

La course à la technologie dans l'industrie de la construction

Une vision holistique.



Le premier virage numérique

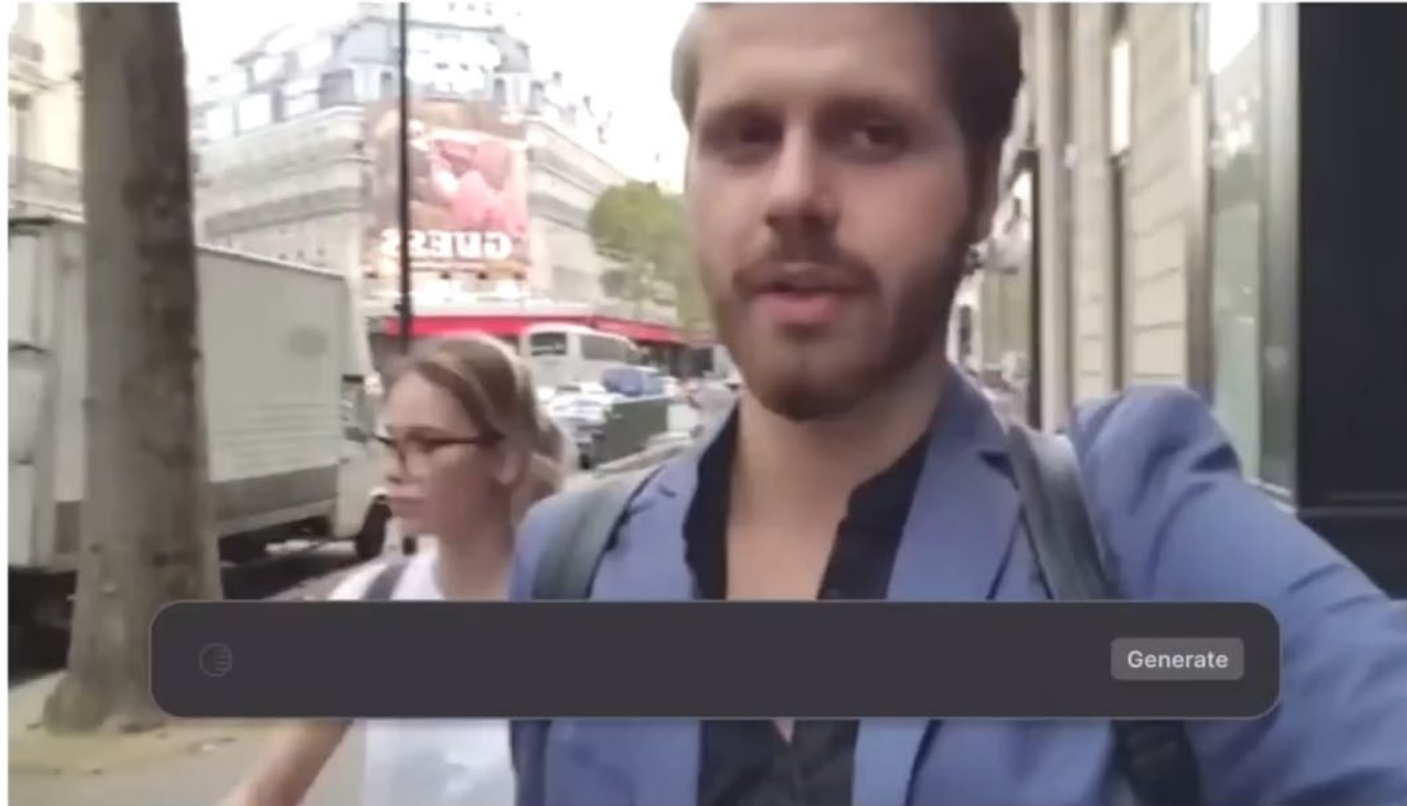
La numérisation de la conception.

Pour commencer..

- Architecte de renommée internationale ;
- Un type de construction ;
- Autre détails..

L'engouement pour l'intelligence artificielle dans l'industrie

Plusieurs cas d'utilisations de l'iA en construction, mais celui du **traitement du langage naturel (NLP)** s'est grandement démarqué en 2023.



Qu'en est-il de la 3D utilisant le traitement du langage naturel ?



Encore au commencement..

Mais une recherche qui ne date pas d'hier.

[Submitted on 16 Dec 2022]

Point-E: A System for Generating 3D Point Clouds from Complex Prompts

Alex Nichol, Heewoo Jun, Prafulla Dhariwal, Pamela Mishkin, Mark Chen

Point-E: A System for Generating 3D Point Clouds from Complex Prompts

Alex Nichol¹, Heewoo Jun¹, Prafulla Dhariwal¹, Pamela Mishkin¹, Mark Chen¹

Abstract

While recent work on text-conditional 3D object generation has shown promising results, the state-of-the-art methods typically require multiple GPU-hours to produce a single sample. This is in stark contrast to state-of-the-art generative image models, which produce samples in a number of seconds or minutes. In this paper, we explore an alternative method for 3D object generation which produces 3D models in only 1-2 minutes on a single GPU. Our method first generates a single synthetic view using a text-to-image diffusion model, and then produces a 3D point cloud using a second diffusion model which conditions on the generated image. While our method still falls short of the state-of-the-art in terms of sample quality, it is one to two orders of magnitude faster to sample from, offering a practical trade-off for some use cases. We release our pre-trained point cloud diffusion models, as well as evaluation code and models, at <https://github.com/openai/point-e>.

1. Introduction

With the recent explosion of text-to-image generative models, it is now possible to generate and modify high-quality images from natural language descriptions in a number of seconds (Hamesh et al., 2021; Ding et al., 2021; Nichol et al., 2021; Ramesh et al., 2022; Galois et al., 2022; Yu et al., 2022; Saharia et al., 2022; Feng et al., 2022; Balaji et al., 2022). Inspired by these results, recent works have explored text-conditional generation in other modalities, such as video (Hong et al., 2022; Singor et al., 2022; Ho et al., 2022a) and 3D objects (Jain et al., 2021; Poole et al., 2022; Lin et al., 2022a; Sanghi et al., 2021, 2022). In this work, we focus specifically on the problem of text-to-3D generation, which has significant potential to democratize 3D content creation for a wide range of applications such as

virtual reality, gaming, and industrial design. Recent methods for text-to-3D synthesis typically fall into one of two categories:

1. Methods which train generative models directly on paired text, 3D data (Chen et al., 2018; Mirdal et al., 2022; Fu et al., 2022; Zeng et al., 2022) or unlabeled 3D data (Sanghi et al., 2021, 2022; Watson et al., 2022). While these methods can leverage existing generative modeling approaches to produce samples efficiently, they are difficult to scale to diverse and complex text prompts due to the lack of large-scale 3D datasets (Sanghi et al., 2022).
2. Methods which leverage pre-trained text-image models to optimize differentiable 3D representations (Jain et al., 2021; Poole et al., 2022; Lin et al., 2022a). These methods are often able to handle complex and diverse text prompts, but require expensive optimization processes to produce each sample. Furthermore, due to the lack of a strong 3D prior, these methods can fall into local minima which don't correspond to meaningful or coherent 3D objects (Poole et al., 2022).

We aim to combine the benefits of both categories by guiding a text-to-image model with an image-to-3D model. Our text-to-image model leverages a large corpus of (text, image) pairs, allowing it to follow diverse and complex prompts, while our image-to-3D model is trained on a smaller dataset of (image, 3D) pairs. To produce a 3D object from a text prompt, we first sample an image using the text-to-image model, and then sample a 3D object conditioned on the sampled image. Both of these steps can be performed in a number of seconds, and do not require expensive optimization processes. Figure 1 depicts this two-stage generation process.

We base our generative stack on diffusion (Sohl-Dickstein et al., 2015; Song & Ermon, 2020b; Ho et al., 2020), a recently proposed generative framework which has become a popular choice for text-conditional image generation. For our text-to-image model, we use a version of GLEUE (Nichol et al., 2021) fine-tuned on 3D renderings (Section 4.2). For our image-to-3D model, we use a stack of diffusion models which generate RGB point clouds conditioned on

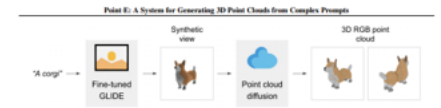


Figure 1. A high-level overview of our pipeline. First, a text prompt is fed into a GLEUE model to produce a synthetic rendered view. Next, a point cloud diffusion stack conditions on this image to produce a 3D RGB point cloud.

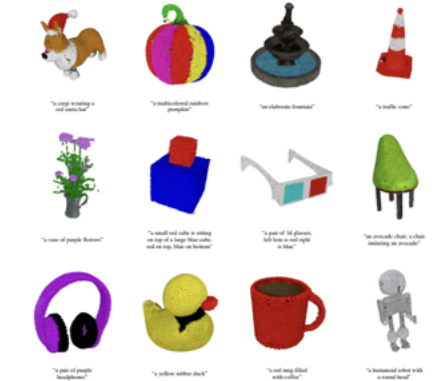


Figure 2. Selected point clouds generated by Point-E using the given text prompts. For each prompt, we selected one point cloud out of eight samples.

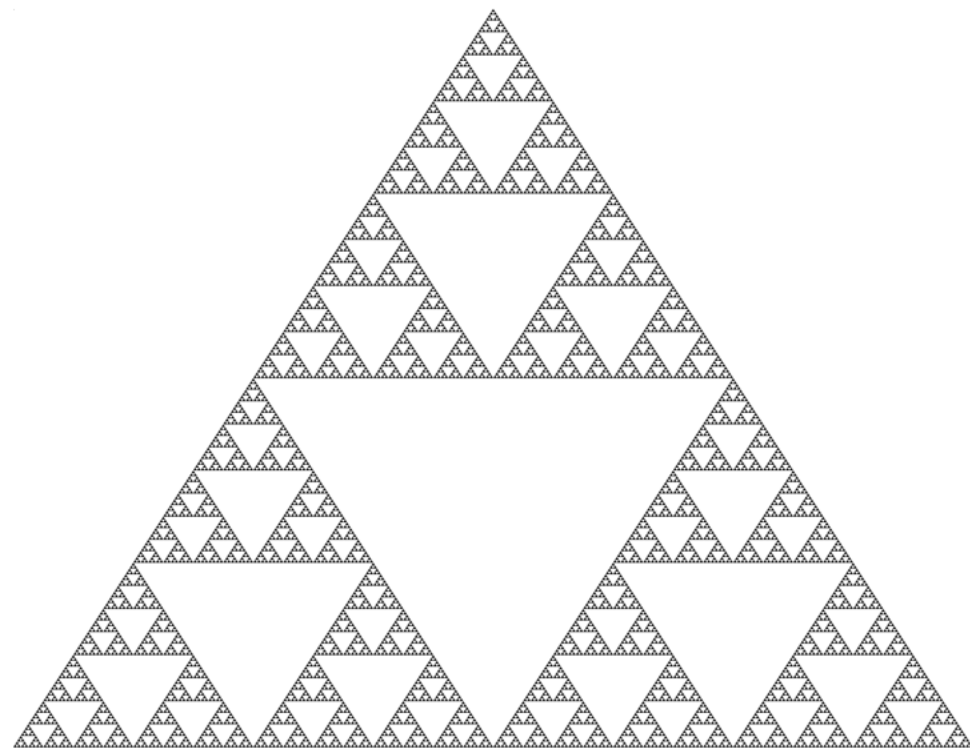
arXiv:2212.08751v1 [cs.CV] 16 Dec 2022

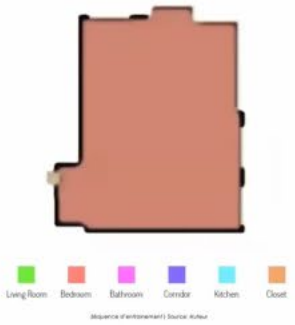
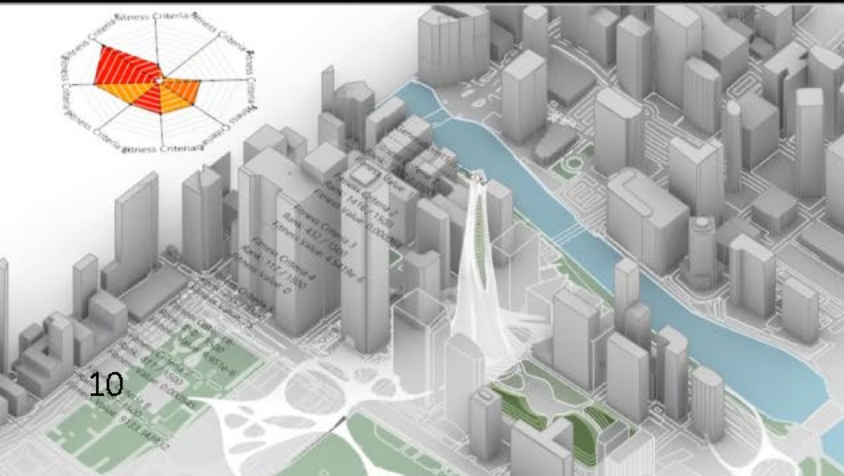
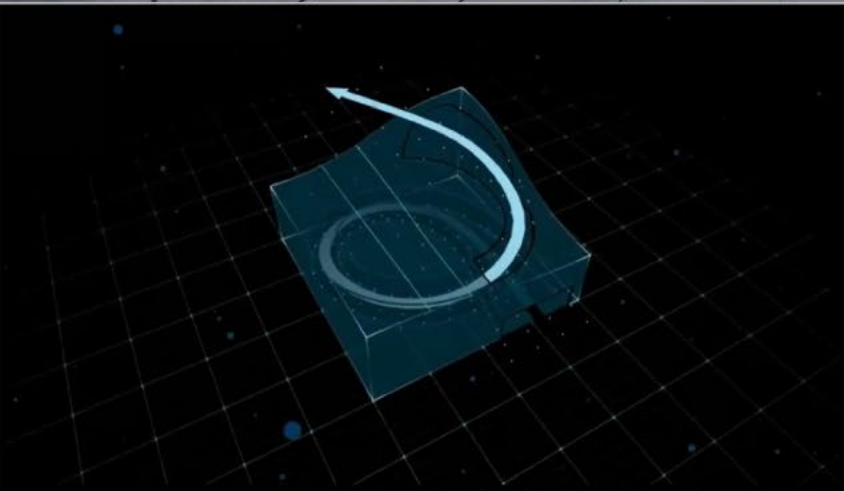
Limites de l'approche :

- Description / Interprétation ;
- Capacité du langage ;
- Disponibilité de la connaissance ;
- Éthique.

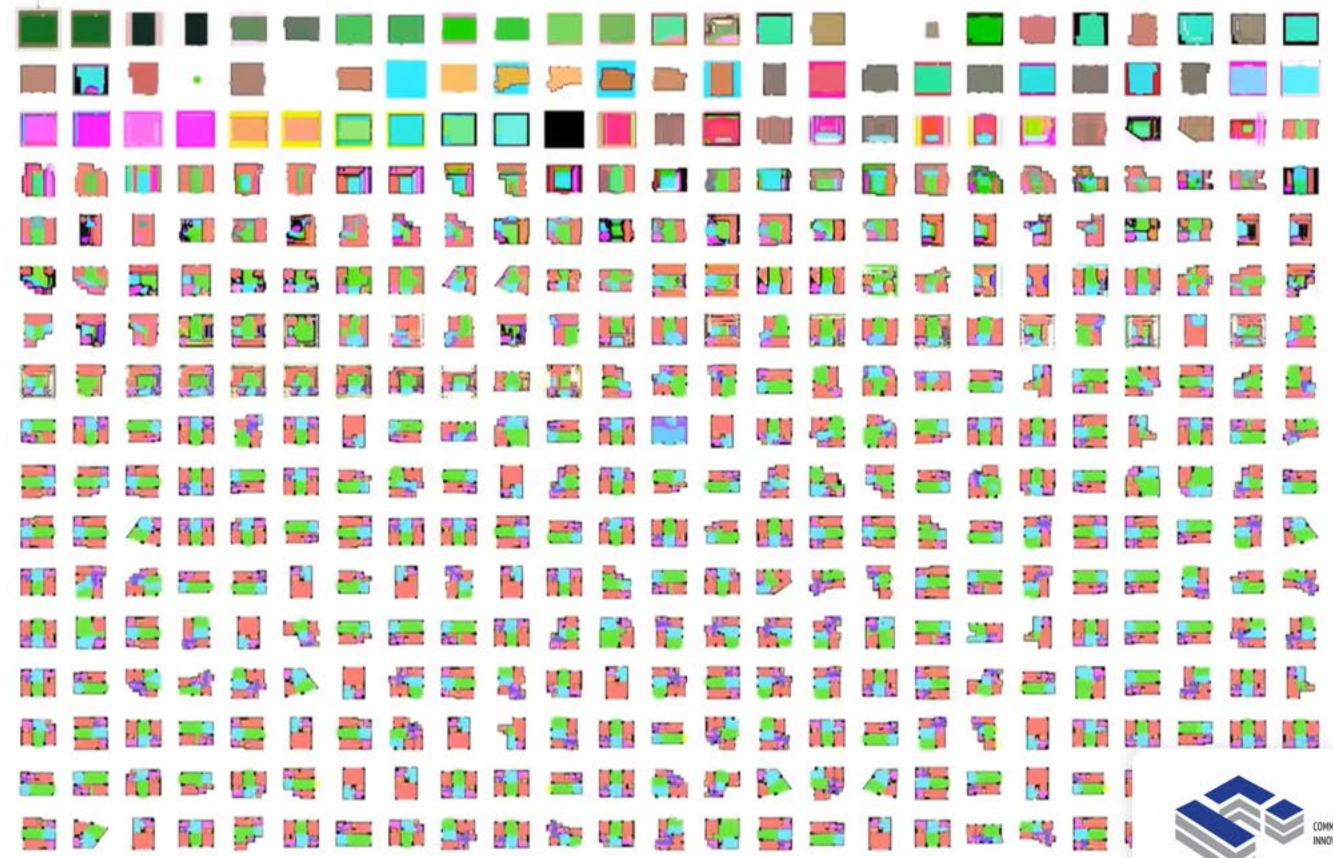
Sait-on ce qu'on veut ?

Cette approche est hautement innovante et présente un grand potentiel, mais elle doit également surmonter de nombreuses limites fondamentales.





Training Start



Le cloud, horizon de la majorité
des applications technologiques.

Pourquoi le cloud ?

- Accessible ;
- Adaptable ;
- Agile ;
- Fiable et résilient ;
- Sécuritaire pour toutes les parties.

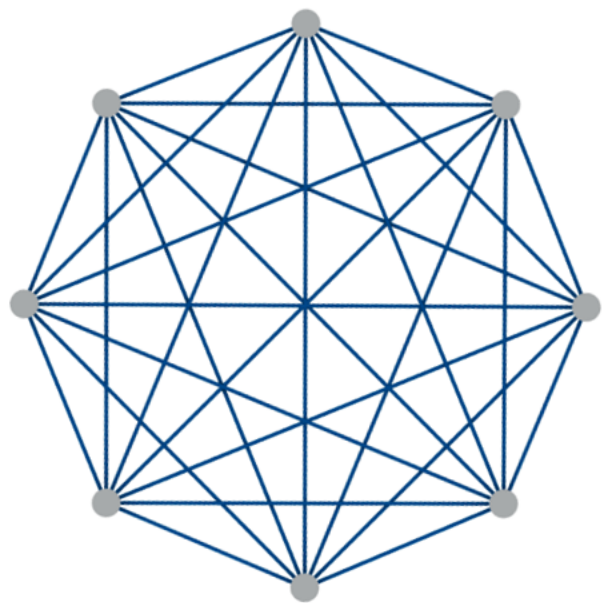


La majorité des applications de la construction
seront disponibles à travers le cloud

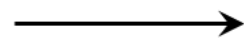
Contexte actuel

Une jungle d'outils numériques.

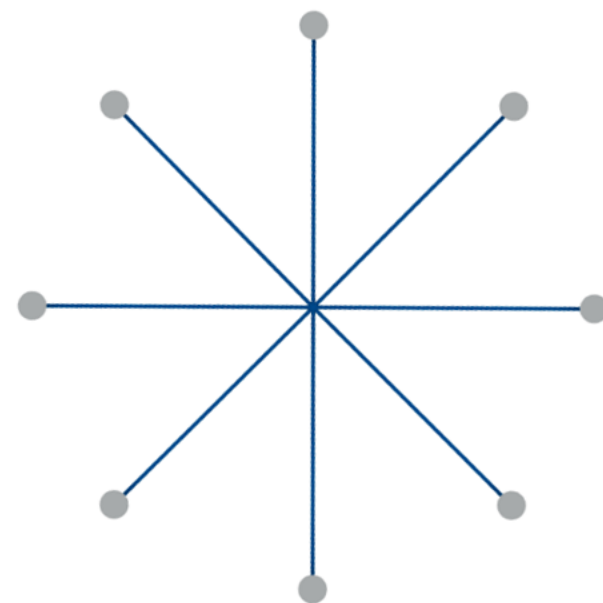
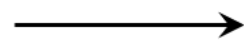




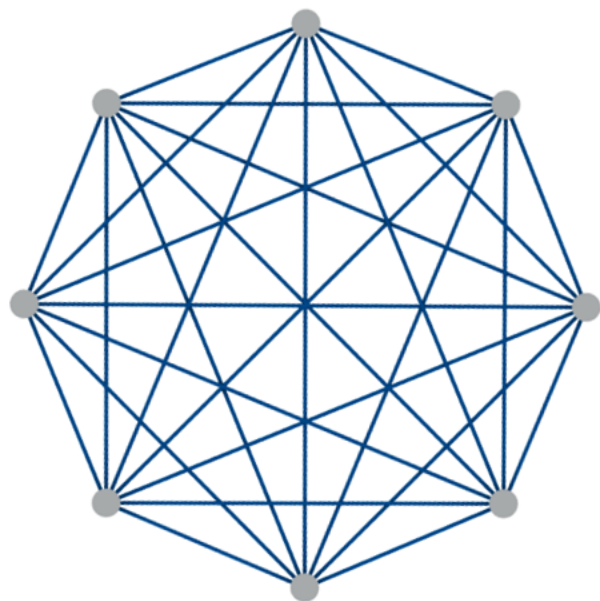
Flux de travail fragmenté



La promesse de chaque outil

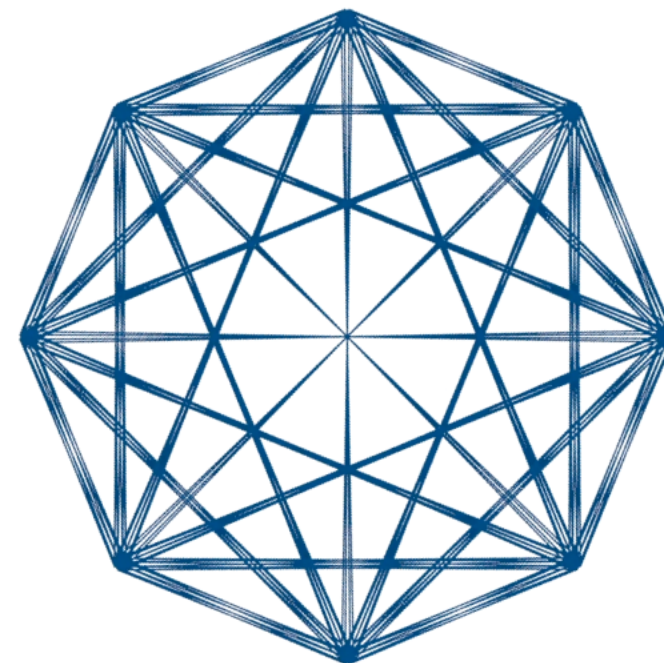


Flux de travail centralisé



Flux de travail fragmenté

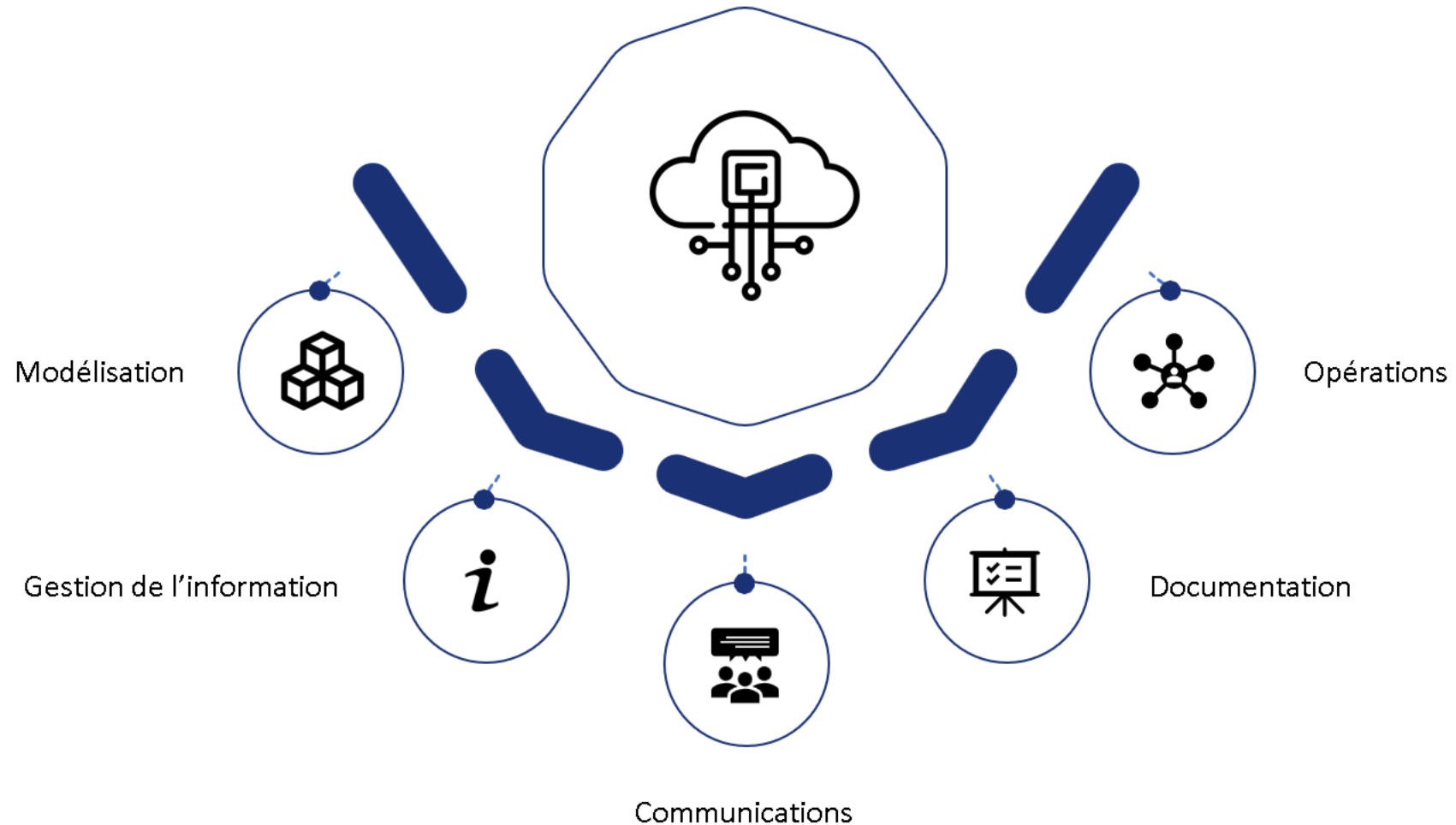
←→
Ce qui arrive au sein
des organisations
←→

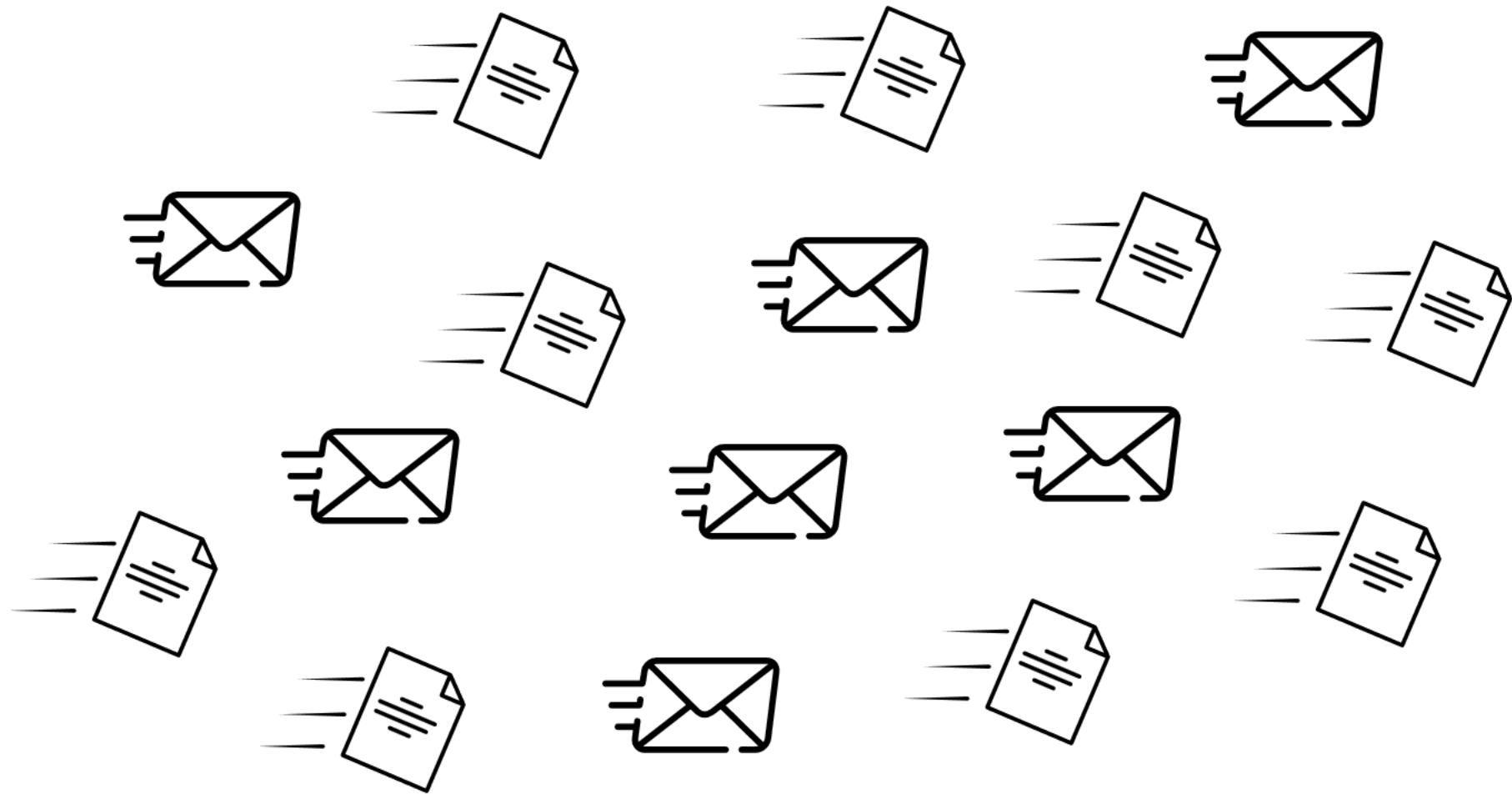


Flux de travail fragmenté

Environnement commun de données (CDE)

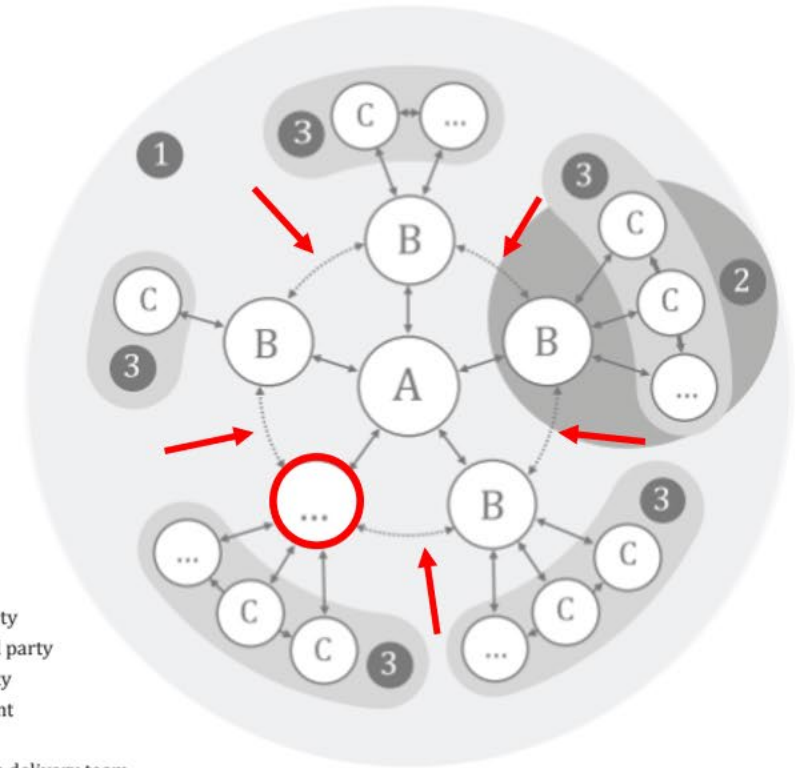
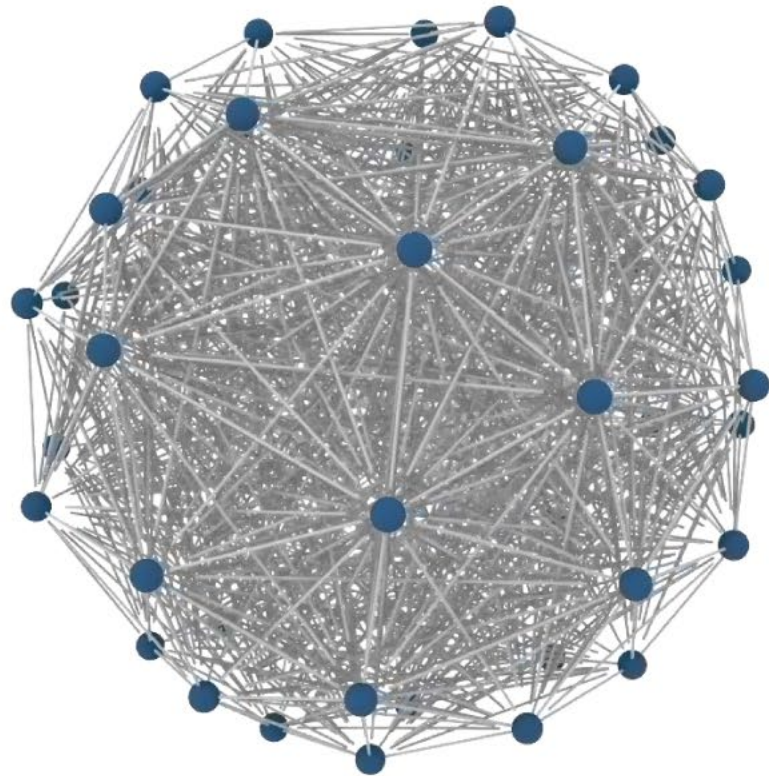
Une nécessité pour l'industrie de la construction.





ISO19650

La norme pour la collaboration et l'échange d'informations.



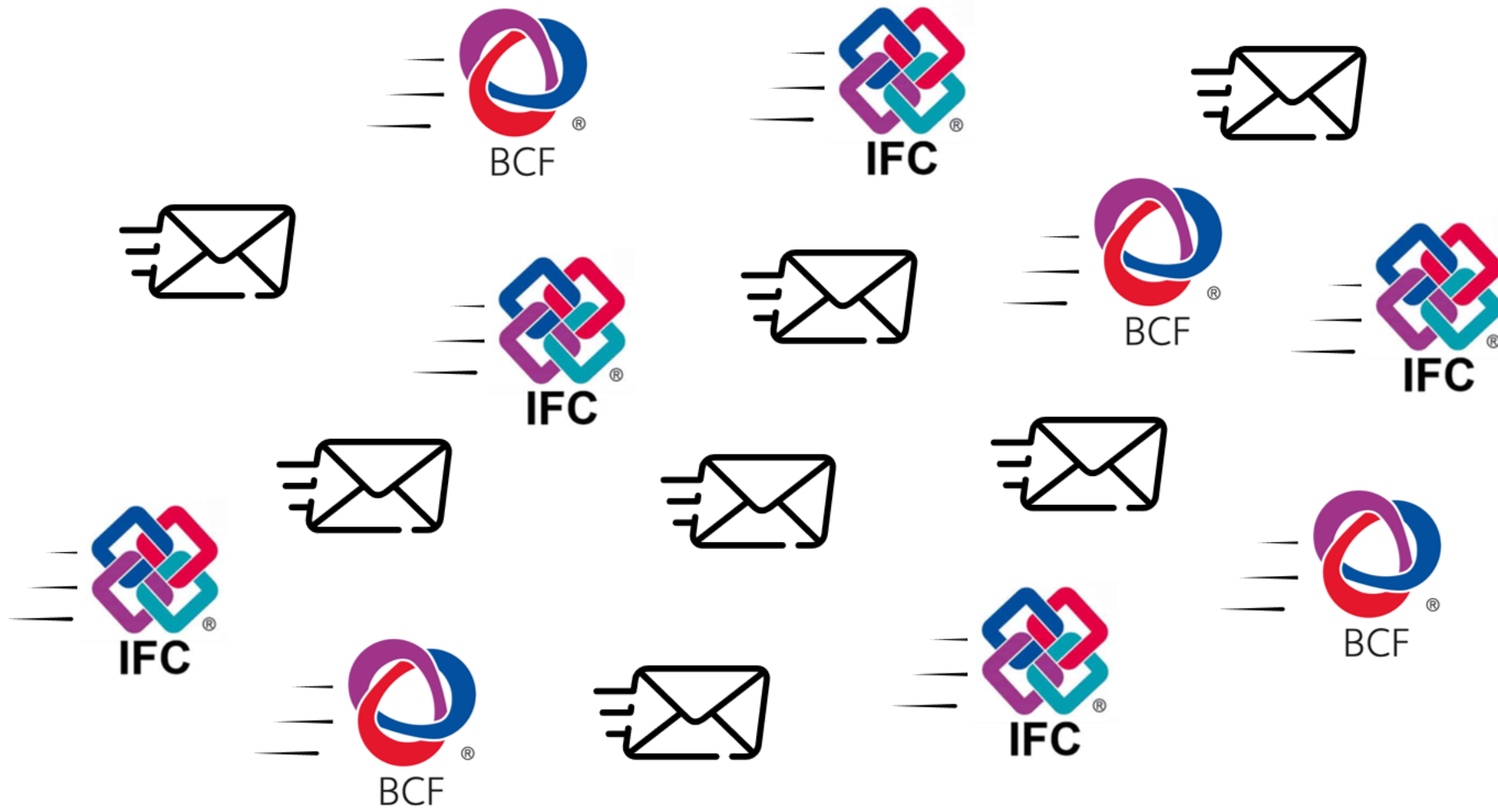
Key

- A appointing party
- B lead appointed party
- C appointed party
- ... variable amount
- 1 project team
- 2 illustration of a delivery team
- 3 task team(s)

↔ information requirements and information exchange

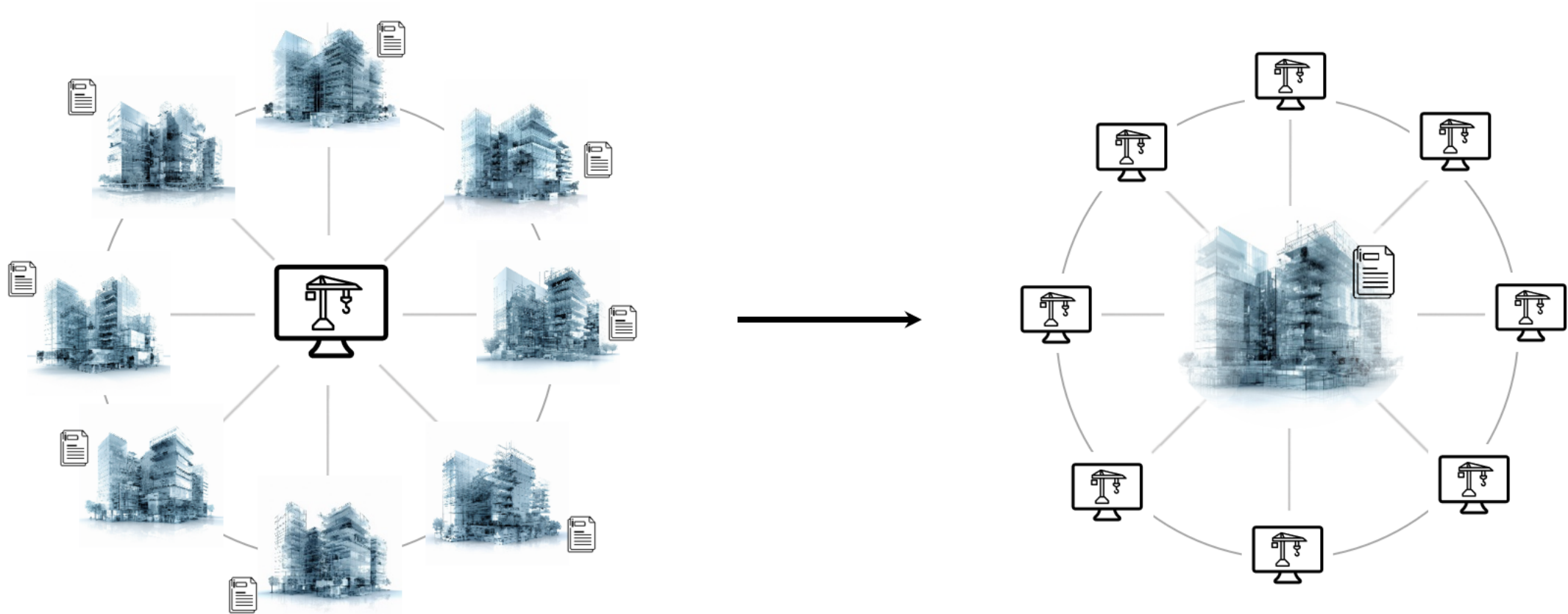
↔ information coordination

Figure 2 — Interfaces between parties and teams for the purpose of information management



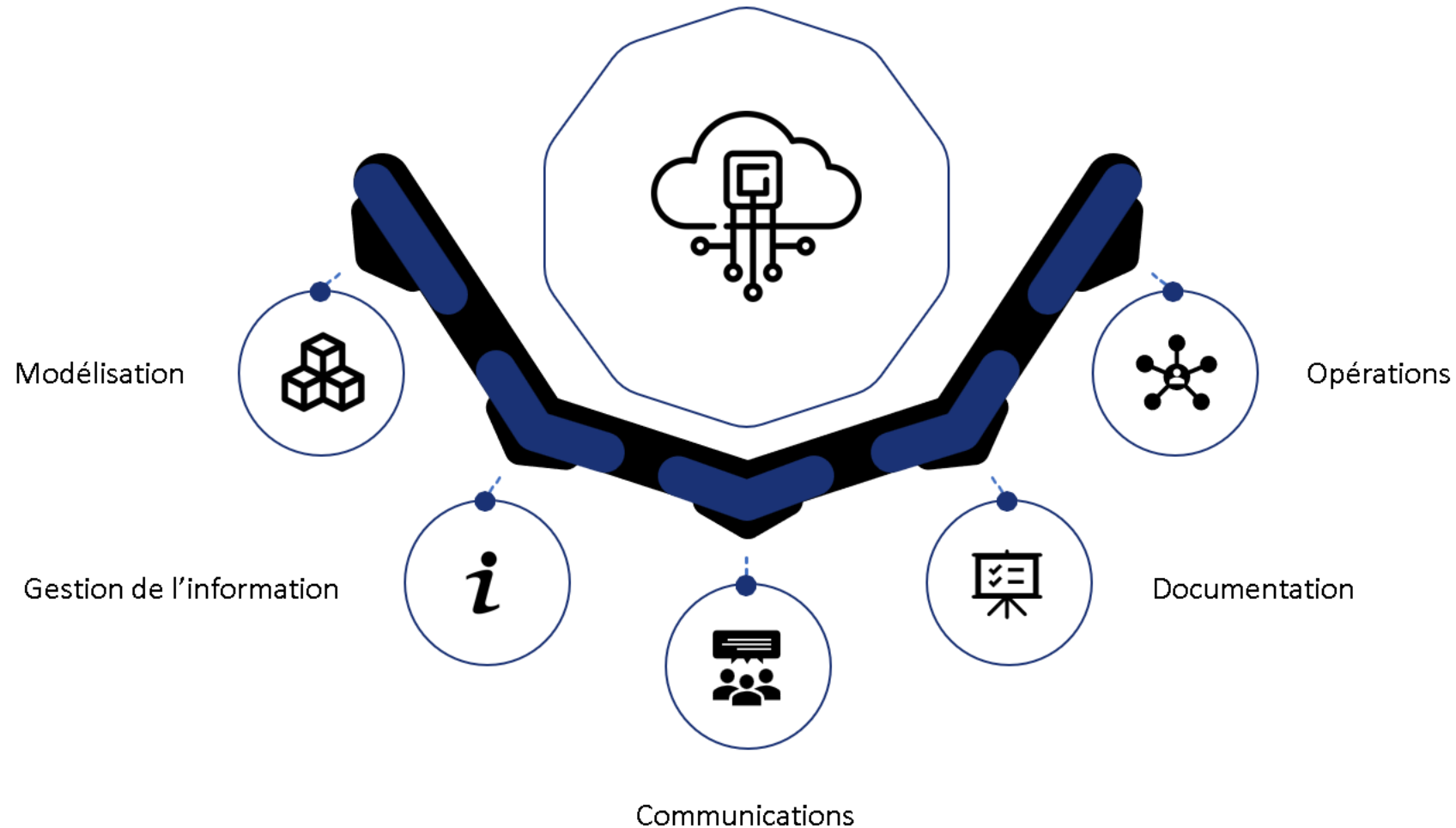
Processus de travail

Basé sur les fichiers vs. Basé sur le modèle.



Environnement commun de données (CDE)

Un flux intégré est une nécessité pour l'industrie de la construction.





Capacité des plateformes cloud

Un enjeu du **présent** et du **futur**.

Le deuxième virage numérique

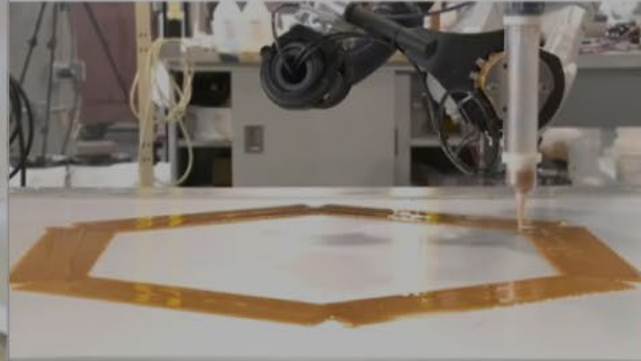
La numérisation des opérations.



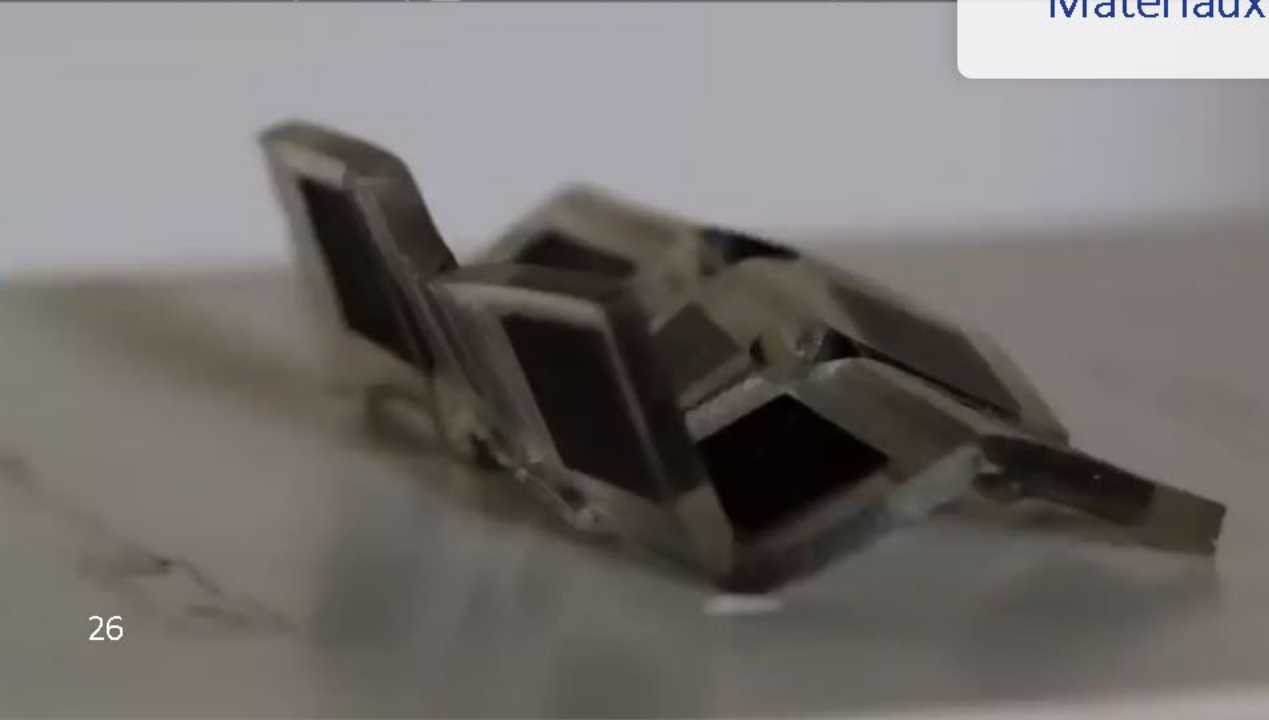
Hors site & Sur site

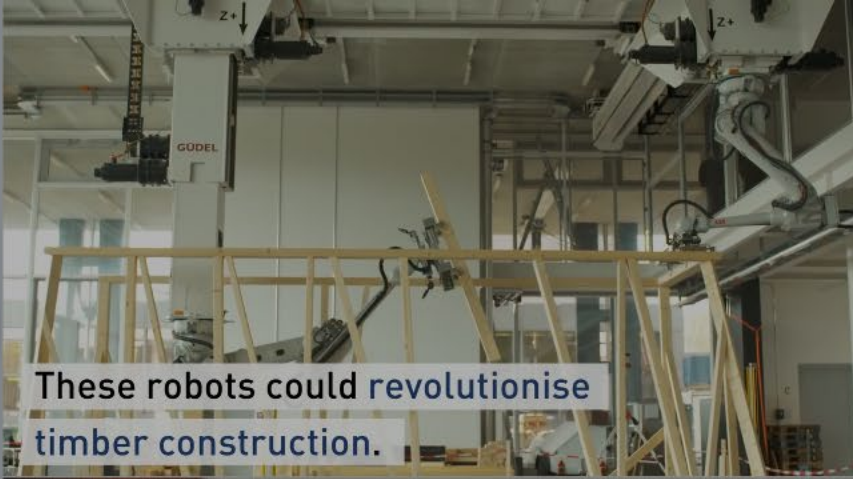
Un boom technologique en perspective.



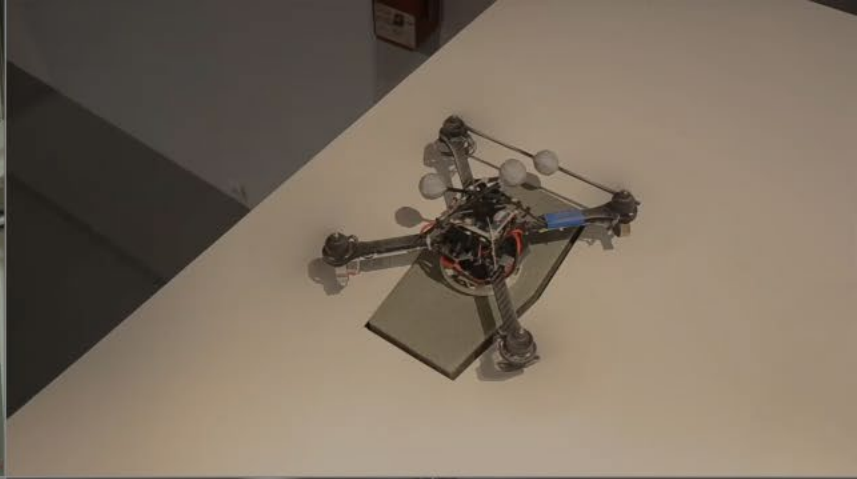


Matériaux Intelligents





These robots could revolutionise timber construction.



Fabrication Robotique



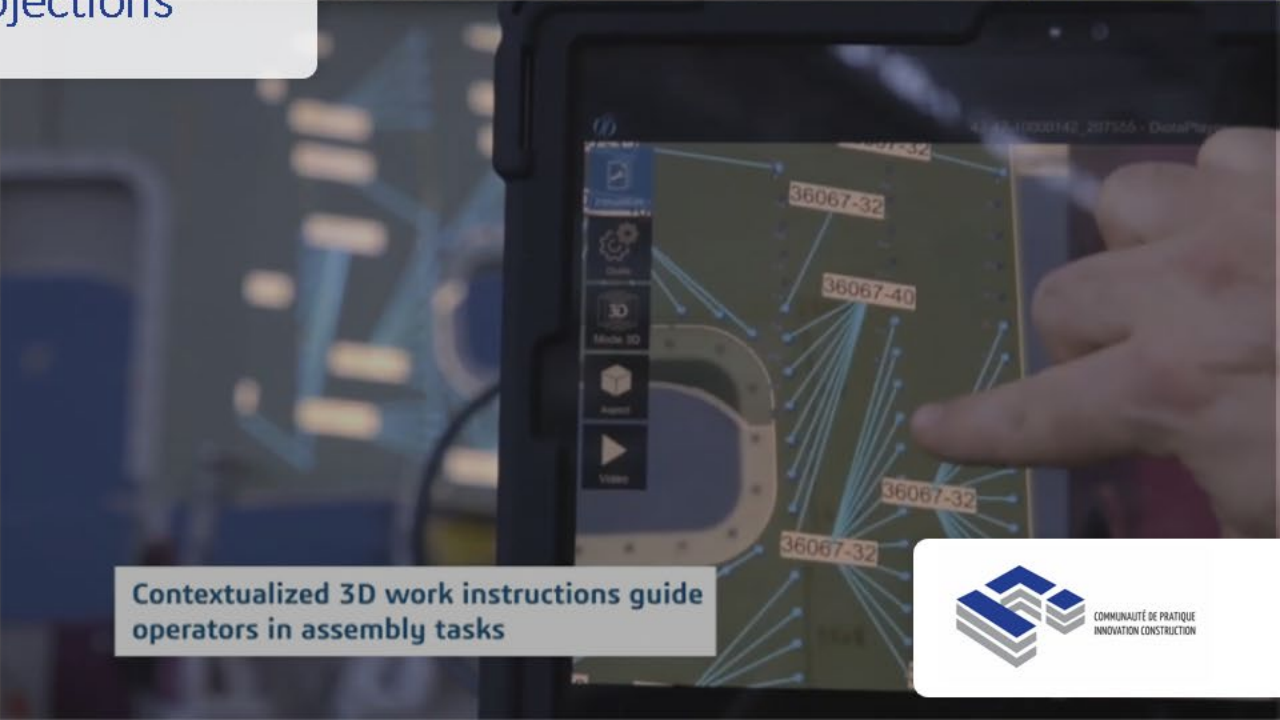


Scan-to-BIM



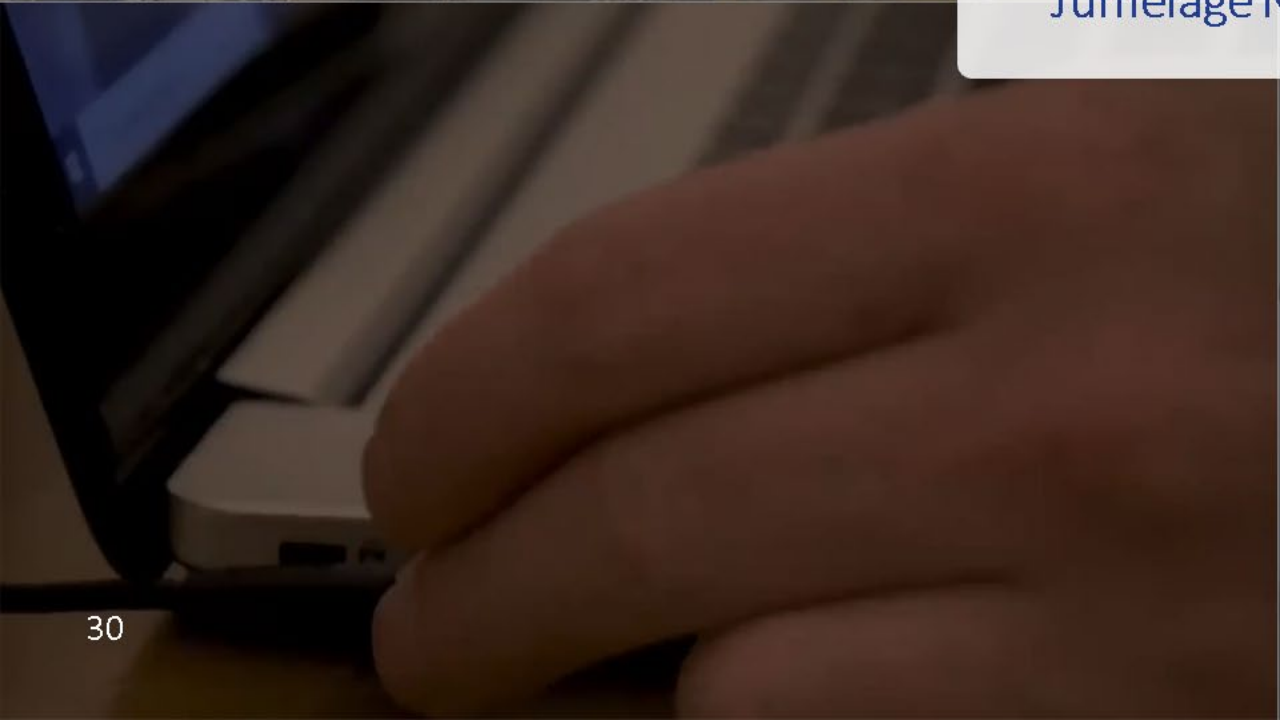


XR & Projections



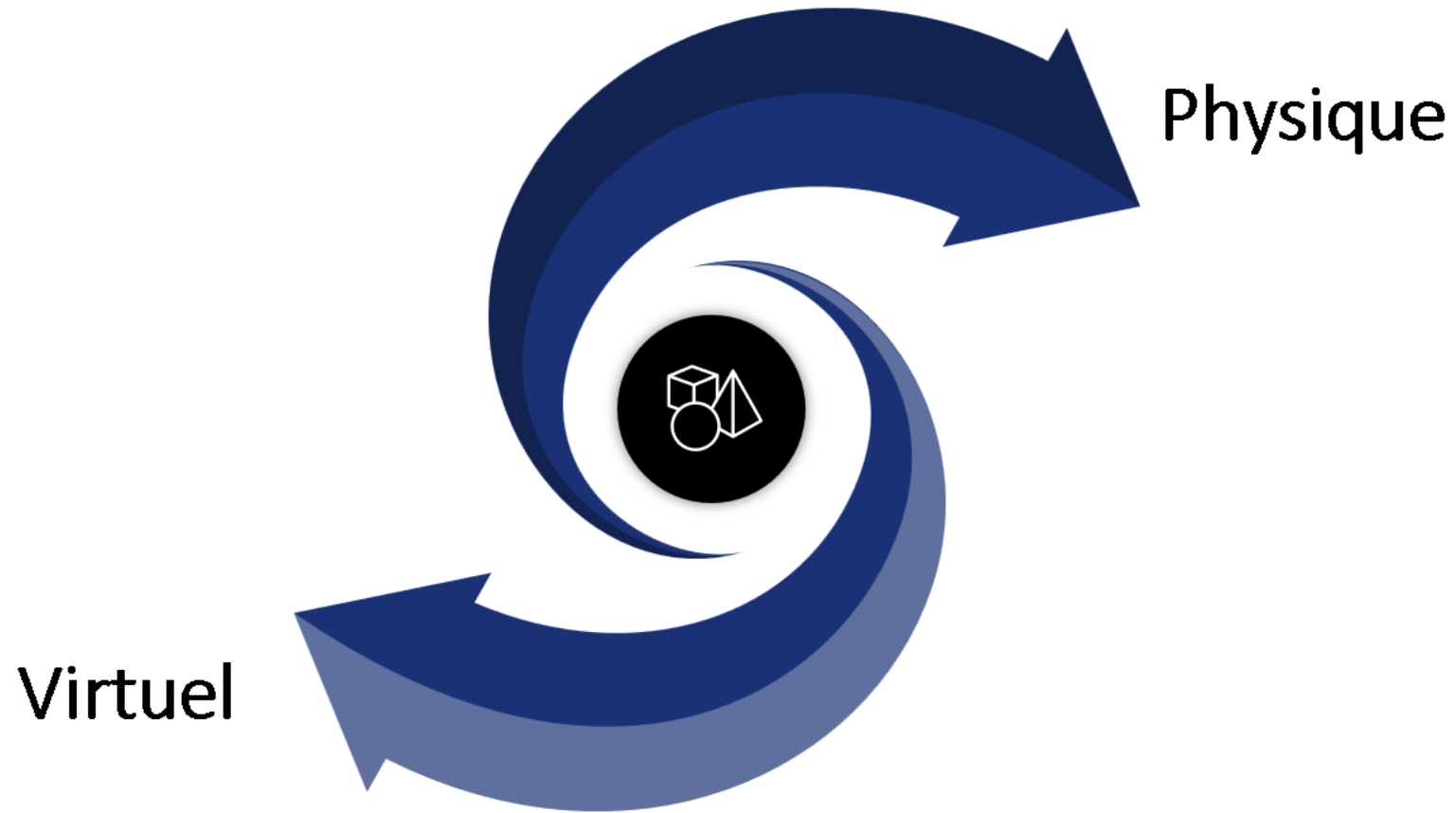


Jumelage Numérique



Horizon pour le deuxième virage numérique

La continuité numérique.



Références

Conception générative : <https://www.youtube.com/watch?v=9A8Q2leo2Jc>
Runway: <https://www.youtube.com/watch?v=fTqgWkHiN0k>
Stanislas Chaillou: https://www.youtube.com/watch?v=xNW_UnSlrqk&t=1172s
Point-E: <https://arxiv.org/abs/2212.08751>
WALLACE: <https://www.youtube.com/watch?v=3-GinXmKOI8>
CATIA: <https://www.youtube.com/watch?v=gAIFWYemeAk&t=665s>
Forma: <https://www.youtube.com/watch?v=JHeYmtHt3NE>
Speckle: <https://www.youtube.com/watch?v=B9humiSpHzM>
Home By Me: <https://www.youtube.com/watch?v=GlgmFDy4gG0&t=58s>
Power BI: https://www.youtube.com/watch?v=By_RmKloMhQ
buildingSMART: <https://buildingsmartcanada.ca/>
Prevu3D: https://www.youtube.com/watch?v=k6h_NGXXIb8
PlanEngage: <https://www.youtube.com/watch?v=TQ81dFIk7DQ&t=63s>
Aconex: <https://www.youtube.com/watch?v=qlyOGjw0c0Q>
Silk Pavilion: <https://www.youtube.com/watch?v=xVGTtV9M6sg&t=4s>
Intelligent Materials: <https://www.youtube.com/watch?v=aV07hCF7-AQ&t=18s>

SmartBuildingMaterials: <https://www.youtube.com/watch?v=Jcmbf6kuR1M>
Neri Oxman: <https://www.youtube.com/watch?v=tBCnRujU5Bc&t=7s>
Bugafibre pavilion: <https://www.youtube.com/watch?v=ggx0TZ3BPIY&t=87s>
House of Design: <https://www.youtube.com/watch?v=O3aI52UWoc0&t=1s>
ICON 3DPrinting: <https://www.youtube.com/watch?v=IRYjZed8ysM>
Laser cutting: <https://www.youtube.com/watch?v=3lLfXX9Xu-0>
Atlas : https://www.youtube.com/watch?v=-e1_QhJ1EhQ
Flight Assembled Architecture: <https://www.youtube.com/watch?v=5wBxXEbnf0s>
IoT projects: <https://www.youtube.com/watch?v=Td1d5iMF3EQ&t=340s>
Fologram: <https://www.youtube.com/watch?v=0bbAwABTzNQ&t=1s>
Gamma AR: <https://www.youtube.com/watch?v=DzFctc7bkCM>
SPOT: <https://www.youtube.com/watch?v=qgHeCfMa39E>
Luma AI: <https://www.youtube.com/watch?v=IZQ2Roor2bk>
Drone scanning: <https://www.youtube.com/watch?v=qdqQmjUk3IM>
Laser scanning: https://www.youtube.com/watch?v=O3B_6U6TQMcc
Diota: https://www.youtube.com/watch?v=ETwJCU_xc9o