

## Segment 1: Pain Relievers and Gain Creators

**Key pains** related to process and results in data collection, summarized from 4Q analysis:

- Small size of hand-collected data sets
- Data sets too narrow and industry specific
- Data sets miss important variables
- Data collection with the help of RAs or purchasing data too expensive
- Data collection takes too long

→ **Pain relievers** from involving crowds?

- Large crowd can collect larger volume of data points
- Large crowd can collect data from multiple industries/ product categories
- Crowd can code additional variables if given good instructions
- Large crowd can collect data faster (parallel work)
- Data collection likely less expensive than hiring RAs or buying from vendors

**Key gains** related to process and results in data collection, summarized from 4Q analysis:

- Data from multiple countries
- Standardized format used by other researchers
- Transparent documentation
- Continuous updating

→ **Gain creators** from involving crowds?

- Distributed crowd can collect data from multiple countries
- Can use standard formats – to be considered when designing tools and interfaces for data collection
- Can use transparent documentation – needed anyways to ensure consistent data collection by crowd members
- Project could run for longer time, crowd members could revisit products periodically
- Outreach to professors to help recruit crowd members may also build reputation as patent scholar using creative methods
- Funding agencies may like approach to involve general public and to educate them about patents and patent policies

**Feasibility check:** Can the design really address the pains/gains you identified?  
**Opportunity check:** Can crowd involvement address pains/gains not considered before?

## Segment 2: Strategic Design Choices

### Crowd Science Paradigm Diamond (Why involve a crowd?)

- Crowd volume: **Very relevant**
- Broadcast search: **Less relevant**
- User crowd: **Less relevant**
- Community production: **Less relevant**
- Crowd wisdom: **Somewhat relevant**



### AKRD Crowd Contribution Matrix (What does the crowd contribute?)

Activities	Collect patent data from products by searching online or products found at home; enter data and submit photographs or links; verify others' submissions; submit complete product-patent lists found on company websites
Knowledge	General knowledge to identify and enter patent numbers
Resources	Computer, internet connection; camera/smartphone; products to analyze
Decisions	Generate options and decide which products or companies to collect data from

### Six Crowd Characteristics (Who is the crowd?)

- **Location:** Different countries → restrict to English-speaking countries to use standard materials
- **Knowledge and skills:** General → English language; ideally some understanding of patent system
- **Time commitment:** Flexible (from a few minutes to many hours)
- **Resources:** Computers and internet, smartphones
- **Size:** As large as possible (at least a few thousands)
- **Diversity:** Location, socio-economic status and consumption habits (gives more diversity in products to collect data from)

**Feasibility check:** Is the design realistic? What adjustments need to be made?

## Segment 3: Implementation Challenges and Solutions

### Key challenges and solutions specific to this particular stage of the project:

- Resource requirements: **Minimal** – should not be a problem
- Invasiveness of data collection: **Should be no problem**
- Expectations of openness: I prefer to keep full data closed, at least initially – explain to crowd why; publish visualizations of interesting subsets of data
- Data protection: **No concerns**

### Organizational challenges and solutions that cut across all stages (see chapters 13–14):

- Dividing and allocating tasks: Tasks are naturally divided (by product); crowd members choose products to work on
- Coordinating crowd members: Will post information on products already submitted; challenge contributors to look for uncommon products (perhaps reward with extra points?)
- Training and enabling learning: Video tutorial; practice examples with automated feedback
- Increasing quality and evaluating contributions: Design submission interface with guiding questions; algorithm to flag unusual patent numbers
- Motivating crowd members: Gamification, e.g., give points for each product submitted; create forum for sharing unusual products (should be fun but also increase coverage)
- Recruiting crowd members: Post project on Scistarter.org; promote projects among colleagues as “hands-on” experience for students in economics and innovation studies

### Research integrity and ethical issues that cut across all stages (see chapter 15):

- Ensuring quality and preventing misconduct: Mandatory video to explain common mistakes and explain why quality is important
- Recognizing effort and sharing project outputs: Acknowledge crowd in publications; feature top performers on website; do not make data open; any financial proceeds re-invested in project
- Role of AI: Automation, augmentation, management: AI to verify submissions; perhaps in future: AI to help crowd members and algorithmic management
- Privacy, safety, institutional oversight: Few issues; instruct participants to exclude personal information when taking pictures