

SEP	5
Title	ItemBuilder API
Author	Ismael Carnales, Pablo Hoffman
Created	2009-07-24
Status	<p>Obsoleted by :doc:`sep-008`</p> <div style="border: 1px solid black; padding: 10px; margin: 10px 0;"> <p>System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\sep\[scrapy-master] [sep] sep-005.rst, line 7); <i>backlink</i></p> <p>Unknown interpreted text role "doc".</p> </div>

SEP-005: Detailed ItemBuilder API use

Item class for examples:

```
#!/python
class NewsItem(Item):
    url = fields.TextField()
    headline = fields.TextField()
    content = fields.TextField()
    published = fields.DateField()
```

gSetting expanders

```
#!/python
class NewsItemBuilder(ItemBuilder):
    item_class = NewsItem

    headline = reducers.Reducer(extract, remove_tags(), unquote(), strip)
```

This approach will override the Reducer class for BuilderFields depending on their ItemField class:

- MultivaluedField = PassValue
- TextField = JoinStrings
- other = TakeFirst

gSetting reducers

```
#!/python
class NewsItemBuilder(ItemBuilder):
    item_class = NewsItem

    headline = reducers.TakeFirst(extract, remove_tags(), unquote(), strip)
    published = reducers.Reducer(extract, remove_tags(), unquote(), strip)
```

As with the previous example this would select join_strings as the reducer for content

gSetting expanders/reducers new way

```
#!/python
class NewsItemBuilder(ItemBuilder):
    item_class = NewsItem

    headline = BuilderField(extract, remove_tags(), unquote(), strip)
    content = BuilderField(extract, remove_tags(), unquote(), strip)

    class Reducer:
        headline = TakeFirst
```

gExtending ItemBuilder

```
#!/python
class SiteNewsItemBuilder(NewsItemBuilder):
    published = reducers.Reducer(extract, remove_tags(), unquote(),
                                strip, to_date('%d.%m.%Y'))
```

gExtending ItemBuilder using static methods

```
#!/python
class SiteNewsItemBuilder(NewsItemBuilder):
    published = reducers.Reducer(NewsItemBuilder.published, to_date('%d.%m.%Y'))
```

gUsing default_builder

```
#!/python
class DefaultedNewsItemBuilder(ItemBuilder):
    item_class = NewsItem

    default_builder = reducers.Reducer(extract, remove_tags(), unquote(), strip)
```

This will use default_builder as the builder for every field in the item class. As a reducer is not set reducers will be set based on Item Field classes.

gReset default_builder for a field

```
#!/python
class DefaultedNewsItemBuilder(ItemBuilder):
    item_class = NewsItem

    default_builder = reducers.Reducer(extract, remove_tags(), unquote(), strip)
    url = BuilderField()
```

gExtending default ItemBuilder

```
#!/python
class SiteNewsItemBuilder(NewsItemBuilder):
    published = reducers.Reducer(extract, remove_tags(), unquote(), strip, to_date('%d.%m.%Y'))
```

gExtending default ItemBuilder using static methods

```
#!/python
class SiteNewsItemBuilder(NewsItemBuilder):
    published = reducers.Reducer(NewsItemBuilder.default_builder, to_date('%d.%m.%Y'))
```