**Critères qualité**

Contrôle qualité par l’équipe de l’OBVIL, en liaison avec les équipes de recherche.

Évaluation au niveau « mot » (nombre de mots erronés) + contrôle du la conformité du balisage (format XML) ou du stylage (format traitement de texte).

● 1er contrôle : évaluation du taux de mots absents des dictionnaires (hors entités nommées).

● 2e contrôle : alignement entre le texte livré par le prestataire et le texte définitif (vérité terrain) établi par les équipes (méthode Text Eval ou ocrevalUAtion).

Lot 1 (saisie/OCR avancé, livrable format XML) :

● Qualité requise : 99,9 %.

● Compte non tenu des mots en doute ou des mots/passages non traitables (balisage en utilisant l’élément <seg).

● Taux de non-faisabilité des textes acceptable : 5 %.

Lot 2 (saisie, livrable format traitement de texte, stylage complexe) :

● Qualité requise : 99 %.

● Compte non tenu des mots en doute ou des mots/passages non traitables (balisage en utilisant l’élément <seg>).

● Taux de non-faisabilité des textes acceptable : 10 %.

Lot 3 (OCR standard, livrable format traitement de texte, stylage standard niveau blocs) :

● Qualité requise : 99,5 %.

● Compte non tenu des mots/passages non traitables (balisage en utilisant l’élément <seg>).

● Taux de non-faisabilité des textes acceptable : 5 %.