

Processing Management Architecture

Sequence of module execution detected dynamically and depend on two field:

1. Processor's module (e.g "Sentiment Analyzer", "Scraper", etc)
2. Processor's Input resource

Input data format

To support processing management's functionality need to two additional site's properties in dc_sites.sites_properties table

1. PROCESSING_MASK - (used by db-task module for operations such as "URL_FETCH", etc.)
2. PROCESSING_DATA - (used by ProcessorTask module as input data for algorithm processing)

PROCESSING_MASK

Each bit of that field points to the single processor's module (e.g. "Scraper", "Sentiment Analyzer", etc.) and can be changed **"on the fly"**.x

len 64 bit

Bit matching set via ini file (../ini/processor-task.ini):

```
[procesistng_mask]
scraper=10000
sentiment=01000
```

PROCESSING_DATA

As JSON with structure:

```
{"processor_modules": [<Processor_module_1>, [<Processor_module_2>, ...,
[<Processor_module_N>]]}
```

Processor_module:

```
{<Processor_module_name>:<Processor_module_data>}
```

Processor_module_data:

```
{"input_resource":<Input_resource_name>}
```

Input_resource_name - (e.g. "raw_content", "description", etc.)