

Colloque Frontières Numériques

5^e ed., Hammamet, 07/06/2023

À l'interface entre l'art et les technologies :
proposition d'un système Typing-to-Scene

Lionel LALOUM & Alain LIORET

laloum.lionel@gmail.com

alainlioret@gmail.com



edesta

Ecole Doctorale Esthétique, Sciences et Technologies des Arts

UNIVERSITÉ
PARIS 8

Ecole Doctorale
d'Esthétique, Sciences et
Technologie des Arts

Une idée créative intéressante

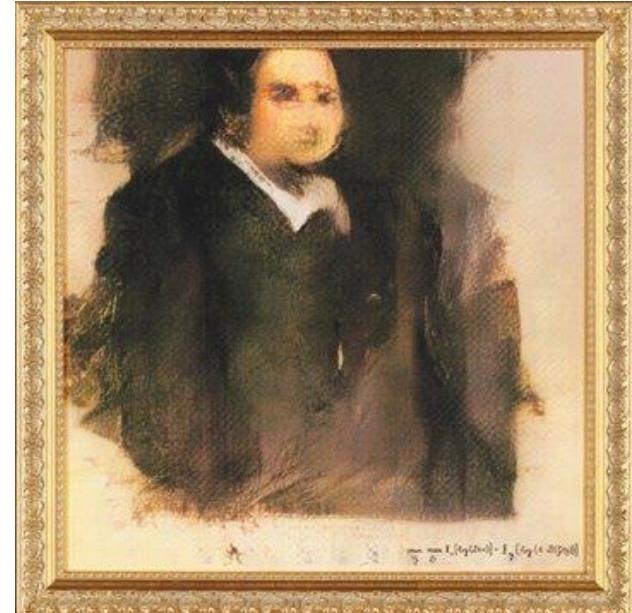


Art génératif



Pseudomnesia: The Electrician,
Boris Eldagsen, 2023
Sony World Photography Awards

Edmond de Belamy
Collectif Obvious, 2018



"Théâtre D'opéra Spatial", Jason Allen, 2022
Colorado State Fair's digital arts competition Award

Modélisation Déclarative

▶ Interface graphique vs textuelle :

Selon (Lee & Yan, 2014), dans le cadre de la conception d'animation, que l'utilisation de l'interface en langage naturel puisse se révéler plus avantageuse sur certains aspects (comme la conception et la manipulation globale des éléments, le temps d'apprentissage et de réalisation).

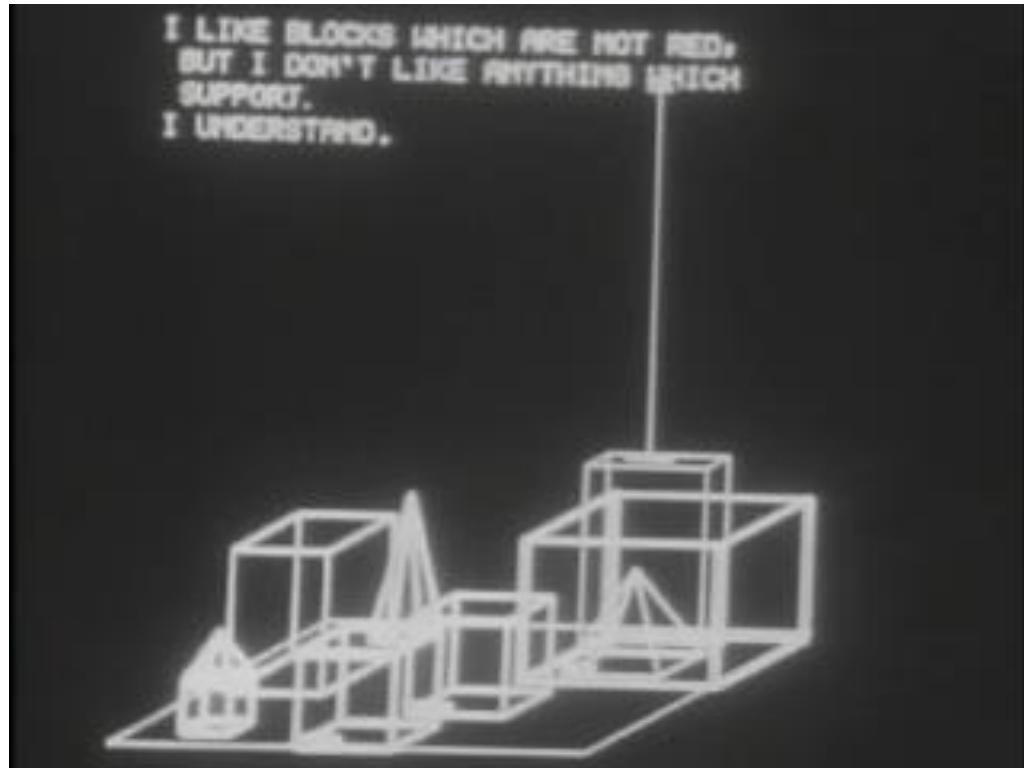
(Lee, S. and Yan, J. 2014. Potential of a Text-Based Interface as a Design Medium: An Experiment in a Computer Animation Environment. *Interacting with Computers*, iwu036. (2014).

▶ Définition proposée par V. Gaildrat :

« Le but de la modélisation déclarative est d'offrir à un concepteur la possibilité de décrire une entité ou une scène à modéliser, non plus en fournissant des données numériques, mais en exprimant l'image mentale qu'il a de la scène grâce à un ensemble de propriétés énoncés dans un langage quasi-naturel [...]»

(in *Modélisation déclarative d'environnements virtuels : Création de scènes et de formes complexes par l'énoncé de propriétés et l'emploi d'interaction gestuelle*, 2013)

Le Langage Naturel comme outil



SHRLDU, Tenny Winograd, MIT, ~1970

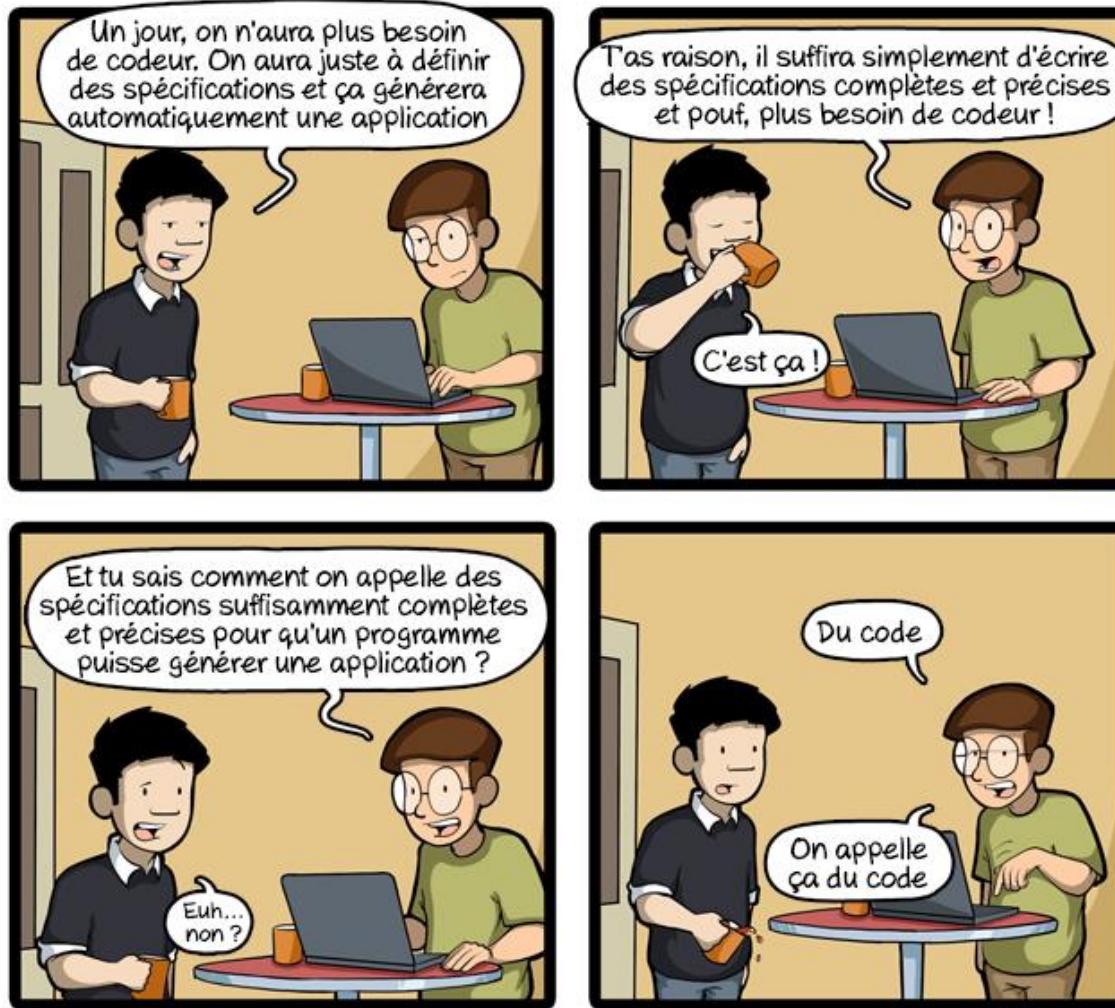


Siri, iPhone (Apple)

A screenshot of a ChatGPT interface. On the left, a sidebar lists various AI models and projects: New chat, Besoins éducatifs particuliers, Extraire des chaînes de caractères, Etudes d'IA et Arts Numériques, Etapes de conception, Préparation cours séminaire, Mini, Systèmes Text-to-Scene, Générer des scènes 3D, Travaux de Gademer, and Correction d'article AI. The main area shows a message from the user: "Peux-tu m'expliquer simplement ce que Kant appelle 'la chose en soi' et me dire dans quels livres il en parle ?". A response from ChatGPT follows: "Kant a développé sa philosophie en opposition au dogmatisme métaphysique, qui affirmait la connaissance de la chose en soi, c'est-à-dire la réalité telle qu'elle est en dehors de toute expérience possible. Selon Kant, la chose en soi ne peut pas être connue car elle est au-delà des limites de notre expérience. Cependant, Kant utilise le concept de "chose en soi" pour désigner ce qui est à l'origine de nos sensations et de notre expérience, mais qui est inaccessible à notre connaissance directe. Pour Kant, la chose en soi est donc une réalité transcendantale qui nous échappe mais qui est à la base de notre expérience. Kant développe sa conception de la chose en soi dans plusieurs de ses ouvrages, notamment dans la "Critique de la raison pure" (1781), la "Critique de la raison pratique" (1785) et la "Critique de la faculté de juger" (1790). Dans ces œuvres, Kant montre que la connaissance de la chose en soi est impossible, mais qu'elle est néanmoins nécessaire pour comprendre notre expérience." At the bottom, there are buttons for "Regenerate response", "Web access", "10 results", "Any time", "France", and "Défaut prompt".

ChatGPT, GPT4 (OpenAI), 2023

Objectif ou écueil à éviter ?



CommitStrip.com

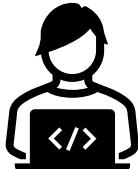
« Des spécifications très complètes et très précises »

<http://www.commitstrip.com/fr>

/2016/08/25/a-very-comprehensive-and-precise-spec/

Langages informatiques et langage naturel

```
/**  
 * Simple HelloButton() method.  
 * @version 1.0  
 * @author john doe <doe.j@example.com>  
 */  
HelloButton()  
{  
    JButton hello = new JButton( "Hello, wor  
hello.addActionListener( new HelloBtnList  
  
// use the JFrame type until support for t  
// new component is finished  
JFrame frame = new JFrame( "Hello Button"  
Container pane = frame.getContentPane();  
pane.add( hello );  
frame.pack();  
frame.show();  
} // display the fra
```



The screenshot shows the Piet programming environment. On the left is a color palette titled 'Colors'. In the center is a 'Buffers' panel with an input value of 5 and an empty output field. Below it is a 'Canvas' panel displaying a pixelated version of the 'Hello, world!' logo. To the right is a 'Debug' panel with sections for registers and stack, and buttons for step operations. A list of operations (ops) is also visible.

ops:

- push
- pop
- subtract
- multiply
- mod
- not
- pointer
- switch
- in(number)
- in(char)
- out(number)
- out(char)

10nnnjjabffntnjataitnnjj
abaitntddaaiaiaaddatataat

Registers:

- dp:
- cc:
- Δh:
- Δl:

Stack:

next op:

Debug buttons:

- Step Forward
- Step Backward
- Run

« Dis Siri, programme-moi un réveil pour 16h »

« Un chat ailé dans le ciel de minuit »

« A clock eating time »

Exemple avec un text-to-image



Prompt :

« An artist is typing a text on his computer, a cloud scale emerges from the computer, and at the end of the scale begins a bucolic world »

DALL-E 2 (OpenAI)

Différentes sortes de « Text-to »

Text-to-Text

- *ChatGPT ; Eliza ; Siri ; Bixbi ; Bard*

Text-to-Picture/Image

- *VizStory ; Dall-E ; Midjourney ; Stable Diffusion ; Muse*

Text-to-Video

- *Imagen (Google) ; Make-A-Video (Meta)*

Text-to-Speech/Voice

- *Vonage ; Lifelike ; TTSReader ; Murf AI*

Text-to-Music/Song

- *Mubert ; MusicLM; Melobytes; Typatone*

Text-to-Mesh

- *Text2Mesh (threedle) ; CLIP-Mesh ; Point-E ; Shape-E*

Text-to-Animation

- *SceneMaker (/CONFUCIUS) ; projet GITAN*

Text-to-Scene

- *WordsEye : Text2Scene (Stanford) ; Gaudi ; MAV3D ;*

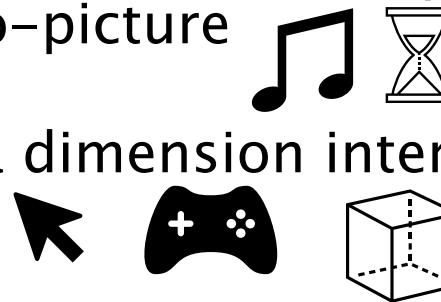
Text-to-Game

- *MarioGPT ; Dwarf Quest (Rafael Bidara)*

Liste non-exhaustive

Particularités des text-to-scene

- ▶ Visualisation à 360°*
- ▶ Exploration de la scène (curseur, avatar ...)
- ▶ Sons
- ▶ Exportabilité granulaire des assets
- ▶ Interaction
 - Comportements autonomes
 - Stockage des éléments
 - Complexité de calcul
 - ...
- ▶ Les **text-to-video** apportent la dimension temporelle et sonore par rapport aux text-to-picture
- ▶ Les **text-to-scene** apportent la dimension interactive par rapport aux text-to-video

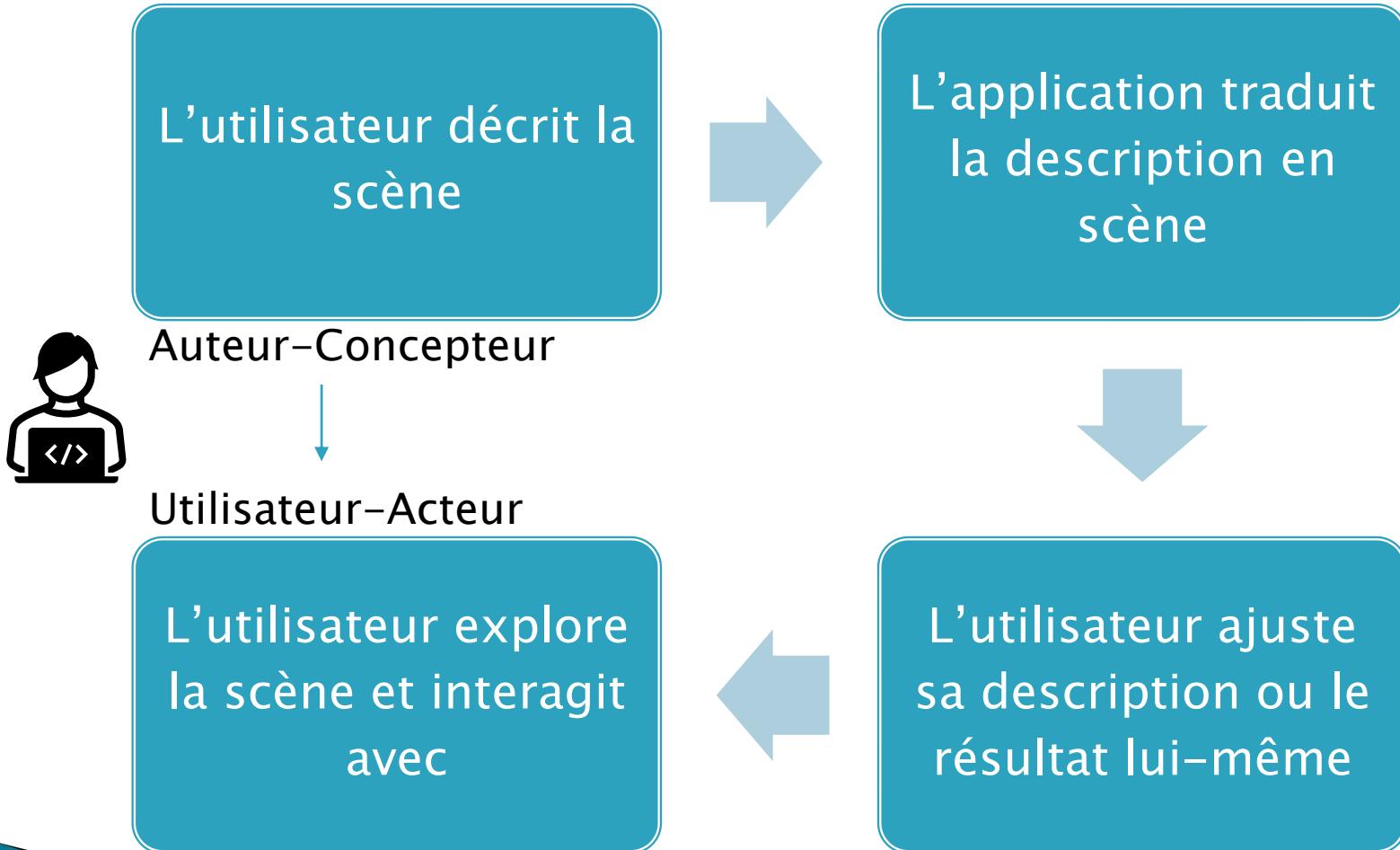


Text-to-Scene en scène

Domaine	Auteurs	Système
Jeu vidéo	(Linden et al., 2013)	Dwarf Quest
Simulation d'accident de la route à partir d'un rapport	(Dupuy et al., 2001)	Carsim
Mise en scène théâtrale collaborative	(Andriamarozakaniaina, 2012)	DRAMA
Progrès en littératie	(Coyne et al., 2011)	Wordseye
Aide aux troubles en apprentissage	(Rugma and Sreeram, 2016) ; (Vyas et al., 2020)	Enabl
Smart learning	(Delair et al., 2016)	/
Smart education	(Pardhi et al., 2021)	/
'Envision': Visual Representation System	(Jaybhay et al. (2020))	Envision
Apprentissage des langues	Projet GITAN & UnimaSoft	Projet GITAN & U Slic

Liste non-exhaustive

Processus du texte à la scène



Exemple avec un text-to-scene

wordseye™

CREATE GALLERY Help Sign up Login

How to type a picture (recommended)

Clear

3 yellow ducks are aligned on water.
They are facing west.
A pink crocodile is near the ducks.
A huge water sphere is 10 feet behind the ducks.
A 10 feet tall salmon is -30 feet left of the sphere.
It leans 50 degrees to the southeast.
It faces northwest.

show tips Examples Scene Objects My Images Library

Pick alternative objects

water	crocodile	water	sphere
salmon	sky texture	duck	duck
duck	ground		

wordseye.com

Notifications

Free Text Template Undo **Display**

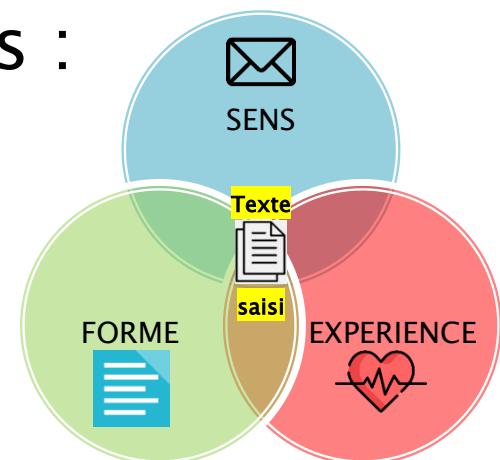
Sign Up to save scenes, add effects, more...

Créativité humaine ?

- ▶ Oppenlaender (2022) analyse de la créativité sur Midjourney (text-to-image), en se basant sur les 4 « P » de Mel Rhodes (Product, Person, Process, Press)
 - Midjourney est centré « Produit » mais ne brasse pas toute la créativité humaine impliquée de ce processus
 - Exemple de paroles de musique recopiées.
 - Exemple du smiley ou de la lettre unique en prompt
 - Oppenlaender évoque la part de collaboration dans la communauté des utilisateurs pour parfaire ses prompts, l'originalité du texte, etc.

Autour du texte en entrée

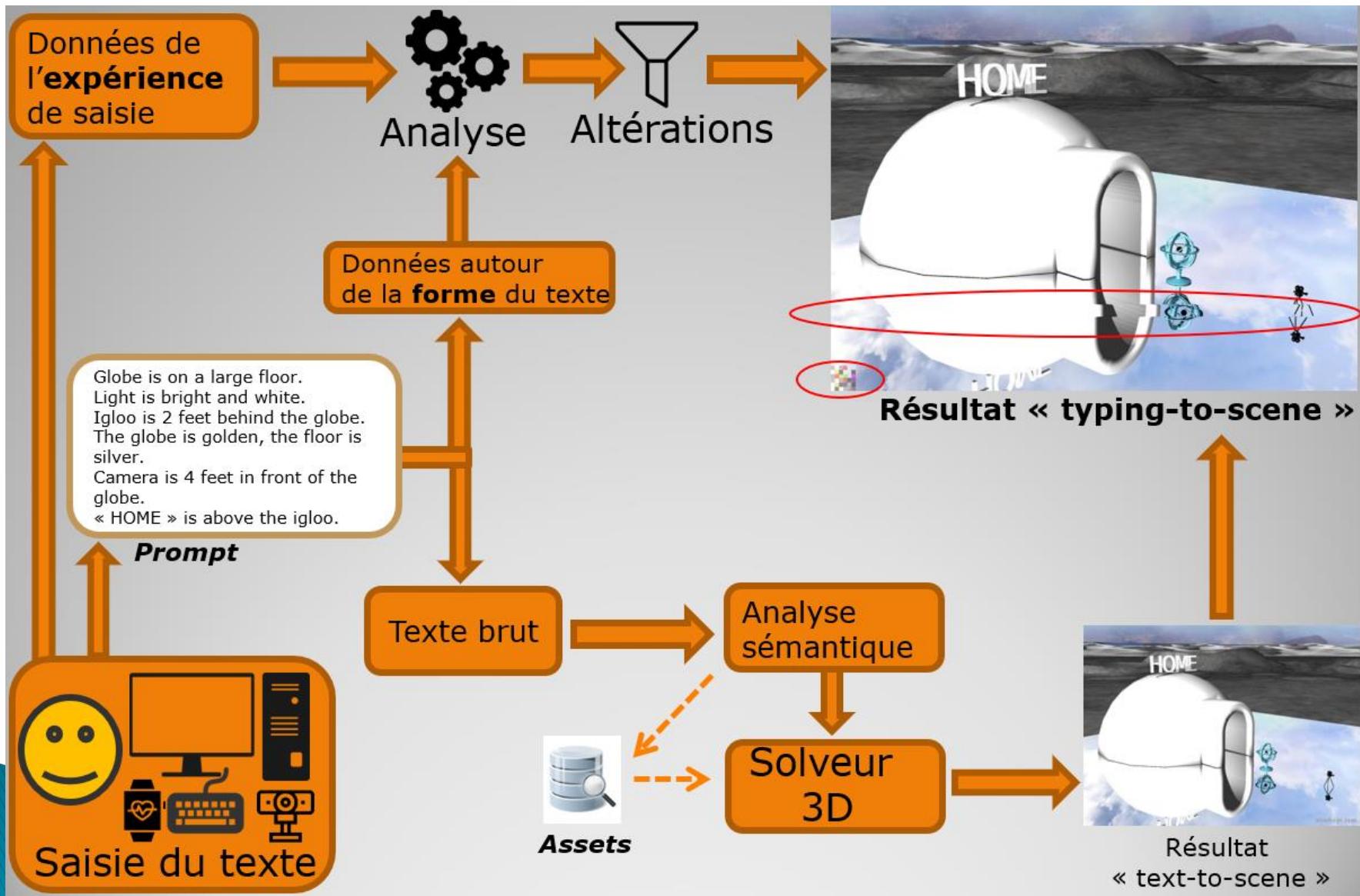
- ▶ Les text-to-image/scene illustrent le sens du prompt, mais sans tenir compte de la forme et de l'expérience de saisie.
- ▶ Nous considérons que dans la conception d'un texte saisi par un humain, trois éléments sont fondamentalement présents :
 - le sens du texte
 - la forme du texte
 - l'expérience de saisie du texte



Du text-to-scene au typing-to-scene

- ▶ Nous proposons un système « Typing-to-Scene » : son objectif est de mettre en valeur la présence humaine dans la saisie du prompt en intégrant dans l'image générée une trace du triptyque
Sens–Forme–Expérience
- ▶ A partir d'un export text-to-scene, nous utilisons les données récoltées depuis le début de la saisie du prompt pour altérer légèrement la scène en post-processing sans trahir la cohérence entre la scène et sa description textuelle.

Architecture du « Typing-to-Scene »



La forme du texte

- ▶ Nous considérons la forme du texte selon 3 perspectives :
 - Le texte en tant que forme esthétique (mise en page, ...)
 - Le texte en tant qu'ensemble de caractères et de mots (approche statistique)
 - Le texte en tant que code (approche stéganographique)

La forme esthétique du texte



FORMES D'ÉCRITURES

À chaque époque, son écriture.
Alors l'écriture, reflet de son époque ?

La première écriture, l'écriture manuelle, incorpore une dimension humaine, inconstante et variée. De la main, à son gré, les lignes vont montantes, descendantes, au bout de la feuille ou à mi-parcours, selon les personnes, selon les jours ou les humeurs.

Puis, de la tête d'acier vint l'écriture mécanique, accélérant la cadence, encore avec une dimension humaine, de par la force de frappe, de par les libres césures. L'écriture ne glisse plus en silence, elle crépite. Elle se fait entendre.

Enfin, la troisième, l'écriture informatique. Normée, elle se diffuse plus vite, plus loin, en tout lieu, en tout temps. Moins audible, aux angles vifs, tirée au cordeau, elle est efficace, rapide parce que dépersonnalisée.

L'écriture informatique met fin aux oscillations, elle va droit. La forme s'uniformise. Serait-elle le reflet d'une société régentée par la norme ou serait-ce par souci d'efficacité ?

Figure 3 — Claude-CLd, reproduction agrandie du texte *Les 3 Écritures*

MEROFLIGHTCOMEINTOTHEROOMFROMUNDERTHEDOORLORDSAVE METHINK
ATTINGATSTREETCORNERSSAVAGESOUTRIGHTMANYOFWHOMYETCARRYON
DALLTHOUGHINSLIGHTMOTIONSTILLMAINTAINSAPERMANENTOBLIQUIT
ULLURKEDASTRONGDESIRETOSEESOMETHINGMOREOFCRISTENDOMTHAN
ESTHEAREABEFORETHEHOUSEWASPAVEDWITHCLAMSHELLSMRSHUSSEYWO
REDANDSEVENTYSEVENGOLDDOUBLELOONSANDSOITHOUGHTATTHETIMEWHY
THEDARTEDTHEIRONRIGHTOVEROLDBILLADSBROADBRIMCLEANACROSST
APSTANBLOODANDTHUNDERJUMPWADTHENEXTCOMMANDANDTHECREWSPRA
DHERBREADANDNOTTOBEFOOLISHLYWASTEDWHEREFOREHEHADNOFANCYF
WAYSOMECONSIDERATINGTOUCHOFHUMANITYWASINHIMFORATTIMESLIK
FTEDWITHSUCHWONDROUSPOWERANDVELOCITYINSWIMMINGASTODEFYAL
LESSSQUALLOFAHORNPIPERIGHTOVERTHEGRANDTURKSHEADANDTHENBY
GSENTIMENTINTARANDELURBERCHILLEDEHAROLDNOTUNFREQUENTLYPERC
HEARTSASLIGHTTOLOVEASGAYANDFLEETINGASBUBBLESTHATSWIMONT
LRAGEANDHATEFELTBYHISWHOLERACEFROMADAMDOWNANDTHENASIFHIS
ASINESSENCEWHITENESSISNOTSOMUCHACOLORASTHEVISIBLEABSENCE
SAGAINSTTHESKYWASITNOTSOONDONMIGUELTHOUCHILIANWHALEMARKED
AVINGLYWORKINGTOGETHERTHESTRAIGHTWARPOFNCESSITYNOTTOBES
MORTALINTHESEJAWSOEDEATHINVAINWEHAILEDTHEOTHERBOATSASWEL
SIDEDARTEDAWAYWITHWHATSEEMEDSHUDDERINGFINSANDRANGEDTHEMS
FSEAUSAGE SANDTHEINSTINCTIVELOVEOFNEATNESSINSEAMENSOMEOFW
HAINSALLATONCESHOUTEDOUTTHERESHROLLSTHERESHROLLSJESUWH
HALLATVERSAILLESWHERETHEBEHOLDERFIGHTSHISWAYPELLMELLTHRO
DSOUNDSODEEPASTOTHREATENTOCARRYOFFTHEENTIRELINEORIGINAL
HESHIPISTOBEMOOREDТИEDBYTHEHEADTOTHESTERNANDBYTHETAILTOT

Exemples d'utilisation de la forme du texte

- ▶ – En considérant la forme du texte :
 - – le décompte des sauts de lignes peut définir des normes de taille dans la scène ou pour pondérer la distance entre des éléments 3D dans la scène
 - – le décompte des espaces superflus dans le texte peut définir un nombre de pixels aléatoires "abîmés" (glitch) dans le rendu visuel
 - – méthodes de stéganographie : application de la méthode des Suites de Lettres Équidistantes de Haralick (2006) pour trouver un mot-clé et altérer la scène en fonction de ce mot-clé (filigrane, glitch...).

L'expérience de saisie

- ▶ L'expérience de saisie correspond à un contexte éphémère qui peut conditionner la façon d'écrire (Chiche, 2013) et porter une différence entre deux textes constitués de façon similaire.



- Données relative au temps : date & heure de début et de fin de saisie, temps écoulé, pauses longues...



- Données liées au clavier : ensemble des touches pressées (comme les suppression de caractère)



- Données concernant l'utilisateur : captation de la fréquence cardiaque, détection vidéo d'émotion faciale...

Exemples d'utilisation de l'expérience de saisie

- ▶ En considérant l'expérience de l'utilisateur :
 - Chaque appui sur une flèche directionnelle déplace la position initiale de la caméra avec un angle et une direction
 - Chaque touche de correction (backspace ou suppression) « supprime » la couleur d'un pixel aléatoire
 - La date et la durée de la saisie permettent d'établir une matrice de chiffres qui associés à une couleur, donne une image « signature temporelle » formée de pixels, à insérer avec dans le HUD (Head-up display).

Méthode pour le tampon temporel

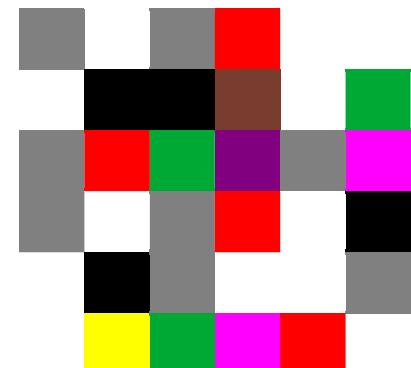
Year_start	Year_start	Year_start	Year_start	Difference of days	Difference of days
Month_start	Month_start	Day_start	Day_start	Difference of hours	Difference of hours
Hour_start	Hour_start	Minute_start	Minute_start	Seconds_start	Seconds_start
Year_end	Year_end	Year_end	Year_end	Difference of minutes	Difference of minutes
Month_end	Month_end	Day_end	Day_end	Difference of seconds	Difference of seconds
Hour_end	Hour_end	Minute_end	Minute_end	Seconds_end	Seconds_end

Start 19/January/2023 at 23h57m28s
 End 20/January/2023 at 04h58m30s

2	0	2	3	0	0
0	1	1	9	0	5
2	3	5	7	2	8
2	0	2	3	0	1
0	1	2	0	0	2
0	4	5	8	3	0

Color-Number correspondance table

0	1	2	3	4
5	6	7	8	9



Exemples d'utilisation de l'expérience de saisie (2)

- ▶ Données physiologiques et émotionnelles :
 - Captation de la fréquence cardiaque de l'utilisateur durant sa saisie par une montre intelligente connectée. L'enregistrement du rythme de pulsation peut être utilisé en boucle pour faire légèrement briller l'image de « signature temporelle » sur ce même rythme.
 - Captation vidéo du visage de l'auteur durant la saisie du texte pour appliquer un algorithme de détection des émotions faciales (Kang et al., 2020; Microsoft, 2021). De cette manière, il est possible de faire varier la luminosité de la scène générée en fonction de l'évolution de la joie ou de la tristesse détectée.

Résultat de ces traitements

Globe is on a large floor.

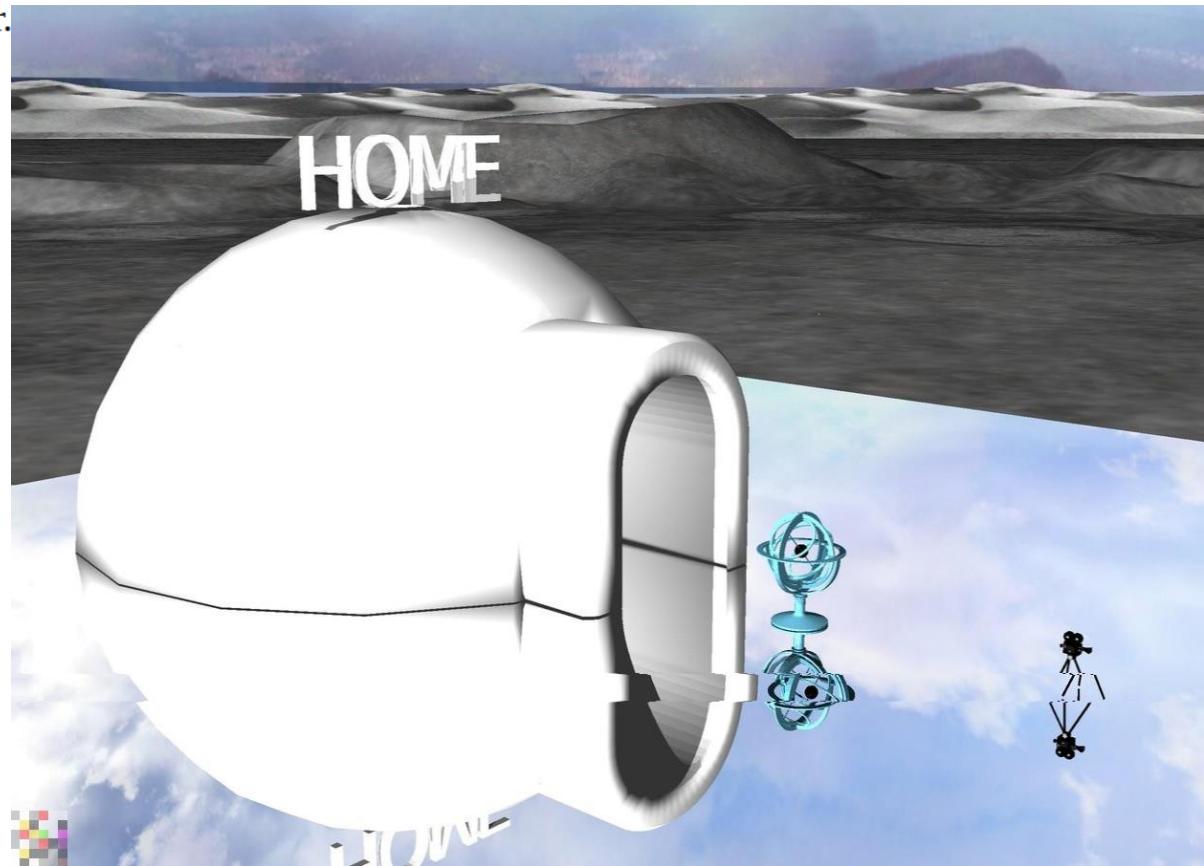
Light is bright and white.

Igloo is 2 feet behind the globe.

The globe is golden, the floor is silver.

Camera is 4 feet in front of the globe.

"HOME" is above the igloo.



Conclusion

- ▶ Notre proposition de système typing-to-scene tend à se démarquer de l'aspect utilitaire du text-to-scene par sa composante créative, mettant en valeur la présence de l'utilisateur, à l'interface entre art et technologies.
- ▶ La prise en compte de la forme du texte ainsi que des données non-textuelles recueillies au cours de la saisie offrent la perspective de modifier la scène avec des éléments uniques du triptyque sens-forme-expérience, qui font ainsi office de signature discrète au sein de l'œuvre.

Perspectives

- ▶ Pour éviter que la nécessité de saisir le texte ne mette à l'écart les personnes en situation d'illettrisme ou d'illectronisme, il est possible de penser à la dictée orale via une application speech-to-text / text-to-speech, lié à un text-to-scene.
- ▶ Cette réflexion peut être étendue dans le cadre d'un "speeching-to-scene" où la nature des données captées mène à repenser les façons d'altérer la scène lors de sa génération.
- ▶ L'export de cet environnement virtuel pourrait être réalisé en réalité virtuelle (VR), favorisant l'immersion, où l'on se permettrait d'imaginer qu'une commande vocale puisse permettre d'intégrer de nouveaux éléments dans la scène (ou d'agir dessus) en temps réel.

Futures reflexions

- ▶ Réflexion sur l'aspect collaboratif dans l'usage des systèmes typing-to-scene comme l'écriture collaborative (inspiré de (Oppenlaender, 2022))
- ▶ Approfondir l'analyse du typing-to-scene au regard de la notion de créativité, avec les 4 « P » de Rhodes ou encore par le benchmark proposé par (Russo, 2022) pour les systèmes text-to-image.
- ▶ Etendre notre étude des caractéristiques du typing-to-scene au regard de domaines diverses tels que la philosophie ou encore la psychologie des arts.

Merci pour votre attention

À l'interface entre l'art et les technologies :
proposition d'un système Typing-to-Scene

Lionel LALOUM & Alain LIORET

laloum.lionel@gmail.com

alainlioret@gmail.com



edesta

Ecole Doctorale Esthétique, Sciences et Technologies des Arts

UNIVERSITÉ
PARIS 8

Ecole Doctorale
d'Esthétique, Sciences et
Technologie des Arts

Références principales

- ▶ Bishop, P., and C.-Y. Chen (2017). La génération procédurale en temps réel dans le jeu vidéo : présentation, concepts et confrontation au gameplay, Revue ATIINREV 5, 183–202.
- ▶ Chiche, S. (2013). Les petits rituels des romanciers, Sciences Humaines 253, no. 11, 24–24.
- ▶ Coyne, R. E. (2017). Painting Pictures with Words – From Theory to System.
- ▶ Coyne, R. E., C. Schudel, M. Bitz, and J. B. Hirschberg (2011). Evaluating a Text-to-Scene Generation System as an Aid to Literacy.
- ▶ Delair, R., H. Sojitra, K. Patel, K. Bhatt, K. Gharia, and R. Mahajan (2016). Text to Scene conversion for smart learning, in 2016 IEEE International Conference on Advances in Electronics, Communication and Computer Technology (ICAECCT), IEEE, 76– 81.
- ▶ Gaildrat, V. (2003). Modélisation déclarative d'environnements virtuels : Création de scènes et de formes complexes par l'énoncé de propriétés et l'emploi d'interactions gestuelles, Habilitation à diriger des recherches, Université Paul Sabatier (Toulouse III).
- ▶ Groupe de recherche GITAN <<http://gitan.polymtl.ca/www/fr/index.html>> (accessed June 4, 2017).
- ▶ Haralick, R. M. (2006). Basic concepts for testing the Torah code hypothesis, in Pattern Recognition, 2006. ICPR 2006. 18th International Conference on, IEEE, 104–109.
- ▶ Hassani, K., and W.-S. Lee (2016). Visualizing Natural Language Descriptions: A Survey, ACM Computing Surveys 49, no. 1, 1–34, doi: 10.1145/2932710.
- ▶ Jaybhay, S., D. Jain, S. Bhagat, P. Chaudhari, and V. Dbrito (2020). 'Envision': Visual Representation System, in 2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA), IEEE, 62–66.
- ▶ Linden, R., R. Lopes, and R. Bidarra (2013). Designing procedurally generated levels, 47.
- ▶ Lombardo, E., and S. Agostinelli (2014). Entre virtualité & réalité. Quelles frontières numériques, Saleh., Bouhāï N., & Hachour H., Les Frontières numériques, 17– 32.
- ▶ Microsoft (2021). Perceived Emotion Recognition Using the Face API – Xamarin: <<https://learn.microsoft.com/en-us/xamarin/xamarin-forms/data-cloud/azurecognitive-services/emotion-recognition>> (accessed February 6, 2023).
- ▶ Midjourney (n.d.): <<https://midjourney.com/home/?callbackUrl=%2Fapp%2F>> (accessed February 3, 2023).
- ▶ Nemmich, M. A. (2016). L'accessibilité numérique pour les personnes souffrant d'un handicap visuel, report, Université Mustapha Stambouli de Mascara, Algérie.
- ▶ OpenAI (2023). GPT-4 Technical Report, arXiv:2303.08774, arXiv, doi : 10.48550/arXiv.2303.08774.
- ▶ Oppenlaender, J. (2022). The Creativity of Text-to-Image Generation, in Proceedings of the 25th International Academic Mindtrek Conference New York, NY, USA, Association for Computing Machinery, Academic Mindtrek '22, 192–202, doi: 10.1145/3569219.3569352.
- ▶ Pardhi, V., K. Shah, J. Vaghasiya, and V. Hole (2021). Generating a scene from text for smart education, in 2021 International Conference on Communication information and Computing Technology (ICCICT), IEEE, 1–6.
- ▶ Ramesh, A., M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, M. Chen, and I. Sutskever (2021). Zero-Shot Text-to-Image Generation, arXiv:2102.12092, arXiv, doi : 10.48550/arXiv.2102.12092.
- ▶ Rugma, R., and S. Sreeram (2016). Text-to-Scene Conversion System for Assisting the Education of Children with Intellectual Challenges, IJIRSET 5, no. 8, 15041– 15046, doi: DOI:10.15680/IJIRSET.2016.0508144.
- ▶ Russo, I. (2022). Creative Text-to-Image Generation: Suggestions for a Benchmark, in Proceedings of the 2nd International Workshop on Natural Language Processing for Digital Humanities Taipei, Taiwan, Association for Computational Linguistics, 145–154.
- ▶ Ulinski, M., B. Coyne, and J. Hirschberg (2018). Evaluating the WordsEye text-to-scene system: imaginative and realistic sentences, in Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018).
- ▶ Vyas, G., P. Kadam, and M. Thaker (2020). Enabl : Visual Aid for Dyslexic based on Natural language processing, in 2020 International Conference on Smart Electronics and Communication (ICOSEC), IEEE, 52–57.
- ▶ Winograd, T. (1971). Procedures as a Representation for Data in a Computer Program for Understanding Natural Language, 235, MIT, MIT AI Technical Report.
- ▶ Yashaswini, S., and S. S. Shylaja (2021). Metrics for Automatic Evaluation of Text from NLP Models for Text to Scene Generation, 4, European Journal of Electrical Engineering and Computer Science 5, no. 4, 20–25, doi: 10.24018/ejece.2021.5.4.341.

Annexes & Bonus

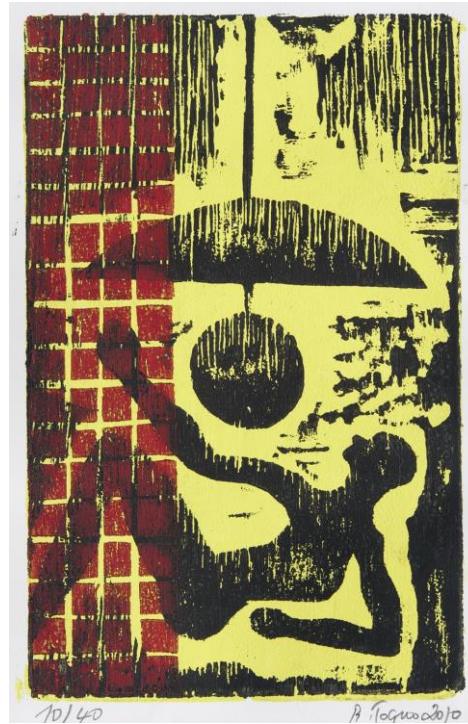


Exemples de liens entre texte et image dans l'art



La trahison des images

René Magritte, 1929
Musée d'Art du comté
de Los Angeles



The Long Night
Barthélémy Toguo
2010, bois gravé
(Galerie Lelong)



Caïn tuant Abel
Anonyme romain (peintre)
4e quart du 17e siècle

Réalisations à partir d'un texte biblique

And Cain spoke to Abel his brother, and it came to pass when they were in the field, that Cain rose up against Abel his brother and slew him. - @lee_onel (fast)



Caïn tuant Abel
Anonyme romain (peintre)
4e quart du 17e siècle

© Collection Musée d'art Roger-Quilliot [MARQ]
Clermont-Ferrand

Créativité ?

RUSSO	OPPENLAENDER
Psychologie des arts, Neurologie, Philosophie...	Rhodos' 4 « P » (Person, Press, Product, Process)
Propositions de Benchmark pour les T2i	Aspect crowdsourcing pour midjourney
Les images générées par les systèmes T2I sont-elles perçues comme créatives par les humains ?	Génération progressive de l'image
La créativité peut-elle être mimée de manière computationnelle et augmentée de manière empirique ?	Pas de différence entre la saisie spontanée ou la copie de paroles de musiques existantes. Ni entre un texte sémantique et un simple smiley
Etudier si des prompt plus créatifs donnent lieu à des images plus créatives ?	

La forme esthétique du texte

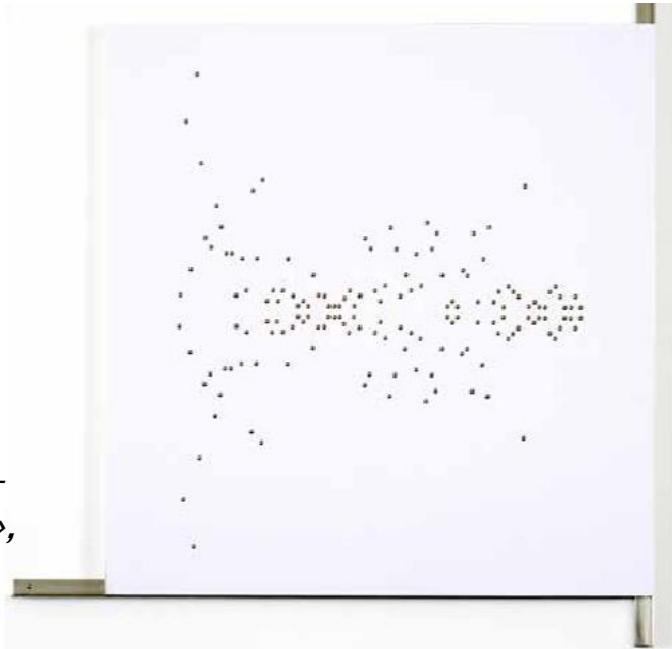
Claude-CLd,
« *AU RYTHME DU RECIT –
Enregistrements visuels* »,
Narration

7.1 Points sur blanc & fluo -
La Partie de trictrac, 2014
clous, carton-plume,
aluminium, 49 x 51 cm

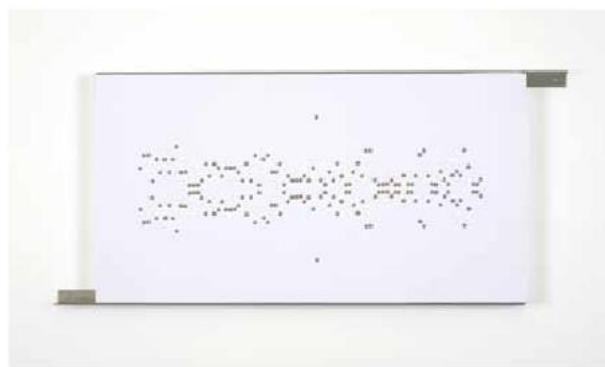
7.2 Points sur blanc & fluo -
Le Horla, 2014
clous, carton-plume,
aluminium, 44 x 30,3 cm

Claude-CLd
© Raphaële Krieger

De part et d'autre d'une ligne horizontale centrale, un clou est enfoncé. Plus le paragraphe est long plus les clous sont symétriquement éloignés de l'axe horizontal. Le sens de lecture va de gauche à droite, à gauche le premier paragraphe, à l'extrême droite, le dernier. C'est en reliant ces points que s'est esquissé un autre dessin spécifique au récit que l'on peut visualiser à la page 97 (Signature).



7.1.a



7.2.a



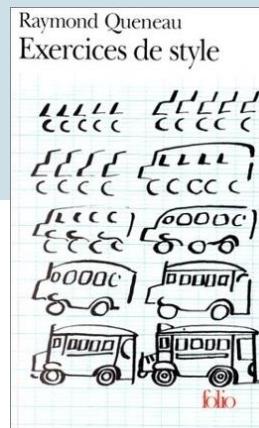
Claude-CLd,
« *Un art visuel cognitif, au rythme
de la longueur des paragraphes
des textes littéraires* »,
Narration

Le texte comme ensemble de caractères

► Exercice de style (R. Queneau)

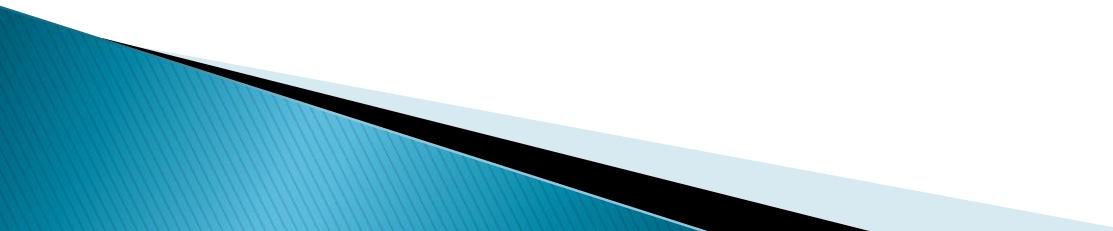
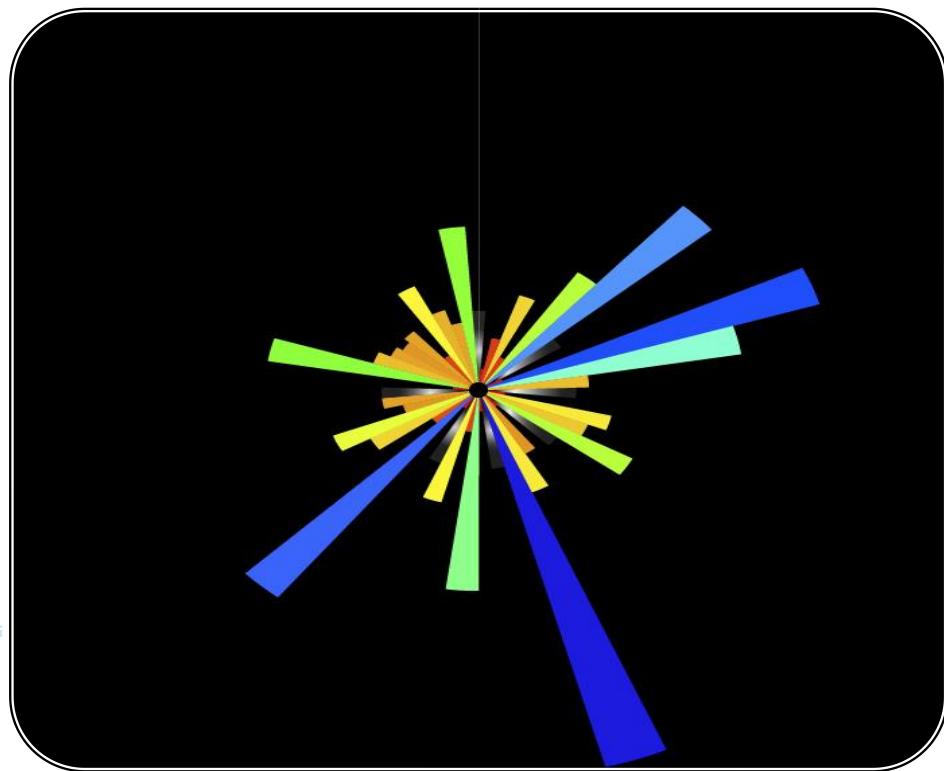
L'arc-en-ciel

Un jour, je me trouvai sur la plate-forme d'un autobus violet. Il y avait là un jeune homme assez ridicule: cou indigo, cordelière au chapeau. Tout d'un coup, il proteste contre un monsieur bleu. Il lui reproche notamment, d'une voix verte, de le bousculer chaque fois qu'il descend des gens. Ceci dit, il se précipite, vers une place jaune, pour s'y asseoir. Deux heures plus tard, je le rencontre devant une gare orangée. Il est avec un ami qui lui conseille de faire ajouter un bouton à son pardessus rouge.



Précisions

A 12h27, dans un autobus de la ligne S, long de 10 mètres, large de 2,1, haut de 3,5, à 3 km. 600 de son point de départ, alors qu'il était chargé de 48 personnes, à 12 h. 17, un individu de sexe masculin, âgé de 27 ans 3 mois 8 jours, taille de 1 m 72 et pesant 65 kg et portant sur la tête un chapeau haut de 17 centimètres dont la calotte était entourée d'un ruban long de 35 centimètres, interpelle un homme âgé de 48 ans 4 mois 3 jours et de taille 1 m 68 et pesant 77 kg., au moyen de 14 mots dont l'énonciation dura 5 secondes et qui faisaient allusion à des déplacements involontaires de 15 à 20 millimètres. Il va ensuite s'asseoir à quelque 2 m. 10 de là. 118 minutes plus tard il se trouvait à 10 mètres de la gare Saint-Lazare, entrée banlieue, et se promenait de long en large sur un trajet de 30 mètres avec un camarade âgé de 28 ans, taille 1 m. 70 et pesant 71 kg. qui lui conseilla en 15 mots de déplacer de 5 centimètres, dans la direction du zénith, un bouton de 3 centimètres de diamètre.

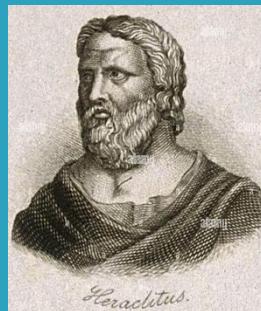


Expérience de saisie

- ▶ Ainsi, lorsque deux textes sont identiques dans le sens et dans la forme, leurs données d'expérience de saisie porteront la différence de l'un à l'autre.

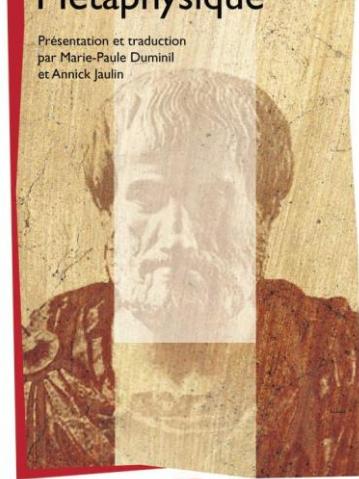
« On ne se baigne jamais deux fois dans le même fleuve »

Héraclite



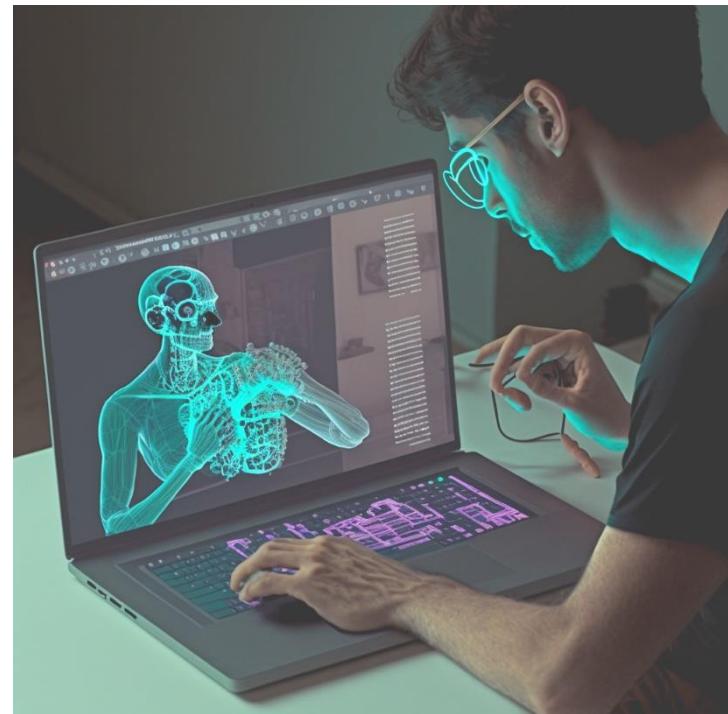
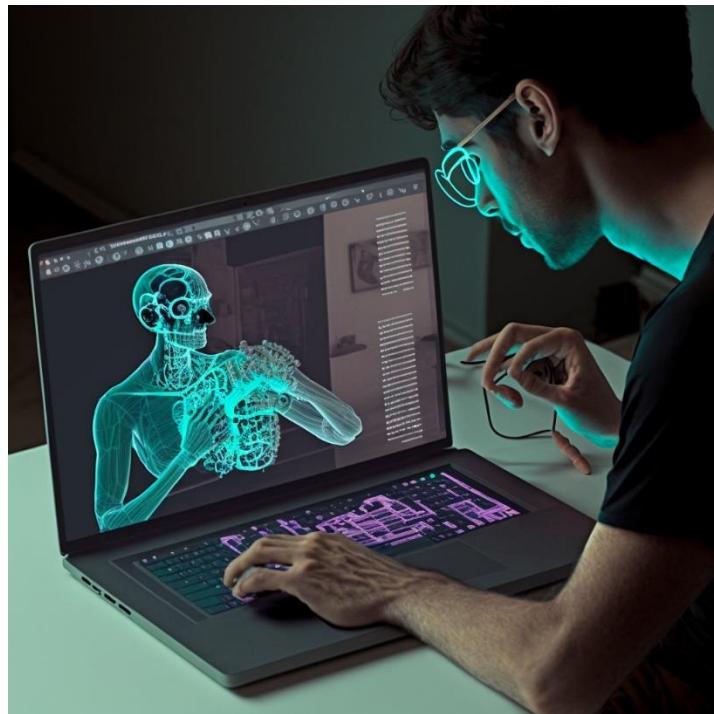
Aristote
Métaphysique

Présentation et traduction
par Marie-Paule Duminiel
et Annick Jaulin



Altération : SadOrJoy

- ▶ Lors de la captation vidéo (pendant la saisie du prompt), si l'émotion la plus valorisée est Joy, on augmente la luminosité de 25% et on diminue d'autant le contraste

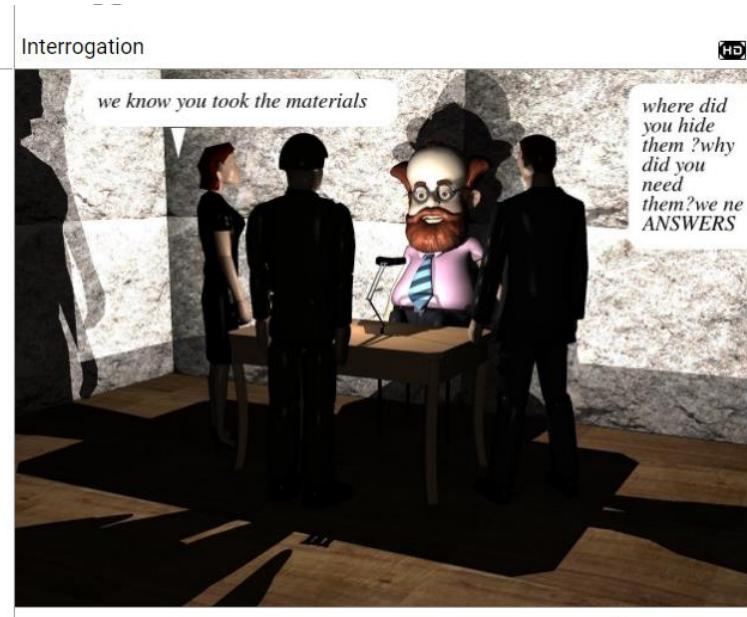


Prompt :

*An artist is typing a text
on his laptop with webcam*

D'autres scènes de WordsEye

a table .
a chair is 1 feet behind the table .
a first man is -22 inches above the chair .
the table is 3 feet tall .
a light is on the table .
the light is facing the first man .
a lamp is on the table .
the lamp is facing the first man .
a white light is above the lamp .
a first wall is 1 feet behind the chair .
a second wall is left of the first wall .
the second wall is facing left .
the first wall is 50 feet tall .
the second wall is 50 feet tall .
the second wall is -8 inches in front of the fir
a 7 feet tall woman is left of the table .
the woman is facing northeast .
a second 7 feet tall man is right of the table .
the man is facing northwest .
a third 7 feet tall man is in front of the table
the third man is facing north .
the first wall is rock .
the second wall is rock .
it is night .
the ground is wood .
the camera light is dark .



watcher570 typed a picture: a table.a chair is 1 feet behind the table.a first man is -22 inches above the chair.the table is 3 feet tall.a light is on the table.the light is facing the first man.a lamp is on the table the lamp is facing the first man.a white light is above the lamp.a first wall is 1 feet behind the chair.a second wall is left of the first wall.the second wall is fac

Tags: #HD #totwDialog

Views: 23

Effects: Talk Balloon, Talk Balloon



Share to



[watcher570](#) (follow)

D'autres scènes de WordsEye

wordseye™

CREATE GALLERY Help Sign up Login

How to type a picture (recommended)

bill gates is next to leonard nimoy .
the ground is grass .
a [flower] wall is 3 feet behind
nimoy .

Clear

1. Edit template and Click Display (Note: Free Text will exit Template mode) 2. Adjust view 3. Swap objects (menu)



Click to pick alternatives

head human body human body head

ground grass [flower] wall

sky texture bill gates leonard ni...

Notifications

Free Text (radio button) Undo (button) Display (button)

Sign Up to save scenes, add effects, more...

wordseye.com

L'avenir ?

