Reflection: Accountability and Web Scraping in Airbnb Regulation

Kelsie Backus

SI 206: Project 2

A. Is the current system adequate for accountability?

   a. No, the current system is inadequate in providing full accountability. According to Dutch scholar Mark Bovens, accountability requires that individuals with responsibilities transparently explain their actions to a forum that can evaluate those actions and impose consequences. In this case, while Airbnb requires listers to enter a policy number, it does not verify the authenticity of the number before allowing listings to go live. As shown by the listing ID "16204265," it's possible to input a random string or incorrect policy number and still operate an illegal short-term rental. This represents a gap in the accountability process. Airbnb, as a platform, should act as an intermediary that enforces compliance. Yet, they don't consistently validate policy numbers against city records. As a result, listers can bypass regulations, and third-party accountability (citizens or researchers like us) becomes essential but is not sustainable or systematic.

   b. *Recommendation:* Airbnb developers should implement a validation system that checks submitted policy numbers against a live database provided by the San Francisco Planning Office. Listings should remain "pending" until policy numbers are verified. This change would align Airbnb's responsibilities with Bovens' model by ensuring transparency, justification, and consequences.

B. Additional Stakeholder: San Francisco Residents

   a. An additional stakeholder is San Francisco residents — especially tenants and neighborhood associations. These individuals are directly affected by short-term rentals, which can reduce long-term housing availability and increase local housing costs. Using Bovens' definition of accountability, residents can act as watchdogs by reporting unregistered or suspicious listings. They can also advocate for transparency and collaborate with local journalists or civic tech groups to document issues.

   b. *Recommendation:* The SFPO should create a user-friendly reporting platform and offer responses or status updates to complaints. This would empower residents to participate in the accountability process and make it more robust.

C. Legal Consideration in Web Scraping

   a. A key legal factor in web scraping is whether the data is protected by login or Terms of Service. Under U.S. law, courts have ruled inconsistently. In *hiQ Labs v. LinkedIn*, scraping public-facing data was ruled legal. However, the Computer Fraud and Abuse Act (CFAA) can still be used against scrapers if they bypass technological protections. This makes it critical to ensure scraping targets public, unauthenticated content and is done without breaching a platform's terms. Scraping Airbnb must only be done from pages that are publicly accessible, and not for commercial or exploitative purposes.

D. Ethical Guidelines for Web Scraping

    a. *Guideline 1:* Avoid collecting personally identifiable information (PII) and only scrape public, unprotected data. *Why:* Even public data can be used in harmful ways, as shown by the Cambridge Analytica scandal. Collecting names, faces, or profiles without consent can lead to misuse. *Accountability Enforcement:* Universities, research labs, and institutions should implement ethics reviews before allowing scraping projects. Developers should publicly disclose the scope and use of data.

    b. *Guideline 2:* Use scraped data only for purposes that serve the public interest. *Why:* Examples like Clearview AI show how scraping can fuel surveillance and discrimination. Even if legal, this violates social and ethical norms. *Accountability Enforcement:* Public transparency is key. Platforms like GitHub or academic conferences should require disclosure of data sources, consent, and ethical justifications when publishing scraped datasets or code.