The Scroll of Rough Thinking

+ 3 to critical thinking

+ 10% chance to succeeding in the quest

-10 to artistic skills

-7 Aesthetic

Proxy manager

Worker/Thread

Worker/Thread

Worker/Thread

Worker/Thread

Task Queue

Domain Delay Registry

Domain Backoff Map (redis)

Could have a thing that determines if these are links to be searched or looking for metadata (or could be done by worker analysing the url

List Page Cache   
(For discovery stage)?  
Maybe discovery and extraction should be separate stages

Basically flow should be for the workers

1. Worker pulls task from queue
2. Before making a request it checks
   1. Domain delay (is domain ready for this proxy) if not requeue the task
   2. Domain backoff (is the domain currently blocked)
3. If ok, request proceeds, if 429, update backoff requeue or log and
4. Select a proxy that can hit the domain
5. Make the request
   1. If success, depending on type of results, cache result and/or store results or scrape metadata
   2. If 429 or fail, update backoff, possibly log it, discard or requeue END
6. Store metadata
7. Update domain delay registry
8. Mark task complete