

Cet article présente une nouvelle ressource lexicographique ancienne mise à disposition sur Internet : le *Petit Larousse Illustré* de 1905. Faisant suite à des œuvres de plus grande ampleur et de plus grande renommée (le dictionnaire critique de Féraud, le dictionnaire de Nicot, les différentes éditions de celui de l'Académie, etc.), le *Petit Larousse Illustré* de 1905, bien plus modeste que ses prédécesseurs - en volume tout au moins - a été numérisé et sera mis en ligne prochainement.

L'intérêt de la mise en ligne d'une telle ressource réside dans sa nature. Il s'agit d'un petit dictionnaire illustré, et la présence d'images est importante. Par ailleurs, il est le premier d'une série de dictionnaires grand public, ce qui le rend fondamental dans l'histoire de la lexicographie.

L'informatisation s'est déroulée en plusieurs phases, de façon à permettre une interrogation fine du dictionnaire. Les différents éléments des articles du dictionnaires ont été décrits et listés, puis balisés en XML selon les standards décrits dans la proposition 5 de la TEI. Le texte a ensuite été balisé automatiquement grâce à des programmes écrits en langage Python contenant des expressions régulières. Le balisage s'est déroulé en trois passes, chacune exploitant le résultat de la précédente.

Le résultat de l'informatisation est une base de données lexicales riche qui permet à l'utilisateur deux sortes de consultations : il peut choisir de faire une interrogation plein texte. Dans ce cas, le résultat apparaîtra avec les images associées aux articles répondant à sa requête. L'utilisateur peut aussi faire une recherche avancée, c'est à dire n'interroger qu'un seul champ de l'article du dictionnaire (vedette, prononciation, information grammaticale, étymologie, définitions, définitions encyclopédiques, renvois, proverbes, exemples, expressions figées). Seules les requêtes sur la vedette permettent l'affichage des images.