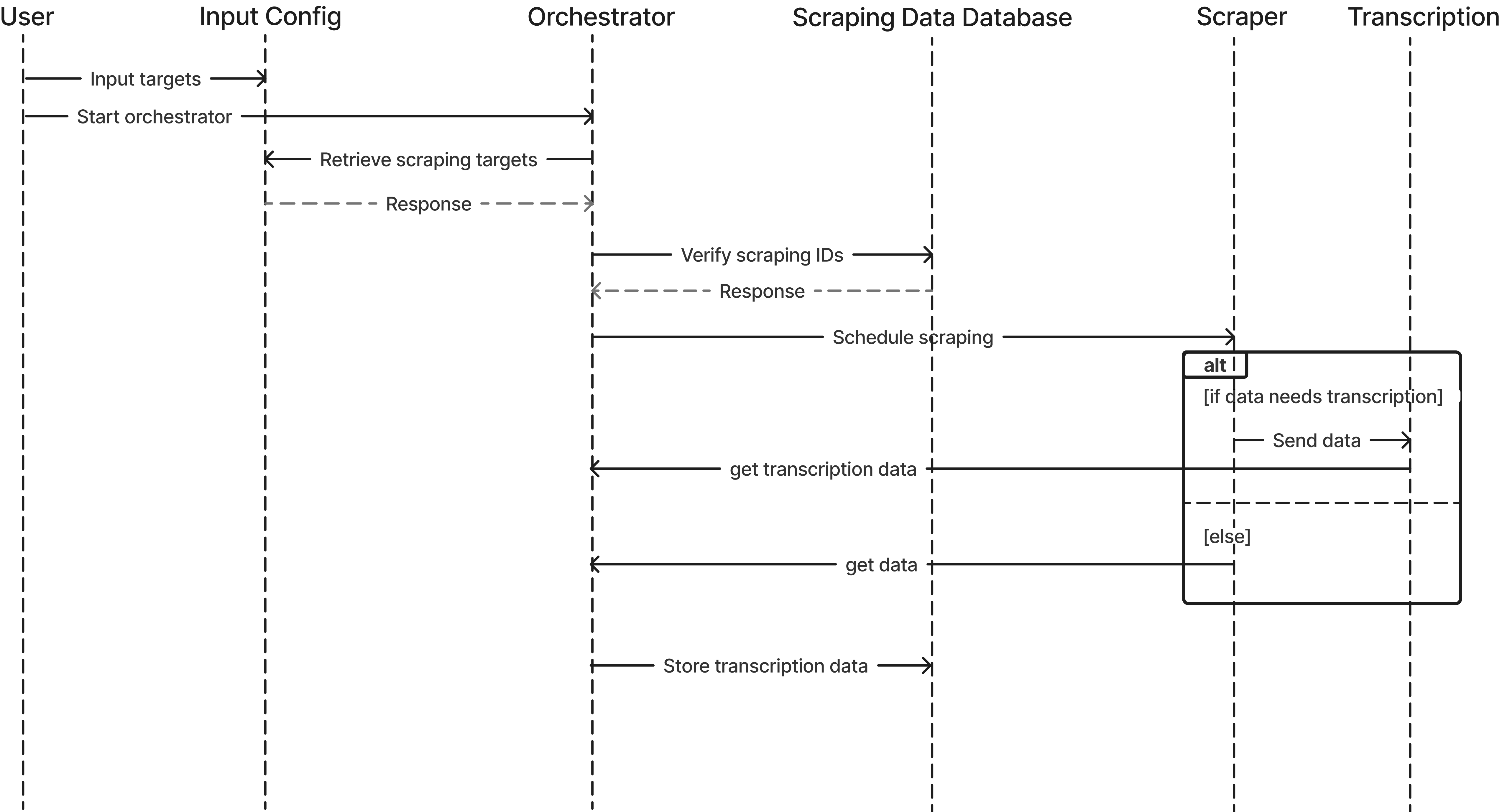


## Scenario 1 - Scraping Pipeline



- 1.) download\_element(DataScrapingElement) : DataScrapingElement**  
gets a DataScrapingElement which contains a key or url which is necessary to scrape eg. a video, the scraper scrapes the resource, attaches the text data to the data field and metadata to the metadata field of the DataScrapingElement and returns it. Orchestrator will store the Element in the DB
- 2.) get\_links(channel\_url): video\_links[]**  
Example: for podcast scraping we get a website link that contains podcasts and keywords. The method should return a list of all the individual podcast links. The orchestrator then checks whether these have already been scraped. Podcasts that have not been scraped will be scheduled for scraping with the download\_element(...) method