

SEP	15
Title	ScrapyManager and SpiderManager API refactoring
Author	Insophia Team
Created	2010-03-10
Status	Final

## SEP-015: ScrapyManager and SpiderManager API refactoring

This SEP proposes a refactoring of ScrapyManager and SpiderManager APIs.

### SpiderManager

- `get(spider_name)` -> Spider instance
- `find_by_request(request)` -> list of spider names
- `list()` -> list of spider names
- `remove fromdomain(), fromurl()`

### ScrapyManager

- `crawl_request(request, spider=None)`
  - calls `SpiderManager.find_by_request(request)` if spider is None
  - fails if `len(spiders returned) != 1`
- `crawl_spider(spider)`
  - calls `spider.start_requests()`
- `crawl_spider_name(spider_name)`
  - calls `SpiderManager.get(spider_name)`
  - calls `spider.start_requests()`
- `crawl_url(url)`
  - calls `spider.make_requests_from_url()`
- `remove crawl(), runonce()`

Instead of using `runonce()`, commands (such as `crawl/parse`) would call `crawl_*` and then `start()`.

### Changes to Commands

- `if is_url(arg):`
  - calls `ScrapyManager.crawl_url(arg)`
- `else:`
  - calls `ScrapyManager.crawl_spider_name(arg)`

### Pending issues

- should we rename `ScrapyManager.crawl_*` to `schedule_*` or `add_*` ?
- `SpiderManager.find_by_request` or `SpiderManager.search(request=request)` ?