

02 DQSC GENERAL KPI

DQSC

Étude technique DQSC (Enterprise Data Hub) : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

Présentation DQSC (Data Quality Score Card) : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

Solution Design DQSC - V1.1 : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

Cette conception décrit l'architecture cible de DQSC :

- Backend fournissant datasets et microservices.
- API suivant les recommandations Orange et proche des standards TMF-630.
- Frontend basé sur Angular 11, déployé sur CaaS-CNP (Flexible Engine).
- Base de données NoSQL MongoDB.
- Supervision via Prometheus et Witbe.
- Livraison et déploiement via CI/CD (Orange Forge DIOD).

Éléments Complémentaires : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

Handover Checklist : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

Une longue liste de contacts (RT/DPM, ODPM, etc.) est disponible pour autoriser l'accès aux sources de données

dans le cadre de DQSC. Chaque source a des référents spécifiques, avec noms, mails, ou unités responsables.

Ces contacts couvrent des systèmes variés comme CIBASE, GOLD, FLIP, FILEADE, QUOTO, etc.

DQSC Controls (EN) : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

Liste complète des contrôles effectués dans DQSC, version 121 (créée par Sonia Labour, modifiée par Varun Srivastava).

Chaque contrôle est associé à :

- Description, dimension, niveau d'impact, source de données, fréquence.

Parmi les contrôles :

- Incohérences entre Installed Offers et clients.
- Statuts IOV incorrects, dates de facturation incohérentes.
- Circuits inactifs facturés, clients sans validation de crédit ou de conformité.
- Sites et équipements inactifs ou mal référencés.

GOLD Contacts : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

- DABRA Vivek – Application Manager (India)
- MUKHERJEE Sabby – Application Manager (India)
- RAMPAL Rahul – Customer Inventory Manager (India)
- RISHI Saurav (ext) – Software Engineer (India)

Analyse des Modèles de Données (Images) : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

La migration FINV vers DATAPLACE suit une stratégie par phase (image DQSC-2382). Le mapping des champs est géré avec précision, comme le montre le fichier DQSC-2196.

Chaque champ issu de FINV (ex: InstalledOffers, BillingAccount, CustomerName, etc.) est aligné avec son équivalent dans DataPlace.

Des fichiers de mapping Excel sont utilisés pour assurer la traçabilité.

Les extractions sont planifiées selon une périodicité de type CRON.

Exemple : `8 3 * * *` signifie une injection quotidienne à 03h08.

Un ticket EDHO-11314 a été créé pour centraliser ces informations.

Contacts d'accès DQSC : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

- Jira DQSC board : imane.boulaid@orange.com
- DQSC UAT : abdellatif.hassnaoui.ext@orange.com
- DQSC Hive : imane.boulaid@orange.com

1. Se connecter à : https://samoutils.itn.ftgroup/cgi/gassi/extract_gassi.php
2. Choisir Guardian/Gate4, puis cliquer sur 'Next'
3. Sélectionner l'application DQSC avec la dernière version d'accès

Incidents & Actions Correctives : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

- EDHO-6230 : Ajout de la source Marine à DQSC (03/02/23) → In progress : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

Support Process – DQSC Application : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

Support Process – DQSC Application

Objective:

Define clear responsibilities and actions for each support level (L1, L2, L3) in case of incidents.

Support Levels Overview:

Level	Team Responsible	Role
L1	DL IT Run ASM (itrn.asm@orange.com)	EMAD George OBS/CTIO – Incident detection, ticket creation (flapping) and routing
L2	DL MAC MEMS (mac.mems@orange.com)	FAROUQ Ahmed OBS/CTIO – Technical intervention (based on the shared test scenarios file)
L3	BU AS DATA QUALITY SCORE CARD (DQSC Team)	Deep analysis, application-level issues, escalation if needed

During working hours:

Example Process Flow – Flapping Incident (>30 min / more than 3 alerts in 30 minutes):

Step 1: Detection (Prometheus / Grafana / Monitoring tools)

- **Owner:** L1 – DL IT Run ASM

- **Actions:**

- Monitor and receive flapping alerts via Prometheus
- Create an incident ticket with details (timestamp, scope, logs)
- Route to L2 (MAC team) with appropriate priority (P1)

Critical Components:

- telegraf_exporter – Medium

- mongodb_exporter – Yes (Critical)
- MongoDB – Yes (Critical)
- Kpi Register – Medium
- Data Ingestor – Medium
- API_endpoint – Yes (Critical)

Step 2: Technical Response

- **Owner:** L2 – MAC Team
- **Actions:**
 - Confirm root cause (flapping, load issue, pod crash...)
 - Follow steps from the shared “Test Scenarios” file
 - Restart pods or take necessary action
 - Document actions in the ticket
 - Route to L3 if further analysis is needed

Step 3: Application-Level Analysis

- **Owner:** L3 – DQSC Team
- **Actions:**
 - Analyze application behavior and logs
 - Escalate to dev/vendor if needed
 - Close ticket with detailed Root Cause Analysis

Outside working hours:

- **Process Flow:** Flapping Incident (>1 hour / 3 or 4 repetitions within an hour)
- Open ticket
- Route to L2

Agreement Point:

L1 and L2 support teams are jointly responsible for ensuring the 24/7 availability of the DQSC application, including during weekends and public holidays.

DQSC KPI

Le projet DQSC vise à exploiter les données disponibles dans l'EDH afin de calculer des KPI et de les exposer via des APIs et tableaux de bord.

Objectif : offrir une vision globale de la qualité des données circulant dans le SI OBS.

L'étude aborde :

- Les sources de données, le périmètre et les contraintes GDPR.
- Les engagements en matière de sécurité et conformité (Charte EDH).
- L'architecture technique envisagée pour l'industrialisation.

DQSC est un outil de suivi de la qualité des données, permettant de :

- Identifier les incohérences et erreurs dans les données injectées dans l'EDH (plus de 180 sources).
- Calculer des indicateurs clés (KPI) et scores globaux.
- Fournir des rapports d'écarts téléchargeables pour faciliter les analyses.

Les contrôles portent sur plusieurs dimensions de qualité : complétude, cohérence, exactitude, unicité, disponibilité, validité.

Les résultats sont exposés via un tableau de bord Angular, des APIs et des rapports JSON.

DQSC as a Service – Génération de KPI & Accès : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

Composants Principaux : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

DQSC en tant que service se compose de quatre composants :

1. KPI Generic Template
2. Spark Library
3. Generator
4. Interface utilisateur (Dashboard)

1. KPI Generic Template : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

Projet GitLab : <https://gitlab.tech.orange/dqsc/dqsc-dev/generic-data-processing/dqsc-kpi-template>

Ce projet contient les fichiers JSON de configuration :

- discrepancyConditions.json : règles de détection d'anomalies
- transformations.json : colonnes du rapport, filtres
- inspectionWorkflow.json : règles métier et traitements
- Projections : fichiers décrivant les champs sources à utiliser

Le dossier lib contient le .jar de la Spark Library utilisé comme dépendance.
Contient également les contrats d'enrichissement/exposition à remplir.

2. Spark Library : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

Projet GitLab : <https://gitlab.tech.orange/dqsc/dqsc-dev/generic-data-processing/dqsc-spark-library>

- Point d'entrée : méthode `process()` dans `App.scala`
- Permet d'exécuter les règles décrites dans `inspectionWorkflow.json`
- Composants disponibles :
 - `FilterComponent`, `JoinComponent`, `DropComponent`, `ReplaceComponent`, etc.

3. Generator : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

Projet GitLab : <https://gitlab.tech.orange/dqsc/dqsc-dev/generic-data-processing/dqsc-generator>

Ce projet :

- Reçoit un JSON depuis le dashboard
- Clone le template KPI
- Remplit les fichiers de config, les contrats, `pom.xml`, `gitlab-ci`
- Sauvegarde la config dans MongoDB (collection `kpi-generation-config`)

Services importants :

- `HdfsClient` : lecture WebHDFS (fréquences, tables, champs...)
- `GitlabClient` : push, clone, suppression de projets GitLab
- `ExcelContractBuilder` : remplissage des contrats Excel
- `ConfigFilesBuilder` : remplit les fichiers config JSON + `pom.xml`

4. UI – Dashboard : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

- Permet de générer un fichier JSON avec la configuration complète du KPI
- Ce JSON est utilisé par :
 - Le Generator (stocké dans `kpi-generation-config`)
 - L'API pour les métadonnées du rapport (`report-info`)

Accès Utilisateurs, Suivi Incidents & Statuts de Déploiement des KPI : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

- Interceptor souvent indisponible → Solution : dupliquer l'instance en PROD
- Connexion MongoDB → Alerte Kubernetes configurée (19/12/2022)
- Données KPI non chargées dans MongoDB (CORE FR control) → Alerte K8S configurée
- Données non rafraîchies → Replanification via EDH météo + ajout datasets manquants dans alertes
- Catégories manquantes → Replanification Interceptor
- Données manquantes bloquant test KPI → Utilisation données PROD en STAGING (DQSC-737)
- Cache partagé REDIS en Non Prod → DQSC-772 (21/12/2022)
- Accès bloqué / Date refresh indéfinie sur UAT

Statut de Déploiement des KPI : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.

Les KPI sont répartis par vue et sous-vue, avec déploiement sur les environnements UAT et PROD : : Ce point fait référence à un aspect important du projet, nécessitant une attention particulière pour assurer la bonne mise en œuvre ou suivi de cette fonctionnalité ou activité décrite.