

Introduction aux humanités numériques – L1HN001

Nom et prénom d'étudiant-e :

Numéro d'étudiant-e :

Date d'examen : 19 octobre 2023, 08h-08h30

Question 1

(1 point)

Parmi les options ci-dessous, laquelle **n'est pas** un format de fichier d'image ?

- ☐ JPG
- ☐ PNG
- ☒ DOCX
- ☐ IIIF

Question 2

(1 point)

Qui a inventé le terme *distant reading* ? Franco Moretti

Question 3

(2 points)

Que désigne le terme « métadonnées » ? Donnez quelques exemples de ce terme.

On appelle « métadonnées » des données structurées décrivant une ressource ou une autre donnée. Une métadonnée peut être un titre, un auteur, une date, un sujet, un éditeur, etc.

Question 4

(1 point)

Parmi les options ci-dessous, choisissez celle(s) qui correspond(ent) à (aux) bibliothèque(s) numérique(s) :

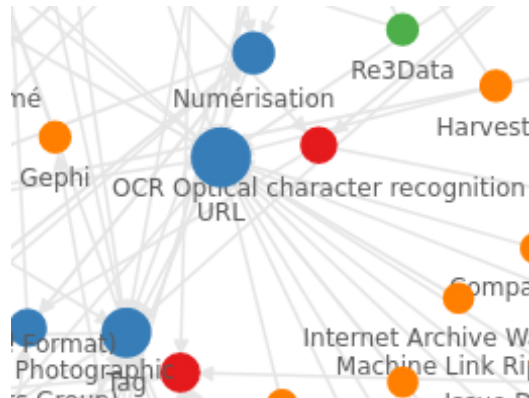
- ☒ Gallica
- ☐ Transkribus
- ☐ Pandore
- ☒ Wikisource

Question 5

(1 point)

L'illustration suivante représente (une seule réponse possible) :

- ☐ une distance
- ☒ un graphe
- ☐ une chaîne de traitement
- ☐ aucune des réponses ci-dessus



Question 6

(6 points)

Reliez les éléments de gauche avec les phrases de droite qui vont ensemble.

- concordance** ———→ occurrences d'un mot dans son contexte
- tokenisation** ———→ découpage d'un texte en unités plus petites
- OCR** —————→ méthode de reconnaissance optique de caractères
- crowdsourcing** ———→ production participative d'une grande quantité de données
- interopérabilité** → possibilité de communication entre deux/plusieurs systèmes
- moissonage** ———→ extraction des données de sites web

Question 7

(2 point)

Dans le segment de texte suivant : `<author>Jean-Jacques Rousseau</author>`, les éléments `<author>` et `</author>` correspondent aux :

- ☒ balises
- ☐ expressions régulières
- ☐ attributs
- ☐ définitions

Ces éléments font partie du langage XML

Question 8

(1 point)

Dans le traitement automatique des langues, les mots qui sont fréquents et n'ajoutent que peu ou pas de valeur sémantique s'appellent aussi les mots vides / outils / fonctionnels / grammaticaux.

Question 9

(1 point)

Quel est l'un des intérêts d'utilisation de l'ordinateur pour faire de la recherche en humanités numériques ?

L'ordinateur se veut un outil indispensable pour automatiser des tâches fastidieuses (p. ex. compter le nombre des mots dans un document) ou même pour faire émerger des tendances et des faits nouveaux, qu'il serait quasiment impossible de découvrir de manière manuelle.

Question 10

(2 points)

Répondez par VRAI (V) ou FAUX (F) aux phrases suivantes :

V F

- (a) Une des particularités des humanités numériques est leur aspect monodisciplinaire. ☐ ☒
- (b) L'un des objectifs de l'OCR (ou de l'HTR) est d'accélérer la transcription du texte. ☒ ☐

Question 11

(2 points)

Dans la phrase to be, or not to be, indiquez le nombre de :

1. *tokens* : 6

2. *types* : 4

(sans compter les signes de ponctuation)