

# Festival Online de la Data



## PROCHAINES DATES

Jeudi 18 Février de 15h à 15h30

**Intégrer du data storytelling dans votre stack décisionnelle**

**Ludovic Narayanin**, DATAROCKS

**Charly Clairmont**, Synaltic

Jeudi 04 Mars de 15h à 15h30

**De Tableau Prep à la visualisation sur Tableau**

**Aya MHADHBI**, Data Analyst, Synaltic

Jeudi 25 Février de 15h à 15h30

**Tableau + R**

**John BONTIT**, Data Analyst, Synaltic

Jeudi 11 Mars de 15h à 15h30

**Data Asso**

**Marc SALLIERES**, SYNALTIC

**Victor Coustenoble - Trifacta**  
Director of Sales Engineering EMEA  
[vcoustenoble@trifacta.com](mailto:vcoustenoble@trifacta.com)

**Charly Clairmont - Synaltic**  
CTO  
[cclairmont@synaltic.fr](mailto:cclairmont@synaltic.fr)



# Intégrer, nettoyer et transformer vos données dans Google Cloud

**Webinaire 11/02/21**

Google Cloud



# Agenda

- Introduction
- Intégrer, nettoyer et transformer vos données dans Google Cloud
- Suivi de la Qualité des Données
- Démonstration de Cloud Dataprep

# SYNALTIC EN QUELQUES MOTS

Acteur innovant, dénicheur de solution et une équipe de collaborateurs engagés



30

**30 collaborateurs formés et certifiés** contribuant aux communautés d'utilisateurs.

15

Spécialiste en Data Management depuis plus de 15 ans et plus de **250 projets** réalisés.

160

Plus de **160 clients** dont certains s'appuient sur nous avec succès depuis plus de **8 ans**.

Experts en Data Management, passionnés d'Open Source et d'Innovation !

# TRIFACTA & SYNALTIC

Plus de 5 années de collaboration



Trifacta est le pionnier de l'industrie et le leader du marché mondial des technologies de préparation des données.



Offre un software qui aide les individus et les organisations à explorer, transformer et rassembler efficacement des données variées.



Apporte la couche "ETL" et son industrialisation du point de vue des métiers. Il s'agit aussi de permettre aux DSI de mieux suivre l'exploitation des données par les métiers.

**Synaltic soutient Trifacta pour sa plateforme de préparation de données pour les métiers et les spécialistes de la donnée. Notre volonté est d'accompagner nos clients dans leur besoin d'exploitation de gros volumes de données à des fins d'analyse.**



“

The data nerds won - but the work isn't done !

Looker CPO Nick Caldwell at JOIN 2019

<https://diginomica.com/looker-cpo-data-nerds-won-if-we-dont-empower-users-data-we-lose>



# EXIGENCES NOUVELLES et FLUX DE DONNÉES

Modernisation des flux de données, Principes DataOps, Inclusivité



Expert DWH



Analyste



Data Scientist



Utilisateur  
Métier



Direction



Expert ETL,  
Développeur



Etude,  
Expérimentation

Plateforme de données

Qualité de donnée continue

# Du code et des scripts pour intégrer, nettoyer et transformer les données, oui mais...

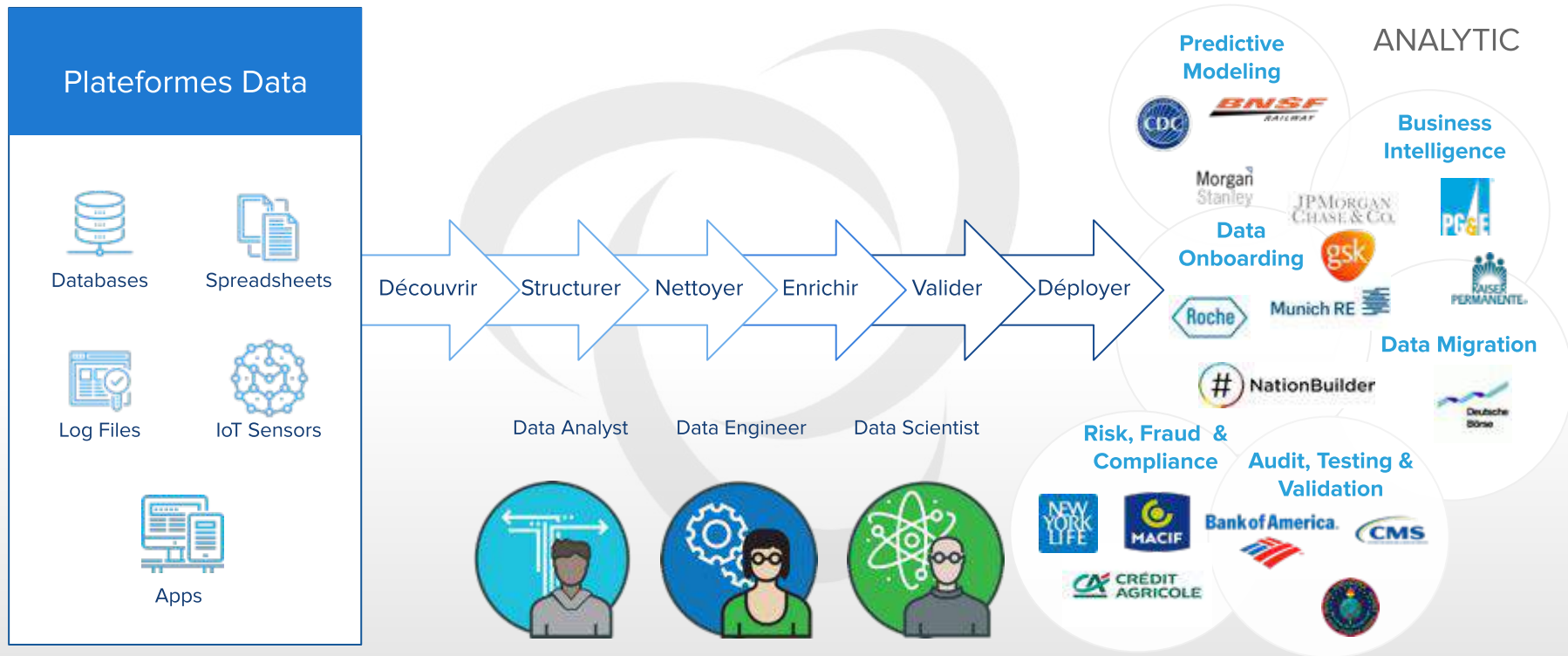
- Besoin de connaissances techniques, et de langages de développement (Python, R, SQL, Spark, Dataflow ...).
- Pas de visualisation des résultats et de la qualité des données manipulées.
- Difficile de maîtriser la performance sur de gros volumes.
- Difficultés d'industrialisation et de maintenance.
- Problèmes de gouvernance, de partage et de collaboration.

**Le nettoyage et la préparation de données peut occuper jusqu'à 80% du temps d'un Data Analyst !**





# Google Cloud Dataprep by Trifacta



# Cloud Dataprep - Bénéfices



**Simplicité du  
Serverless  
Cloud-native**



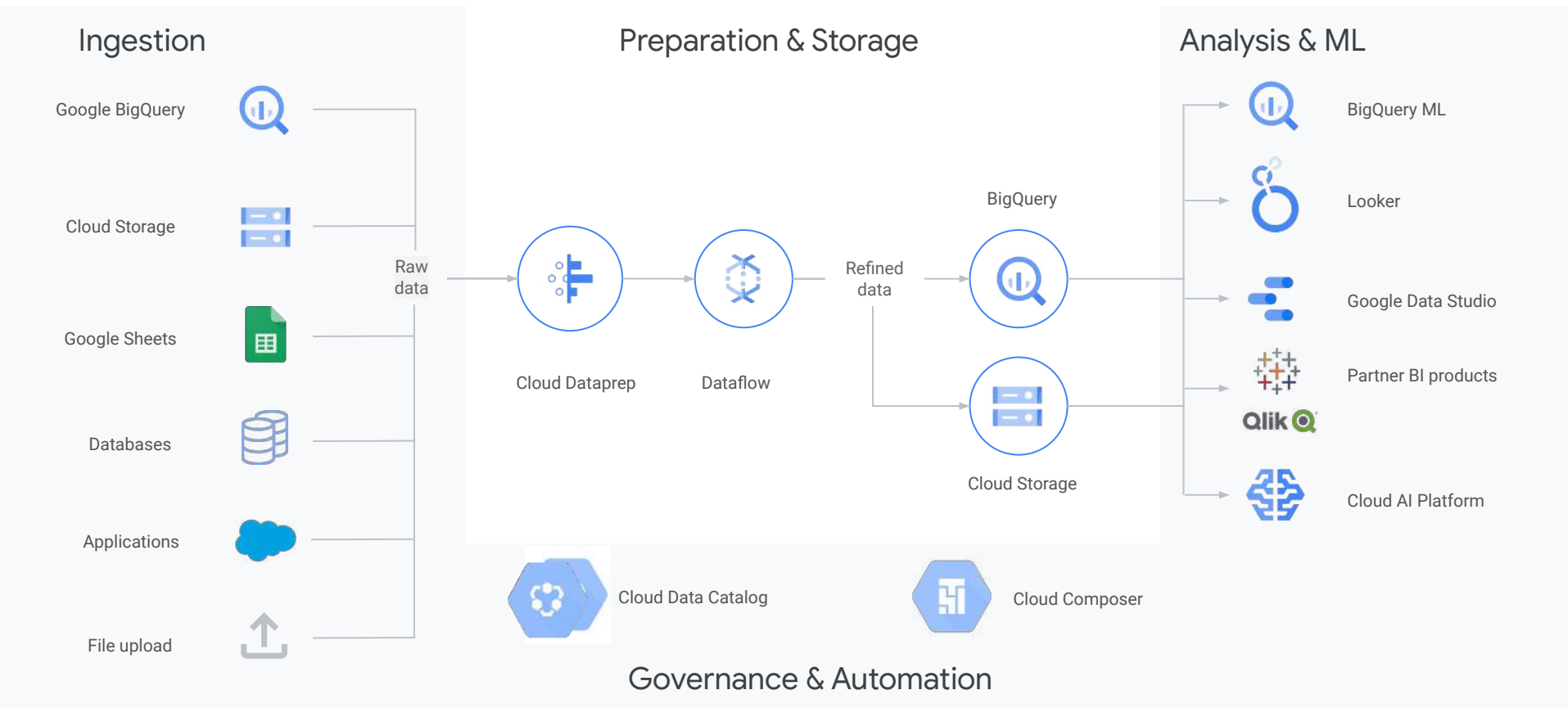
**Exploration  
rapide des  
données**



**Préparation des  
données rapide  
et simple**

# Cloud Dataprep by Trifacta

*Intégration native et disponible dans Google Cloud depuis Septembre 2017*



# Suivi de la Qualité des Données

Profiling et Qualités des Données lors des étapes d'exploration et de transformation

Monitoring de la Qualité des Données lors de l'automatisation des jobs d'intégration

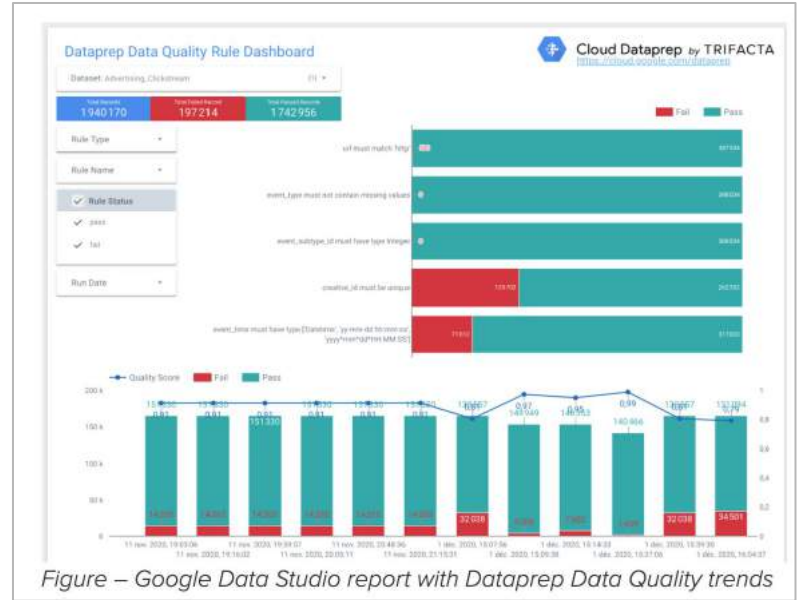
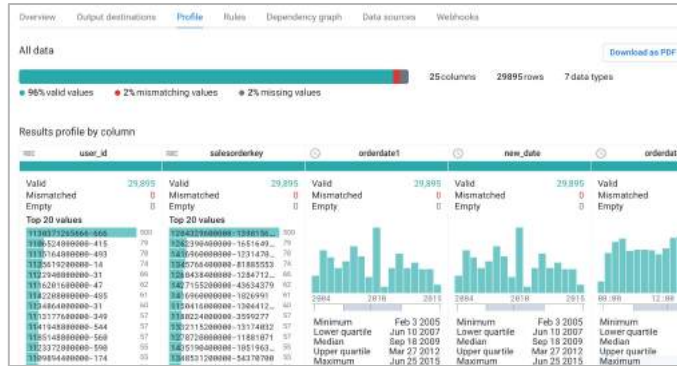
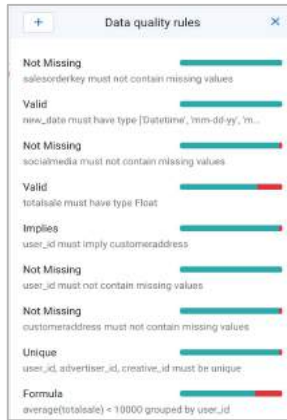
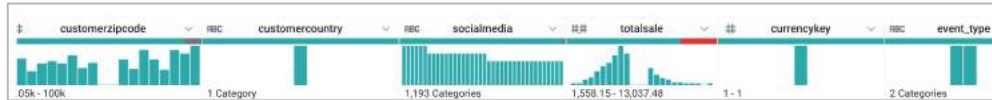


Figure – Google Data Studio report with Dataprep Data Quality trends



*"Dès le début, nous avons vu les avantages de s'appuyer sur Google Cloud Platform et un Cloud Data Warehouse, mais nous avons besoin d'une solution qui permettrait à nos analystes métiers de gérer leurs propres pipelines de données au lieu de s'appuyer uniquement sur notre équipe de data engineer. Avec Cloud Dataprep by Trifacta, nos analystes ont plus d'autonomie que jamais pour gérer et préparer leurs données, qu'il s'agisse d'exploiter les données Google Analytics stockées dans BigQuery ou de migrer de nouvelles données depuis Salesforce, ce qui entraîne d'énormes gains de productivité."*

Thibaut Gadiolet, Product Manager of Data Platforms, HomeServe

# Démonstration

Google  
Cloud Dataprep





Synaltic

Merci

Essayer Dataprep

<https://cloud.google.com/dataprep?hl=fr>

Questions ?

# Festival Online de la Data



## PROCHAINES DATES

Jeudi 18 Février de 15h à 15h30

**Intégrer du data storytelling dans votre stack décisionnelle**

**Ludovic Narayanin**, DATAROCKS

**Charly Clairmont**, Synaltic

Jeudi 04 Mars de 15h à 15h30

**De Tableau Prep à la visualisation sur Tableau**

**Aya MHADHBI**, Data Analyst, Synaltic

Jeudi 25 Février de 15h à 15h30

**Tableau + R**

**John BONTIT**, Data Analyst, Synaltic

Jeudi 11 Mars de 15h à 15h30

**Data Asso**

**Marc SALLIERES**, SYNALTIC