

# Guide d'annotation en concepts wikipedia

V 17 janvier 2019

## Objectif

identifier les concepts/entités/événements de la Wikipedia présents dans chaque diapositive du diaporama du cours fourni et lister (on dit aussi annoter) exhaustivement ceux appartenant au domaine du cours

## Procédure de recherche dans la wikipedia

1. Rechercher librement le concept (via l'expression littérale copier-coller de la diapositive) dans la wikipedia-fr, si vous le trouvez alors retenez celui-ci même si il n'est pas retourné comme choix numéro un par le moteur de wikipedia.
2. Si vous ne trouvez pas l'objet dans la wikipedia fr, regardez dans la en, préciser la valeur 'en' dans la colonne other-language
3. Si vous ne trouvez pas l'objet ni dans la wikipedia -fr ni dans la -en et que vous pensez que le concept est pertinent alors noter seulement l'ancre

## Procédure d'annotation

Dans un tableur, prendre une ligne pour chaque concept identifié dans une diapositive donnée. Puis sur des colonnes distinctes

1. Slide: Indiquer le numéro de la diapositive (1 pour la 1ère)
2. Anchor: Indiquer l'ancre c'est-à-dire la zone de texte de la diapositive qui porte le concept/entité/événement ; Faire un copier-coller le plus fidèle (cela peut couvrir sur plusieurs mots/lignes mais supprimer les retour chariot dans votre copier-coller)
3. Wikipedia\_concept : Indiquer (copier-coller) le titre de la page tel que l'on peut le lire dans la page wikipedia retenue
4. Url : donner l'url de la page
5. Slide-focus : indiquer False si le concept courant n'est pas le concept principal de la diapositive (si vous le retirez de la diapositive sans que cela ne touche à la compréhension du concept principal alors ce n'est pas un concept en focus)

6. Image-based: indiquer True si le concept est inférée de votre analyse d'une image (diagramme, logo ou autre)
7. Where
  - a. indiquer "en" si vous n'avez pas trouvé le concept dans la wikipedia-fr mais dans la wikipedia-en (moyennant éventuellement une traduction manuelle).
  - b. Indiquer "out-of-wikipedia" si le concept que vous avez identifié n'est ni dans la wikipedia 'fr' ni dans celle 'en'.
8. Out-of-domain : Les concepts qui ne sont pas relatifs au domaine ne doivent pas être listés ; si jamais vous le faisiez alors mettre True ici
9. Use : now (pour satisfaire un besoin de clarification immédiate)/later (pour aller plus loin, consommation différée) ; par défaut, indifférent

## Consignes diverses :

- Une même ancre peut donner plusieurs concepts et sera donc sur plusieurs lignes avec des parties dupliquées
- Le même numéro de diapositive peut se trouver sur plusieurs lignes si la diapositive contient plusieurs concepts distincts
- Si le concept n'est pas du domaine (tel que l'attribution d'un auteur d'une photo), ne pas le lister
- Si un concept se produit plusieurs fois dans une même diapositive, ne l'indiquer qu'une fois. Par contre s'il se produit dans plusieurs diapositives distinctes et même si il n'est pas en focus, le lister.
- S'il n'y a pas de concepts évidents pour une diapositive, le numéro de la diapositive ne devra pas apparaître.
- Si le concept émane d'une figure (logo, diagramme, photo, ...) et non d'un texte, le noter aussi. La colonne anchor sera laissée vide mais la colonne from-image sera à True.