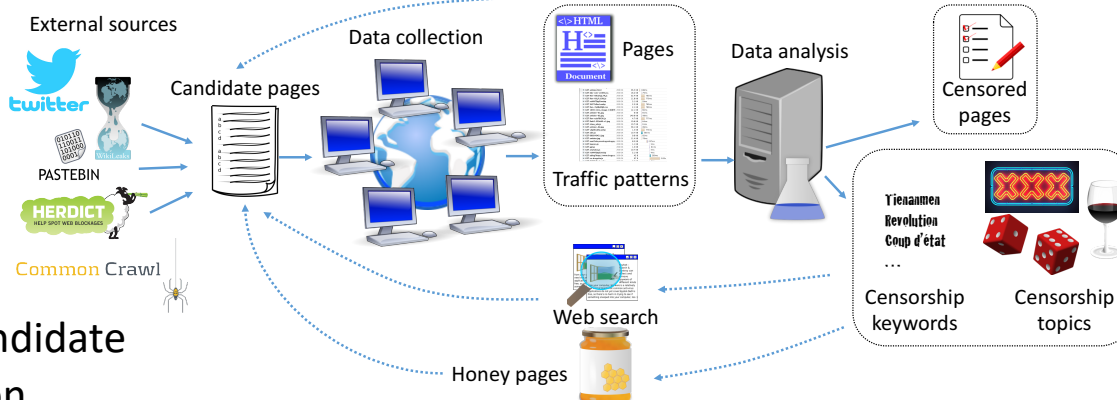
# SaTC: CORE: Small:

## Toward Fully Automated Data-Driven Analysis of Web Censorship



### Challenge:

- Is it possible to **automatically** infer
  - which techniques censors use?
  - which topics censors are after?
- ... with minimal human involvement?



### Scientific Impact:

- Obtain comprehensive picture of the censorship techniques used in practice
- Automatic detection of censorship topic changes
- Understand limits of automation in the censorship measurement context

### Solution:

- Automated candidate page generation
- Automated collection and analysis
  - Distinguishing between overt and covert censorship
  - Used to feed back into candidate pages

### Broader Impact:

- Important to journalists, activists, ...
- Transition to the CS/networking measurement community
- Interface with political science, sociology, geopolitics...
- PhD and undergraduate student training

NSF Award [CNS-1814817](#)  
Carnegie Mellon University  
Nicolas Christin <[nicolasc@cmu.edu](mailto:nicolasc@cmu.edu)>