

Glossary

Definitions of the most important concepts used on the dexi.io platform



Written by Morten Franck
Updated over a week ago

- *Account*: is created when you sign up. Multiple users with different roles can be added to your account.
- *Addon*: additional functionality that can be added to your account, e.g. more step types/pipe actions or integrations.
- *AutoBot*: a type of robot that, given a list of input URLs (on different domains), maps inputs to outputs via an Extractor robot per domain.
- *Configuration*: various properties of a robot, e.g. its concurrency level, inputs, proxy or schedule. Multiple configurations can be defined for a robot. Synonymous with *Run*.
- *Crawler*: a type of robot that visits links given a starting page and extracts basic information about the pages it visits, e.g. the URL and page title.
- *Data set*: a set or table of rows. Similar to a table in a SQL database, a sheet in a spreadsheet or collection in a NoSQL database. See also *Deduplication* and *Record Linkage*.
- *Data type*: field definition of a row used in a data set, dictionary - or as input or output for an Extractor, Pipes or AutoBot robot. Can be useful to standardise input and output when working with many robots.
- *Deduplication*: the process of removing duplicate rows within one data set as defined by the key configuration on the data set.
- *Dictionary*: a mapping of keys to values.
- *Element path*: a CSS3 selector expression used in steps in Extractors specifying how to traverse the DOM tree of a web page to reach the particular element to interact with.
- *Execution*: the process of running a specific configuration of a robot. An execution has one result per input.
- *Extractor*: a type of robot that extracts information from a web page and interacts with the page in various ways, e.g. fills in forms, clicks buttons and much more.
- *Input*: the values to be used when executing a robot, e.g. a URL or a search query value. A data type can be used. In an Extractor the input fields are defined and in a configuration of the robot the actual values are specified. In Pipes robots input fields are automatically calculated given the starting nodes of the Pipes graph.
- *Integration*: a type of addon that specifically integrates (or connects) with an external/3rd-party service, e.g. Amazon S3, Facebook or Google Drive.
- *Key configuration*: field definition of how duplicates in a data set should be identified.
- *Output*: the definition of fields in a robot that should be saved as results. A data type can be used. In an Extractor the output fields are defined. In Pipes robots output fields are automatically calculated given the exit nodes of the Pipes graph.

- *Pipes*: a type of robot that performs various actions in a sequence or workflow, e.g. reads data from a source, performs some processing/transformation and saves results in a data store.
- *Pipes action*: a part of a Pipes robot that performs some action, e.g. executes a robot, iterates rows in a data set, makes HTTP requests or looks up details of a Facebook page by name.
- *Project*: an entity on your account, e.g. a robot or data set.
- *Proxy*: a server performing requests on behalf of a robot execution.
- *Record Linkage*: the process of combining two data sets using the key configuration on the data set combining into. For more details, see [How do I use Data Sets for deduplication/record linkage?](#)
- *Results*: the data, in row format, saved by a robot. See also *Execution* and *Output*.
- *Result log*: a text file containing all events pertaining to the particular result.
- *Robot*: a type of entity that performs an automated process, e.g. extracts information from a web page. See *Extractor*, *Pipes*, *Crawler* and *AutoBot*.
- *Run*: synonymous with *Configuration*.
- *Schedule*: the recurrence with which a configuration is executed. Can be expressed in cron syntax.
- *Scraper*: deprecated. See *Extractor*.
- *Step*: a part of an Extractor robot that performs some action, e.g. visits a URL, clicks a link, waits for an element or extracts a piece of information.
- *Timetable*: the earliest and latest possible times that a configuration can be executed.
- *Trigger*: is added to a project entity, e.g. a configuration or data set, which causes an action to be performed when some event occurs, e.g. adding a row to a data set when an execution completes.
- *Webhook*: a type of integration that notifies an external endpoint about some event, e.g. when an execution completes.
- *Worker*: a part of the [dexi.io](#) platform that does the work of a robot, e.g. extracts information from a web page. The number of workers on your account determines how much execution work can be done concurrently. Visit our [price plans](#) to see how many workers are included in each price plan.

Did this answer your question?



 We run on Intercom