

Présentation des blocs permettant l'écriture des scénarios du SMART PATH

Préliminaire :

- L'écriture des scénarios s'effectue au format YAML.
- Une liste YAML correspond à un scénario de site.
- Une liste YAML est rédigée pour chaque site dans un fichier commun "*scenarii.yaml*"
- Il est possible de tester les scénarios sur le site : "<http://www.yamllint.com/>"
- Un exemple de fichier "*scenarii.yaml*" se trouve à la fin de ce document.

Liste des blocs

1. GoPage()
2. ClickPage()
3. SaveJob()
4. Scroll()
5. GoBack()
6. FindDate()

Gestion des nodes

1. La première ligne du scénario renvoie au nom du site traité.
2. La seconde ligne du scénario renvoie au premier node du scénario, indexé "o"
3. Typologie des nodes :
 - *actualNode* = node de la ligne traitée
 - *nextNode* = repère pour envoyer à un autre node à la fin du bloc, par défaut "nextNode" renvoie au node de la ligne suivante
 - *possibleNode* = repère pour envoyer à un node si une condition interne au bloc est validée
4. En cas de problème lors de l'exécution du bloc, le "possibleNode" référencé sera le node utilisé par défaut pour poursuivre le scénario.
5. En cas de problème lors de l'exécution de la dernière ligne du scénario et en l'absence d'un "possibleNode", le scénario s'arrête.

Avertissement:

Attention aux éventuelles boucles infinies avec la gestion des nodes.

Gestion des tags

1. Un tag est défini selon deux paramètres.
2. "*class*", "*name*", "*target*" sont les types de balises recherchées
3. Les types de balises sont liées à une chaîne de caractère
4. En cas de recherche dans des balises en cascade, les tags sont imbriqués dans un dictionnaire
5. En cas de recherche dans une balise d'un niveau équivalent, on utilisera « index : n » pour renvoyer à la balise *n*-suivante

Description des blocs :

Bloc GoPage() :

Le bloc **GoPage()** permet d'accéder à la page web des offres. Il nécessite en entrée un lien internet qui

renvoie à la page d'offre d'emplois de l'entreprise visée.

Exemple :

```
GoPage: {'url': "https://www.safran-group.com/fr/emplois?pays=France"}
```

GoPage : référence du bloc utilisé

'url' : variable d'entrée du bloc

'https://www.safran-group.com/fr/emplois?pays=France' : adresse url renvoyant à la page web

Bloc ClickPage()

Le bloc **ClickPage()** permet de cliquer sur un lien url spécifique : fonction recherche, accéder à l'offre d'emploi, accéder à la page suivante du site. Il nécessite en entrée la liste des tags nécessaires à la navigation dans la page HTML.

Exemple :

```
ClickPage: {'tag' : {'class' : 'type type-28 viewport', 'tag' : {'name' : 'a'}}, 'possibleNode' : 5}
```

ClickPage : référence du bloc utilisé

'tag' : marqueur lié à la recherche HTML

'class' : type du tag recherché

'type type-28 viewport' : valeur de la variable 'class'

'tag' : marqueur lié à la recherche HTML

'name' : type du tag recherché

'a' : valeur de la variable 'name'

Note : Ici on remarque l'utilisation en cascade des tags permettant de trouver l'adresse html nécessaire à la poursuite du scénario et l'appel au node 5 en cas d'échec dans la recherche des tags.

Bloc SaveJob()

Le bloc **SaveJob()** permet de sauvegarder l'offre d'emploi. Il ne nécessite pas de variable d'entrée. Le programme python est chargé d'effectuer la sauvegarde locale puis le transfert sur la base de données.

Exemple :

```
SaveJob: {}
```

SaveJob : référence du bloc utilisé

Bloc Scroll()

Le bloc **Scroll()** permet de simuler l'action de la souris afin de charger les données dynamiques du site. Il nécessite en entrée un entier int relatif à la distance nécessaire pour afficher les nouvelles informations.

Exemple :

```
Scroll : {'size' : 10, 'possibleNode' : 5}
```

Scroll : référence du bloc utilisé

'size : 10' : paramétrage de la taille du scroll nécessaire.

'possibleNode : 5' : en cas d'absence de chargement de donnée, on se rend au node 5, si aucune donnée n'est trouvée.

Bloc GoBack()

Le bloc **GoBack()** permet d'effectuer un retour en arrière pour retourner sur la page url précédente. Il nécessite en entrée la valeur du "nextNode".

Exemple :

```
GoBack: {'nextNode' : 2}

GoBack : référence du bloc utilisé
'nextNode : 2' : paramétrage du node appelé à la suite
```

Bloc FindDate()

Le bloc **FindDate()** permet de repérer la date présente dans la page. En interne, il déterminera si l'offre d'emploi est intéressante ou non. Il nécessite en entrée la liste des tags nécessaires à la navigation dans la page HTML.

Exemple :

```
FindDate: {'tag' : {'class' : 'ts-offer-card-content offerContent', 'tag' : {'name' : 'li'

FindDate : référence du bloc utilisé
'tag' : marqueur lié à la recherche HTML
'class' : type du tag recherché
'ts-offer-card-content offerContent' : valeur de la variable 'class'
'tag' : marqueur lié à la recherche HTML
'name' : type du tag recherché
'li' : valeur de la variable 'name'
'index : 1' : marqueur et valeur lié à la recherche dans la balise soeur n°1
```

Note : Exemple ici de l'utilisation de l'index pour la recherche d'une balise au même niveau que la précédente.

Récapitulatif des blocs généralisés :

- **NomSite:**
 - **GoPage:** {'url': "url_site"} # Je vais sur La page emploi
 - **Scroll:** {'size' : int, 'possibleNode' : int} # Je scroll pour charger la page, en cc
 - **FindDate:** {'tag' : {'class' : str, 'tag' : {'name' : str, 'index' : int }}} # Je che
 - **ClickPage:** {'tag' : {'class' : str, 'tag' : {'target' : str}}, 'possibleNode' : int}
 - **SaveJob:** # Je sauvegarde l offre
 - **GoBack:** {'nextNode' : int} #Je reviens en arrière et lance le bloc du node référence
 - **ClickPage:** {'tag' : {'class' : str}, 'nextNode' : int, 'possibleNode' : int} # Je cl

Exemples de scénarios :

SAFRAN

- **SAFRAN:**
 - **GoPage:** {'url': "https://www.safran-group.com/fr/emplois?pays=France"}
 - **FindDate:** {'tag' : {'class' : 'date'}, 'possibleNode' : 5}
 - **ClickPage:** {'tag' : {'class' : 'offer-card'}}
 - **SaveJob:**
 - **GoBack:** {'nextNode' : 1}
 - **ClickPage:** {'tag' : {'class' : 'next', 'tag' : {'name' : 'a'}}, 'nextNode' : 1}

BNP

- **BNP:**
 - **GoPage:** {'url': "https://group.bnpparibas/emploi-carriere/toutes-offres-emploi/france"
 - **Scroll:** {'size' : 10, 'possibleNode' : 5}
 - **ClickPage:** {'tag' : {'class' : 'type type-28 viewport', 'tag' : {'name' : 'a'}}, 'pos
 - **SaveJob:**
 - **GoBack:** {'nextNode' : 1}
 - **ClickPage:** {'tag' : {'class' : 'progress-buton elastic show-more'}, 'nextNode' : 1, 'p
 - **ClickPage:** {'tag' : {'class' : 'next', 'tag' : {'name' : 'a'}, 'nextNode' : 1}}

SODEXO

- **SODEXO:**
 - **GoPage:** {'url': "https://fr.sodexo.com/home/nous-rejoindre/rejoignez-nos-equipes.html"
 - **FindDate:** {'tag' : {'class' : 'ts-offer-card-content offerContent', 'tag' : {'name' : 'a'}}
 - **ClickPage:** {'tag' : {'class' : 'ts-offer-card_title'}}}
 - **SaveJob:**
 - **GoBack:** {'nextNode' : 1}
 - **ClickPage:** {'tag' : {'class' : 'ts-ol-pagination-list-item_link ts-ol-pagination-list

TOTAL

- **TOTAL:**
 - **GoPage:** {'url' : 'https://krb-sjobs.brassring.com/tgnewui/search/home/home?partnerid=3'
 - **ClickPage:** {'tag' : {'class' : 'primaryButton ladda-button ng-binding'}}}
 - **FindDate:** {'tag' : {'class' : 'jobProperty position1'}}}
 - **ClickPage:** {'tag' : {'class' : 'jobProperty jobtitle'}}}
 - **SaveJob:**
 - **GoBack:** {'nextNode' : 2}
 - **ClickPage:** {'tag' : {'class' : 'showMoreJobs UnderLineLink ng-binding'}, 'nextNode' : 2}

CANAL

- **CANAL:**
 - **GoPage:** {'url' : 'https://www.vousmeritezcanalplus.com/metiers.html'}
 - **FindDate:** {'tag' : {'class' : 'srJobListPublishedSince'}}}
 - **ClickPage:** {'tag' : {'class' : 'srJobListJobOdd'}}}
 - **SaveJob:**
 - **GoBack:** {'nextNode' : 1}

DASSAULT

- **DASSAULT:**
 - **GoPage :** {'url' : 'https://careers.3ds.com/fr/jobs?woc=%7B%22pays%22%3A%5B%22pays%2Ffr'
 - **ClickPage :** {'tag' : {'class' : 'ds-card ds-card--lines ds-card--image ', 'tag' : {'tag
 - **SaveJob:**
 - **GoBack:** {'nextNode' : 1}
 - **ClickPage:** {'tag' : {'class' : 'ds-pagination_next', 'tag' : {'name' : 'a'}, 'nextNoc