

Nom propre - correction

Correction de l'Exercice 1 – Affiner une requête

Étape 2 : Rechercher les formes anglaises et françaises *nom propre*, *proper name* et *proper noun*.

```
"nom propre" OR "proper name"
OR "proper noun"
```

31 213¹ résultats. Cette syntaxe utilise l'opérateur booléen OR (« ou »). La requête est insensible à la casse. Une analyse des résultats révèle du bruit et du silence.

Étape 3 : Limiter le bruit et le silence.

```
"nom propre" OR "noms propres" OR
"proper name" OR "proper names" OR
"proper noun" OR "proper nouns"
```

65 113 résultats. Pour limiter le silence, on peut chercher les variantes morphosyntaxiques (singulier/pluriel).

```
title:("nom propre" OR "noms
propres" OR "proper name" OR
"proper names" OR "proper noun" OR
"proper nouns") OR abstract:("nom
propre" OR "noms propres" OR
"proper name" OR "proper names" OR
"proper noun" OR "proper nouns") OR
subject.value:("nom propre" OR
"noms propres" OR "proper name" OR
"proper names" OR "proper noun" OR
"proper nouns")
```

1 425 résultats. Les parenthèses permettent de faire des groupes de priorité. La spécification des champs fournit un corpus plus pertinent.

```
title:("nom propre" OR "noms
propres" OR "proper name" OR
"proper names" OR "proper noun" OR
"proper nouns") OR abstract:("nom
propre" OR "noms propres" OR
"proper name" OR "proper names" OR
"proper noun" OR "proper nouns") OR
```

1 148 résultats. Les *dictionnaires de noms propres*, *index*, *errata*, *nomenclature* et *back matter* ne sont pas des documents pertinents pour l'objectif visé.

¹ Chiffres en date du 22 mars 2024.

```
subject.value:("nom propre" OR
"noms propres" OR "proper name" OR
"proper names" OR "proper noun" OR
"proper nouns")NOT
title:(dictionnaire OR dictionary
OR index OR errata OR "back
matter")
```

```
(title:("nom propre" OR "noms
propres" OR "proper name" OR
"proper names" OR "proper noun" OR
"proper nouns") OR abstract:("nom
propre" OR "noms propres" OR
"proper name" OR "proper names" OR
"proper noun" OR "proper nouns") OR
subject.value:("nom propre" OR
"noms propres" OR "proper name" OR
"proper names" OR "proper noun" OR
"proper nouns"))AND publicationDate
:[1913 TO 2019] NOT
title:(dictionnaire OR dictionary
OR index OR errata OR "back
matter")
```

```
(title:("nom propre" OR "noms
propres" OR "proper name" OR
"proper names" OR "proper noun" OR
"proper nouns") OR abstract:("nom
propre" OR "noms propres" OR
"proper name" OR "proper names" OR
"proper noun" OR "proper nouns")
OR subject.value:("nom propre" OR
"noms propres" OR "proper name" OR
"proper names" OR "proper noun" OR
"proper nouns"))AND
publicationDate :[1913 TO 2019]
AND
qualityIndicators.tdmReady:true
NOT title:(dictionnaire OR
dictionary OR index OR errata OR
"back matter")
```

```
(title:("nom propre" OR "noms
propres" OR "proper name" OR
```

930 résultats. Grâce aux informations complémentaires, on peut encore spécifier la requête en précisant les dates de publications pertinentes. Le champ `publicationDate` peut s'interroger par intervalle.

602 résultats. L'adjonction de `qualityIndicators.tdmReady:true` permet de vérifier que les textes seront exploitables.

573 résultats. Le champ `abstract` doit contenir des caractères.

```
"proper names" OR "proper noun" OR
"proper nouns") OR abstract:("nom
propre" OR "noms propres" OR
"proper name" OR "proper names" OR
"proper noun" OR "proper nouns")
OR subject.value:("nom propre" OR
"noms propres" OR "proper name" OR
"proper names" OR "proper noun" OR
"proper nouns")) AND
publicationDate:[1913 TO 2019]
AND
qualityIndicators.tdmReady:true
NOT title:(dictionnaire OR
dictionary OR index OR errata OR
"back matter") AND abstract:*
```

Pour aller plus loin...

Pour limiter le silence, on cherche les variantes morphosyntaxiques. Pour cela on peut utiliser des expressions régulières présentées entre barres obliques / :

- les . signifient n'importe quel caractère ;
- les * signifient n'importe quel nombre fois (s'appliquent au caractère précédent) ;
- les ? signifient entre 0 et 1 fois (s'appliquent au caractère précédent) ;
- les crochets [] sont l'équivalent d'un OR ;
- les .raw considèrent que le token interrogé est le champ dans sa globalité.

Pour en savoir plus et tester des expressions régulières se rendre sur <https://regex101.com/>.

```
(title.raw:(/.*[pP]roper.[nN]ames?.* / .*[pP]roper.[nN]ouns?.* /
.*[Nn]oms?.[pP]ropres?.* /) OR
abstract.raw:(/.*[pP]roper..?[nN]ames?.* /
.*[pP]roper.[nN]ouns?.* / .*[Nn]oms?.[pP]ropres?.* /) OR
subject.value.raw:(/.*[pP]roper.[nN]ames?.* /
.*[pP]roper.[nN]ouns?.* / .*[Nn]oms?.[pP]ropres?.* /)) AND
publicationDate:[1913 TO 2019] AND
qualityIndicators.tdmReady:true NOT title:(dictionnaire OR
dictionary OR index OR errata OR "back matter")AND abstract:*
```