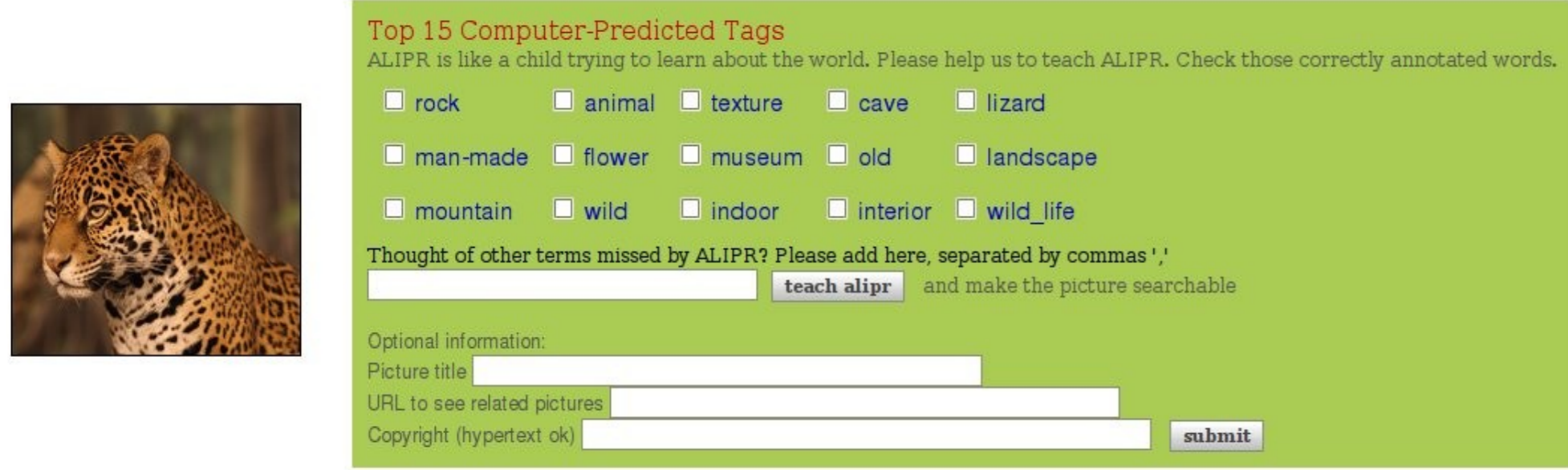


Nicolas James, Céline Hudelot
 {nicolas.james,celine.hudelot}@ecp.fr
 Laboratoire MAS, Ecole Centrale Paris

Contexte:

Le problème du fossé sémantique (Smeulders, 2000).



Objectifs:

- Développer une technique d'annotation d'images utilisant à la fois contenu visuel et connaissance a priori.

➔ Vers une annotation sémantique structurée.

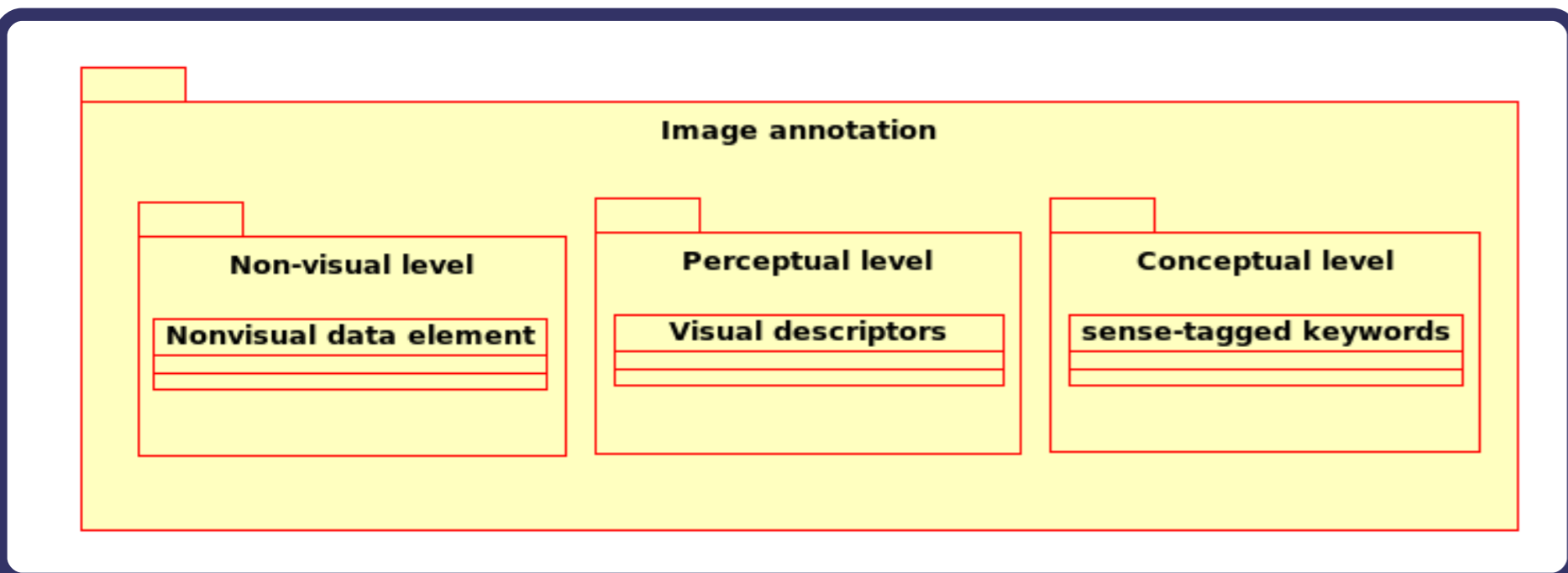


Figure 1 : Modèle d'annotation structurée en 3 niveaux (le niveau non visuel représentant les métadonnées associées au document).

➔ Permettre aux utilisateurs des moteurs de recherche d'images d'écrire des requêtes complexes.
 Par exemple: voiture de la marque jaguar

Notre approche:

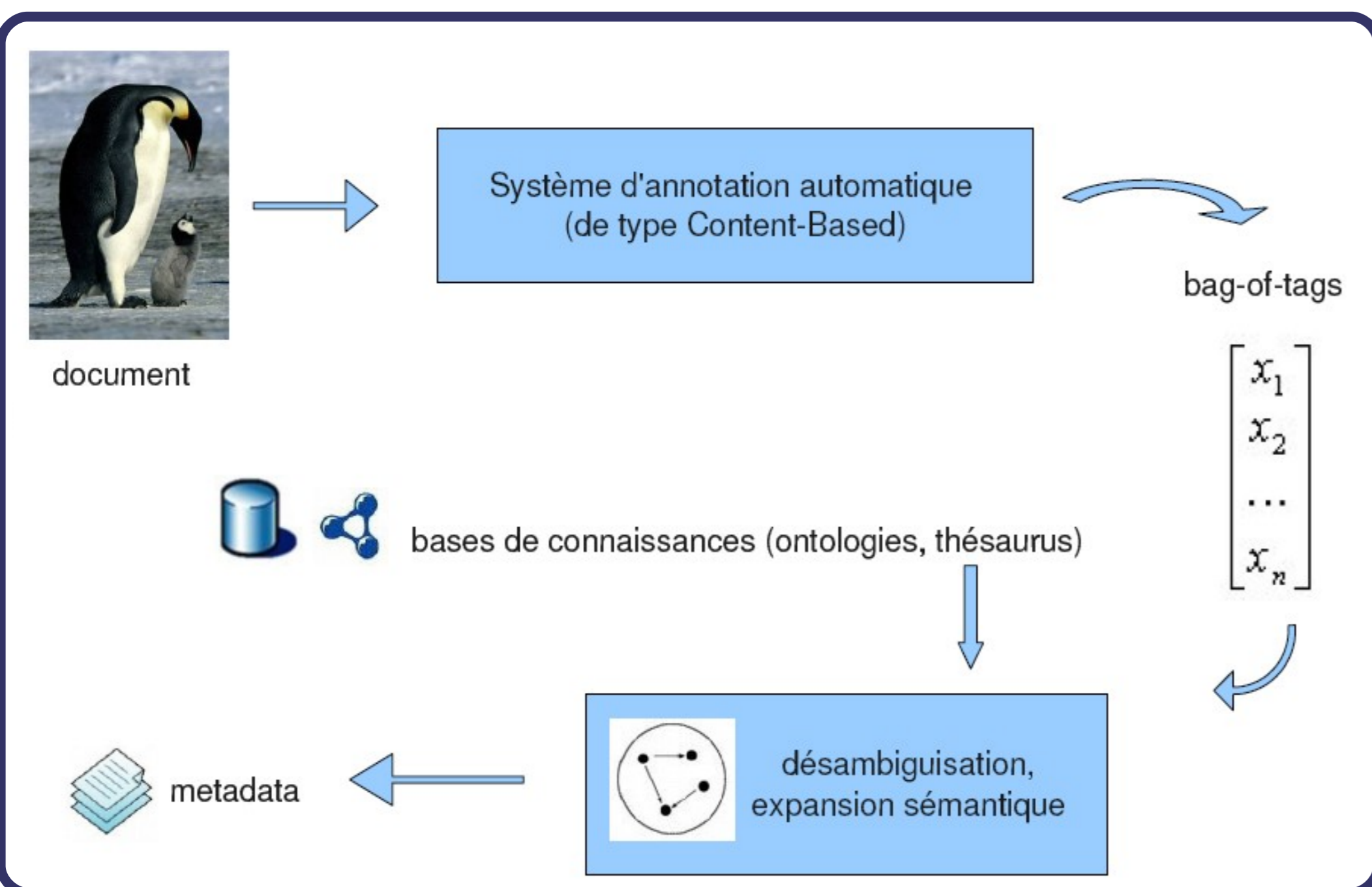


Figure 2 : Composants d'un système d'annotation sémantique automatique d'images

- Basée sur une technique de CBIA, *Content-Based Image Annotation*.
- S'appuie sur des marches aléatoires avec redémarrages sur un graphe représentant une base d'images d'apprentissage.

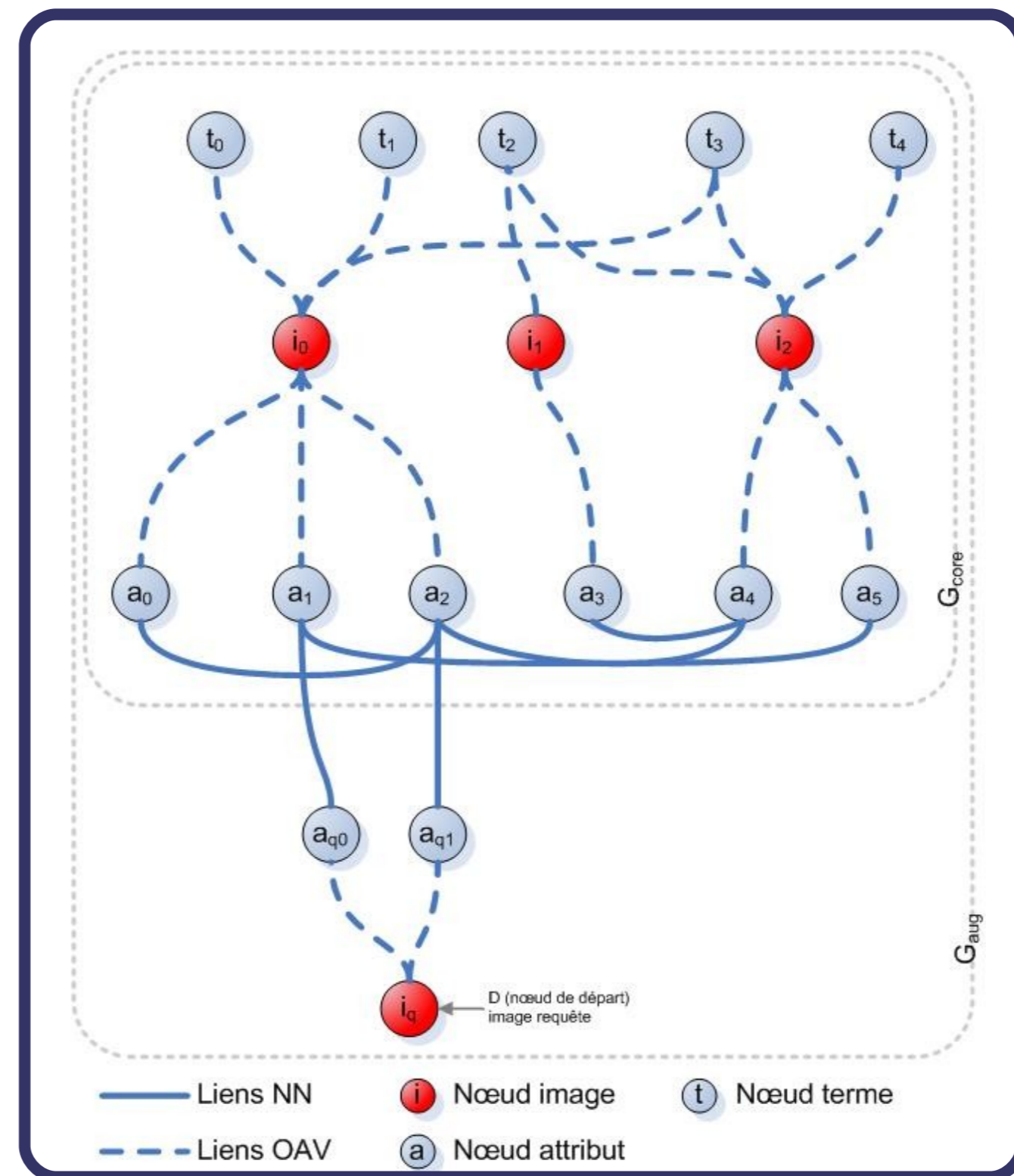


Figure 2 : Modélisation de la base d'apprentissage sous la forme d'un ARG, *Attributed Relational Graph*.

➔ En sortie du système: un ensemble de mot-clés indépendants et dont le sens est inconnu.

➔ Recherche du sens des mot-clés à l'aide de connaissances a priori (*WordNet*).

➔ Nécessite un processus de désambiguïisation.

Mot-clés	Nombre de sens dans WordNet
Fern	1
Street	5
Pole	10
Sign	11
House	12
Tree trunks	1
Bench	7
People	4
Tree	3

Source: <http://www.cs.washington.edu/research/imagedatabase>

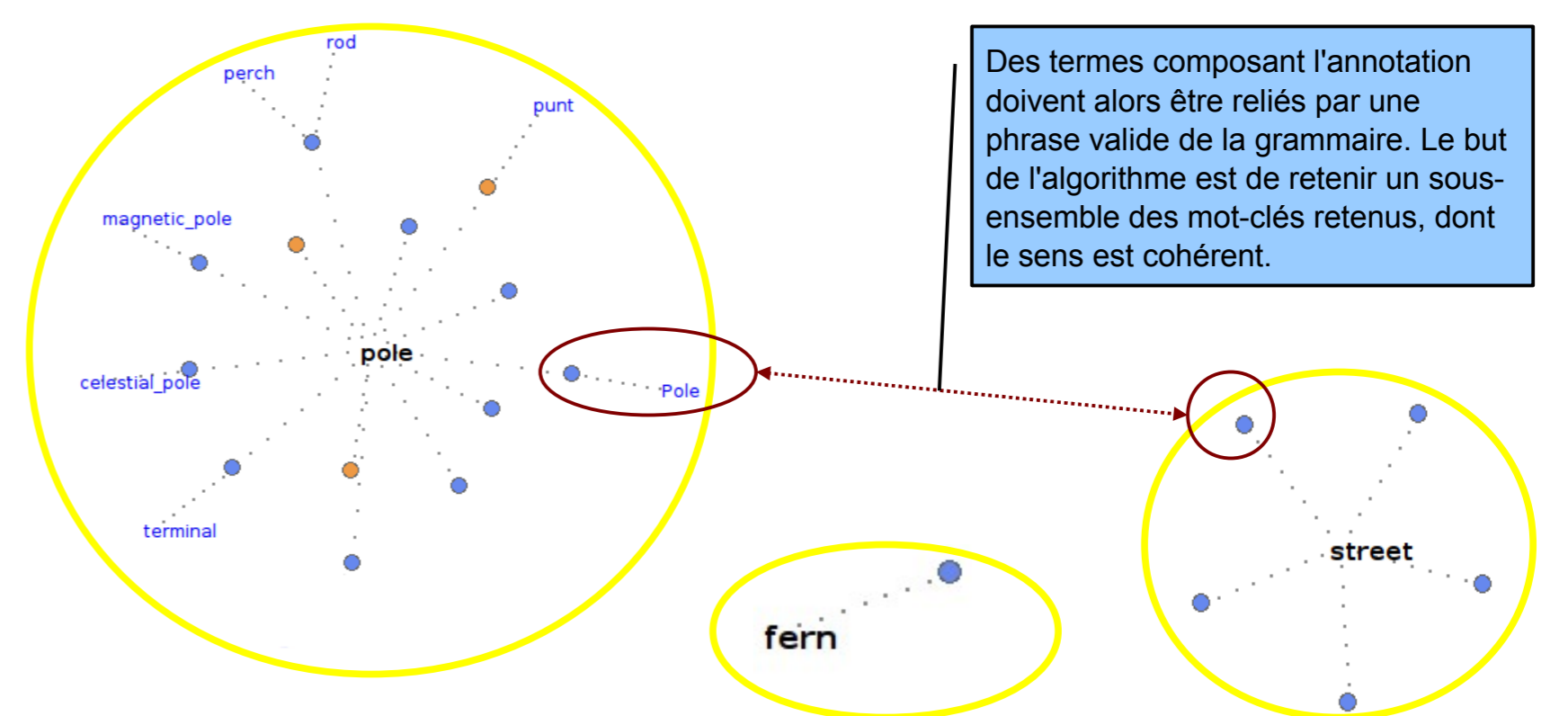
Figure 3 : exemple d'image soumise au système et les 9 mot-clés calculés comme étant les plus pertinents pour annoter l'image.

Algorithme SSIIA

Algorithme de désambiguïisation des mot-clés, le **SSIIA**, Structural Semantic Interconnection for Image Annotation, adaptation du SSI (*Navigli, 2005*) à l'image.

- Permet l'utilisation de multiple relations sémantiques.
- Combinaisons de relations sémantiques pondérées et paramétrées par une grammaire.

S_G	$\rightarrow S_s S_g$	(all the heuristics)
S_s	$\rightarrow S_1 S_2 S_3$	(simple heuristics)
S_1	$\rightarrow E_1 S_1 E_1$	(hypernymy/meronymy)
E_1	$\rightarrow e_{\text{kind-of}} e_{\text{part-of}}$	
S_2	$\rightarrow E_2 S_2 E_2$	(hyponymy/holonymy)
E_2	$\rightarrow e_{\text{has-kind}} e_{\text{has-part}}$	
S_3	$\rightarrow e_{\text{has-kind}} S_3 e_{\text{has-kind}} e_{\text{has-kind}} e_{\text{has-kind-of}}$	(parallelism)
S_g	$\rightarrow e_{\text{gloss}} S_g S_4 S_5$	(gloss)
S_4	$\rightarrow e_{\text{gloss}} e_{\text{topic}}$	(gloss, context)



Références:

- Smeulders, A.W.M. and al., Content-Based Image Retrieval at the End of the Early Years, 2000.
- Pan, J.Y. and Yang, H.J. and Faloutsos, C. and Duygulu, P., Automatic multimedia cross-modal correlation discovery, Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining, 2004.
- Jin, Y. and Khan, L. and Wang, L. and Awad, M., Image annotations by combining multiple evidence & WordNet, Proceedings of the 13th annual ACM international conference on Multimedia, 2005.
- Navigli, R., Velardi, P., Structural Semantic Interconnections: A Knowledge-Based Approach to Word Sense Disambiguation, IEEE Trans. Pattern Analysis and Machine Intelligence, 2005.
- Bartolini, I., Ciaccia, P., Imagination: Accurate Image Annotation Using Link-analysis Techniques, 5th International Workshop on Adaptive Multimedia Retrieval, 2007.