

Automatisation d'une chaîne de traitement

pour un accès rapide aux données du SINP



Introduction - Rôle de l'UMS

- Bancarisation - Valorisation des données du SINP
 - Définir des standards d'échange



Introduction - Rôle de l'UMS

- Bancarisation - Valorisation des données du SINP
 - Définir des standards d'échange
 - Définir des règles de contrôle et de validation des données transmises



Introduction - Rôle de l'UMS

- Bancarisation - Valorisation des données du SINP
 - Définir des standards d'échange
 - Définir des règles de contrôle et de validation des données transmises
 - Diffuser les données selon les règles définies par le SINP



Introduction - Rôle de l'UMS

- Bancarisation - Valorisation des données du SINP
 - Définir des standards d'échange
 - Définir des règles de contrôle et de validation des données transmises
 - Diffuser les données selon les règles définies par le SINP
 - Assurer la communication avec le fournisseur de données à toutes les étapes de la chaîne



Introduction - l'automatisation : pourquoi?

- Enjeux au sein de l'équipe
 - Uniformiser les méthodes



Introduction - l'automatisation : pourquoi?

- Enjeux au sein de l'équipe
 - Uniformiser les méthodes
 - Fluidifier les traitements



Introduction - l'automatisation : pourquoi?

- Enjeux au sein de l'équipe
 - Uniformiser les méthodes
 - Fluidifier les traitements
 - Simplifier la maintenance



Introduction - l'automatisation : pourquoi?

- Enjeux au sein de l'équipe
 - Uniformiser les méthodes
 - Fluidifier les traitements
 - Simplifier la maintenance
- Besoins
 - Une chaîne de traitement accessible



Introduction - l'automatisation : pourquoi?

- Enjeux au sein de l'équipe
 - Uniformiser les méthodes
 - Fluidifier les traitements
 - Simplifier la maintenance
- Besoins
 - Une chaîne de traitement accessible
 - Des traitements exécutables en une seule fois ou indépendamment



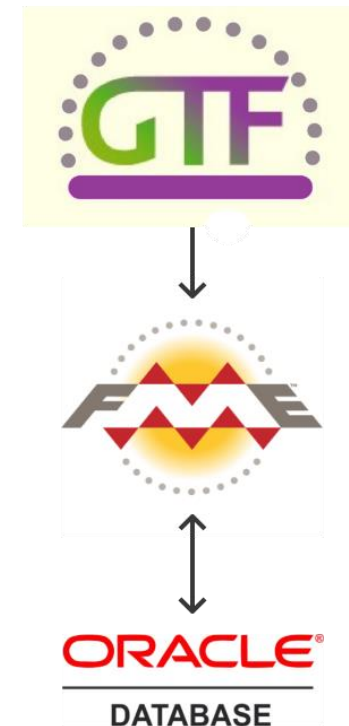
Introduction - l'automatisation : pourquoi?

- Enjeux au sein de l'équipe
 - Uniformiser les méthodes
 - Fluidifier les traitements
 - Simplifier la maintenance
- Besoins
 - Une chaîne de traitement accessible
 - Des traitements exécutables en une seule fois ou indépendamment
 - Une prise en main et des traitements rapides



Introduction - l'automatisation : comment?

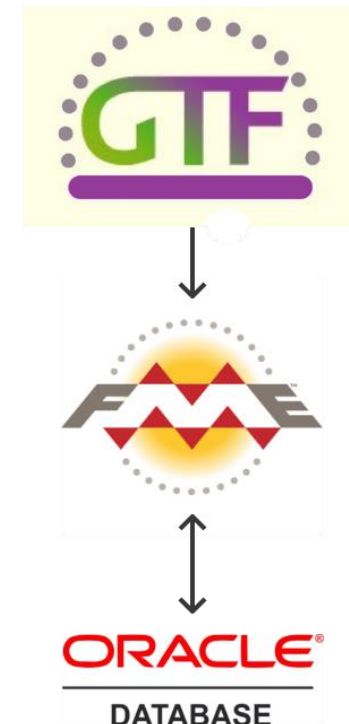
- Blocs indépendants
 - Maintenance simplifiée





Introduction - l'automatisation : comment?

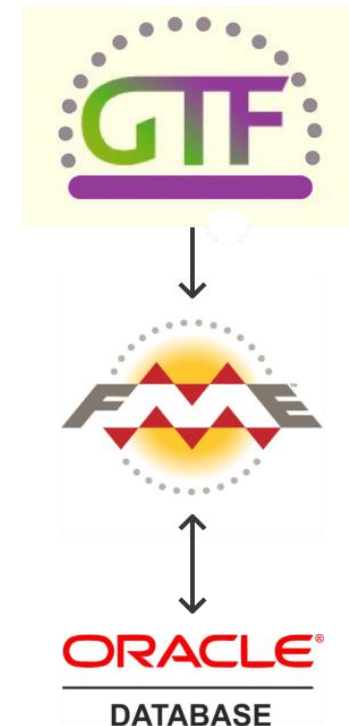
- Blocs indépendants
 - Maintenance simplifiée
 - Possibilité d'exécution indépendante





Introduction - l'automatisation : comment?

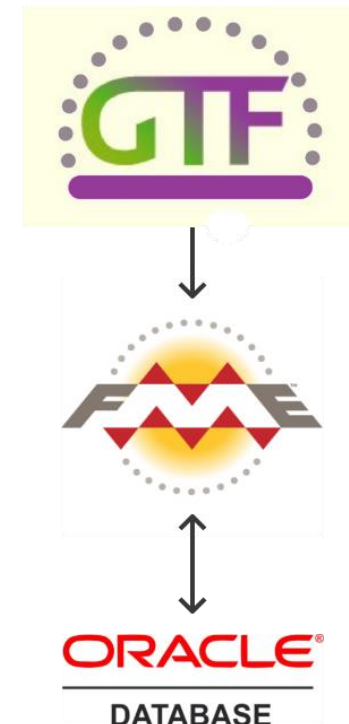
- Blocs indépendants
 - Maintenance simplifiée
 - Possibilité d'exécution indépendante
- Hiérarchie
 - Un ordre d'exécution bien défini





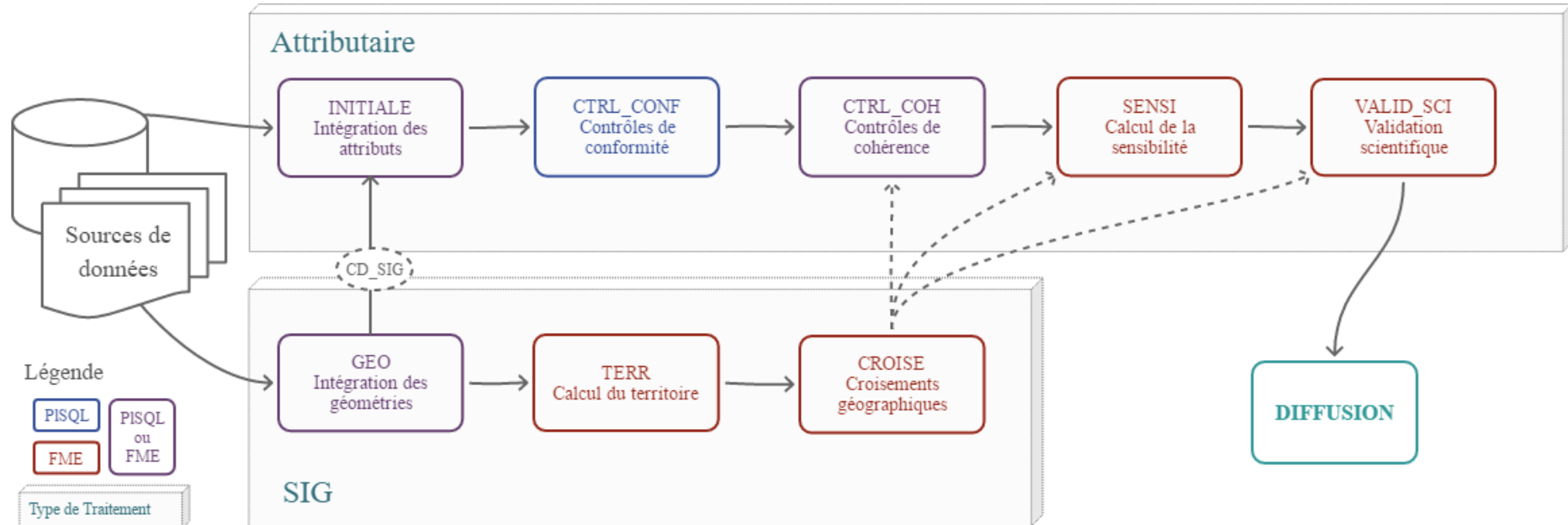
Introduction - l'automatisation : comment?

- Blocs indépendants
 - Maintenance simplifiée
 - Possibilité d'exécution indépendante
- Hiérarchie
 - Un ordre d'exécution bien défini
- Outils



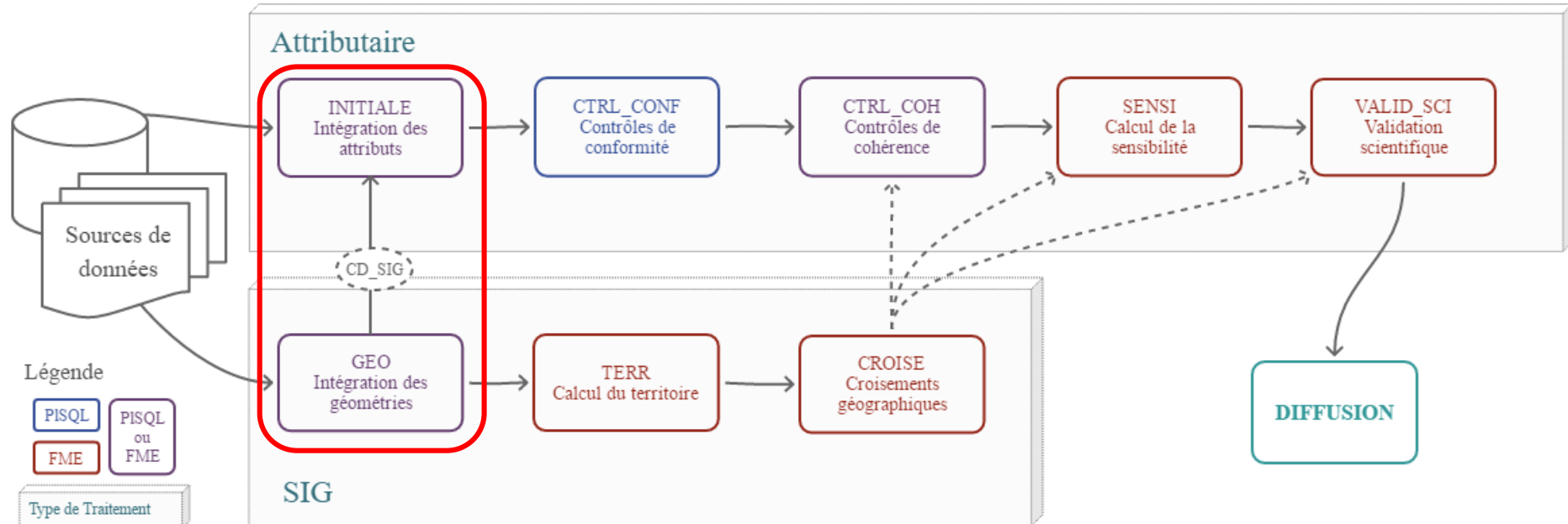


La chaîne de traitement globale





La chaîne de traitement globale



Informations complémentaires

Sujet d'observation

Information géographique

THEMA.COMMUNE	
* CODE_COMMUNE	VARCHAR2 (10 BYTE)
* NOM_COMMUNE	VARCHAR2 (256 BYTE)
* ANNEE_REF	NUMBER (4)
* TYPE_INFO_GEO	NUMBER (3)
* CD_SIG_COM	VARCHAR2 (20 BYTE)
* ID_SINP_OCCATX	VARCHAR2 (255 BYTE)
ID_JDD	NUMBER
F * ID_SUIJET_OBS	NUMBER
COMMUNE_FK1 (ID_SUIJET_OBS)	
COM_ID_SUIJET_OBS (ID_SUIJET_OBS)	
COMMUNE_UID_IDX (ID_SINP_OCCATX)	

THEMA.ESPACE NATUREL	
* CODE_EN	VARCHAR2 (200 BYTE)
* TYPE_EN	VARCHAR2 (200 BYTE)
* TYPE_INFO_GEO	NUMBER (3)
* CD_SIG_EN	VARCHAR2 (250 BYTE)
* ID_SINP_OCCATX	VARCHAR2 (255 BYTE)
* VERSION_EN	VARCHAR2 (200 BYTE)
ID_JDD	NUMBER
F * ID_SUIJET_OBS	NUMBER
ESPACE_NATUREL_FK (ID_SUIJET_OBS)	
ESPACE_NATUREL_UID_IDX (ID_SINP_OCCATX)	
ESP_ID_SUIJET_OBS (ID_SUIJET_OBS)	

THEMA.TYPE ORGANISME	
* ID	NUMBER
* LIBELLE	VARCHAR2 (50 BYTE)
STANDARD	VARCHAR2 (100 BYTE)
TYPE_ORGANISME_FK (ID)	
TYPE_ORGANISME_FK (ID)	

THEMA.TYPE PERSONNE	
* ID	NUMBER
* LIBELLE	VARCHAR2 (50 BYTE)
STANDARD	VARCHAR2 (100 BYTE)
TYPE_PERSONNE_FK (ID)	
TYPE_PERSONNE_FK (ID)	

Personne/organisme

THEMA.DEPARTEMENT	
* CODE_DEP	VARCHAR2 (10 BYTE)
* ANNEE_REF	NUMBER (4)
* TYPE_INFO_GEO	NUMBER (3)
* CD_SIG_DEP	VARCHAR2 (20 BYTE)
* ID_SINP_OCCATX	VARCHAR2 (255 BYTE)
* NOM_DEP	VARCHAR2 (255 BYTE)
ID_JDD	NUMBER
F * ID_SUIJET_OBS	NUMBER
DEPARTEMENT_FK1 (ID_SUIJET_OBS)	
DEPARTEMENT_UID_IDX (ID_SINP_OCCATX)	
DEP_ID_SUIJET_OBS (ID_SUIJET_OBS)	

THEMA.MAILLE 10X10	
* CD_MAILLE	VARCHAR2 (250 BYTE)
* TYPE_INFO_GEO	NUMBER (3)
* CD_SIG_MAILLE	VARCHAR2 (250 BYTE)
* ID_SINP_OCCATX	VARCHAR2 (255 BYTE)
SUPP_V_REF	NUMBER (5)
ID_JDD	NUMBER
F * ID_SUIJET_OBS	NUMBER
MAILLE_10X10_FK (ID_SUIJET_OBS)	
MAIL_10X10_UID_IDX (ID_SINP_OCCATX)	

THEMA.MASSE EAU	
* VERSION_ME	NUMBER (4)
* CODE_ME	VARCHAR2 (200 BYTE)
DATE_ME	DATE
* TYPE_INFO_GEO	NUMBER (3)
* CD_SIG_ME	VARCHAR2 (250 BYTE)
* ID_SINP_OCCATX	VARCHAR2 (255 BYTE)
ID_JDD	NUMBER
F * ID_SUIJET_OBS	NUMBER
MASSE_EAU_FK (ID_SUIJET_OBS)	
MAS_ID_SUIJET_OBS (ID_SUIJET_OBS)	
MASSE_EAU_UID_IDX (ID_SINP_OCCATX)	

THEMA.PERSONNE ORGANISME	
* TYPE_INFO	NUMBER
* IDENTITE	VARCHAR2 (255 BYTE)
ORGANISME	VARCHAR2 (255 BYTE)
MAIL	VARCHAR2 (255 BYTE)
PERSONNE_ORGANISME	NUMBER
F * ID_SINP_OCCATX	VARCHAR2 (255 BYTE)
ID_JDD	NUMBER
F * ID_SUIJET_OBS	NUMBER
PERSONNE_ORGANISME_FK (ID_SUIJET_OBS)	
PERSONNE_ORGANISME_FK1 (TYPE_PERSONNE)	
PERSONNE_ORGANISME_FK2 (TYPE_ORGANISME)	
PER_ID_SUIJET_OBS (ID_SUIJET_OBS)	
PERSONNE_ORGANISME_UID_IDX (ID_SINP_OCCATX)	

THEMA.ELEMENTS DEPLACES	
* ID_SINP_OCCATX	VARCHAR2 (255 BYTE)
* NOM_ATTR	VARCHAR2 (200 BYTE)
* VALEUR_INITIALE	VARCHAR2 (255 BYTE)
ID_JDD	NUMBER
F * ID_SUIJET_OBS	NUMBER
ELEMENTS_DEPLACES_FK (ID_SUIJET_OBS)	
ELE_ID_SUIJET_OBS (ID_SUIJET_OBS)	
ELEM_DEPLACES_UID_IDX (ID_SINP_OCCATX)	

THEMA.SUIJET OBSERVATION	
NOM_CITE	VARCHAR2 (250 BYTE)
DATE_DEBUT	DATE
DATE_FIN	DATE
DATE_DET	DATE
ALT_MIN	NUMBER (10,2)
ALT_MAX	NUMBER (10,2)
ALT_MOY	NUMBER (10,2)
PROF_MIN	NUMBER (10,2)
PROF_MAX	NUMBER (10,2)
PROF_MOY	NUMBER (10,2)
CD_NOM	NUMBER (8)
SUPP_CD_REF	NUMBER (8)
SUPP_V_TAXREF	VARCHAR2 (200 BYTE)
COMMENTAIRE	VARCHAR2 (4000 BYTE)
* CD_ESP	VARCHAR2 (200 BYTE)
* CD_SIG_GDIR	VARCHAR2 (255 BYTE)
NAT_OBJ_GEO	VARCHAR2 (25 BYTE)
TYPE_GEOM	VARCHAR2 (50 BYTE)
* TYPE_DATA_SOURCE	VARCHAR2 (50 BYTE)
FLOUTAGE	VARCHAR2 (5 BYTE)
* OS_PUBLIQUE	VARCHAR2 (255 BYTE)
ID_ORIGINE	VARCHAR2 (10 BYTE)
* STAT_SRC	VARCHAR2 (10 BYTE)
REF_BIBLIO	VARCHAR2 (4000 BYTE)
DENBR_MIN	NUMBER (10)
DENBR_MAX	NUMBER (10)
* CD_DOC	VARCHAR2 (20 BYTE)
* ID_SINP_OCCATX	VARCHAR2 (255 BYTE)
HEURE_DEBUT	DATE
HEURE_FIN	DATE
PRECISION_GEO	VARCHAR2 (20 BYTE)
NIVEAU_VALIDATION_PRODUCTEUR	VARCHAR2 (20 BYTE)
PAYS	VARCHAR2 (30 BYTE)
CD_INVENT	VARCHAR2 (20 BYTE)
TERRITOIRE	VARCHAR2 (20 BYTE)
STATUT_OBS	VARCHAR2 (50 BYTE)
SINP_DATE	DATE
SINP_NIV	NUMBER (3)
SINP_REF	VARCHAR2 (30 BYTE)
SINP_V_REF	VARCHAR2 (30 BYTE)
TYPE_SINP	VARCHAR2 (10 BYTE)
* ID_SINP_JDD	VARCHAR2 (1000 BYTE)
UID_SENSIBLE	VARCHAR2 (255 BYTE)
CD_NOM_OLD	NUMBER (8)
DATE_CTRL_PRODUCTEUR	VARCHAR2 (10 BYTE)
DENBR_OBUE	VARCHAR2 (50 BYTE)
SUPP_SENSIBLE	VARCHAR2 (2000 BYTE)
PROC_CTRL_PRODUCTEUR	VARCHAR2 (1000 BYTE)
DENBR_TYPE	VARCHAR2 (10 BYTE)
ID_SOURCE_JDD	NUMBER
ID_JDD	NUMBER
NOM_LIEU	VARCHAR2 (4000 BYTE)
ID_SUIJET_OBS	NUMBER
ID_GRP_OBS	NUMBER
F * ID_SINP_GRP	VARCHAR2 (255 BYTE)
SUIJET_OBSERVATION_FK (ID_SUIJET_OBS)	
SUIJET_OBSERVATION_FK1 (ID_GRP_OBS)	
SUIJET_OBSERVATION_FK2 (ID_SUIJET_OBS)	

THEMA.HABITAT TYPE	
V_REF	VARCHAR2 (200 BYTE)
CODE_HABITAT_REF	VARCHAR2 (200 BYTE)
CODE_HABITAT	VARCHAR2 (200 BYTE)
REF_HABITAT	VARCHAR2 (200 BYTE)
* ID_SINP_OCCATX	VARCHAR2 (255 BYTE)
ID_JDD	NUMBER
F * ID_SUIJET_OBS	NUMBER
HABITAT_TYPE_FK (ID_SUIJET_OBS)	
HAB_ID_SUIJET_OBS (ID_SUIJET_OBS)	
HABITAT_TYPE_UID_IDX (ID_SINP_OCCATX)	

THEMA.REGROUPEMENT OBSERVATION	
* TYPE_GRP	VARCHAR2 (255 BYTE)
* ID_SINP_GRP	VARCHAR2 (255 BYTE)
METHODE_GRP	VARCHAR2 (255 BYTE)
DATE_MODIF_GRP	DATE
ID_ORIGINE_GRP	VARCHAR2 (255 BYTE)
ECHANT_EFFORT	VARCHAR2 (255 BYTE)
ECHANT_TAILLE	NUMBER
ECHANT_TECH	VARCHAR2 (255 BYTE)
ECHANT_UNIT	VARCHAR2 (20 BYTE)
SENSI_DAT	DATE
SENSI_NIV	NUMBER (3)
SENSI_REF	VARCHAR2 (30 BYTE)
SENSI_V_REF	VARCHAR2 (30 BYTE)
* ID_GRP_OBS	NUMBER
F * ID_GRP_OBS_SUP	NUMBER
ID_JDD	NUMBER
F * ID_SINP_GRP_SUP	VARCHAR2 (255 BYTE)
REGROUPEMENT_OBSERVATION_FK (ID_GRP_OBS)	
REGROUPEMENT_OBSERVATION_FK1 (ID_GRP_OBS_SUP)	
REGROUPEMENT_OBSERVATION_FK2 (ID_GRP_OBS)	

THEMA.VALIDATION NAT	
AUTO_DATE_CTRL	DATE
AUTO_ID_PROC_REF	NUMBER
AUTO_NIVEAU_VALID	NUMBER
* ID_SINP_OCCATX	VARCHAR2 (255 BYTE)
MANUELLE_COMA_VALID	VARCHAR2 (255 BYTE)
MANUELLE_DATE_CTRL	DATE
MANUELLE_IDENTITE	VARCHAR2 (255 BYTE)
MANUELLE_ID_PROC_REF	NUMBER
MANUELLE_MAIL	VARCHAR2 (255 BYTE)
MANUELLE_NIVEAU_VALID	NUMBER
MANUELLE_ORGANISME	VARCHAR2 (255 BYTE)
MANUELLE_PER_VALID	NUMBER
PRODUC_DATE_CONTACT	DATE
PRODUC_IDENTITE	VARCHAR2 (255 BYTE)
PRODUC_MAIL	VARCHAR2 (255 BYTE)
PRODUC_ORGANISME	VARCHAR2 (255 BYTE)
ST_ABDSM	VARCHAR2 (20 BYTE)
ST_COHERENCE_DATE	VARCHAR2 (20 BYTE)
ST_COHERENCE_SINP	VARCHAR2 (20 BYTE)
ST_CONFORMITE	VARCHAR2 (20 BYTE)
ST_DOUBLON	VARCHAR2 (20 BYTE)
ST_HABITAT	VARCHAR2 (20 BYTE)
ST_STATUT	VARCHAR2 (20 BYTE)
ST_TAXO	VARCHAR2 (20 BYTE)
ST_VSD	NUMBER
* ID_JDD	NUMBER
F * ID_SUIJET_OBS	NUMBER
ST_CTRL_COMMUNE	VARCHAR2 (4000 BYTE)
ST_CTRL_DEP	NUMBER
ST_CTRL_D_MODIF	VARCHAR2 (4000 BYTE)
ST_CTRL_D_NOM_CONF	VARCHAR2 (4000 BYTE)
ST_CTRL_DONNEES	NUMBER
ST_CTRL_EN	NUMBER
ST_CTRL_MAILLE	NUMBER
ST_CTRL_MD_DEPLACE	VARCHAR2 (4000 BYTE)
ST_CTRL_MD_MODIF	VARCHAR2 (4000 BYTE)
ST_CTRL_MD_NOM_CONF	VARCHAR2 (4000 BYTE)
ST_CTRL_ME	NUMBER
ST_CTRL_METADONNEES	NUMBER
ST_CTRL_PRECI	NUMBER
ST_COHERENCE	VARCHAR2 (20 BYTE)
ST_COHERENCE_DENBR	VARCHAR2 (20 BYTE)
ST_COHERENCE_MD	VARCHAR2 (20 BYTE)
ST_COHERENCE_MD_CA_JDD	VARCHAR2 (20 BYTE)
ST_COHERENCE_MD_DONNEES	VARCHAR2 (20 BYTE)
ST_CONFORMITE_DONNEES	VARCHAR2 (20 BYTE)
ST_CONFORMITE_MD	VARCHAR2 (20 BYTE)
ST_COHERENCE_EN	NUMBER
ST_COHERENCE_MAILLE	NUMBER
ST_COHERENCE_ME	NUMBER
VALIDATION_NAT_FK (ID_SUIJET_OBS)	
VALIDATION_NAT_FK1 (AUTO_ID_PROC_REF)	
VALIDATION_NAT_REF_UID_IDX (ID_SINP_OCCATX)	
VALID_ID_SUIJET_OBS (ID_SUIJET_OBS)	

Validation

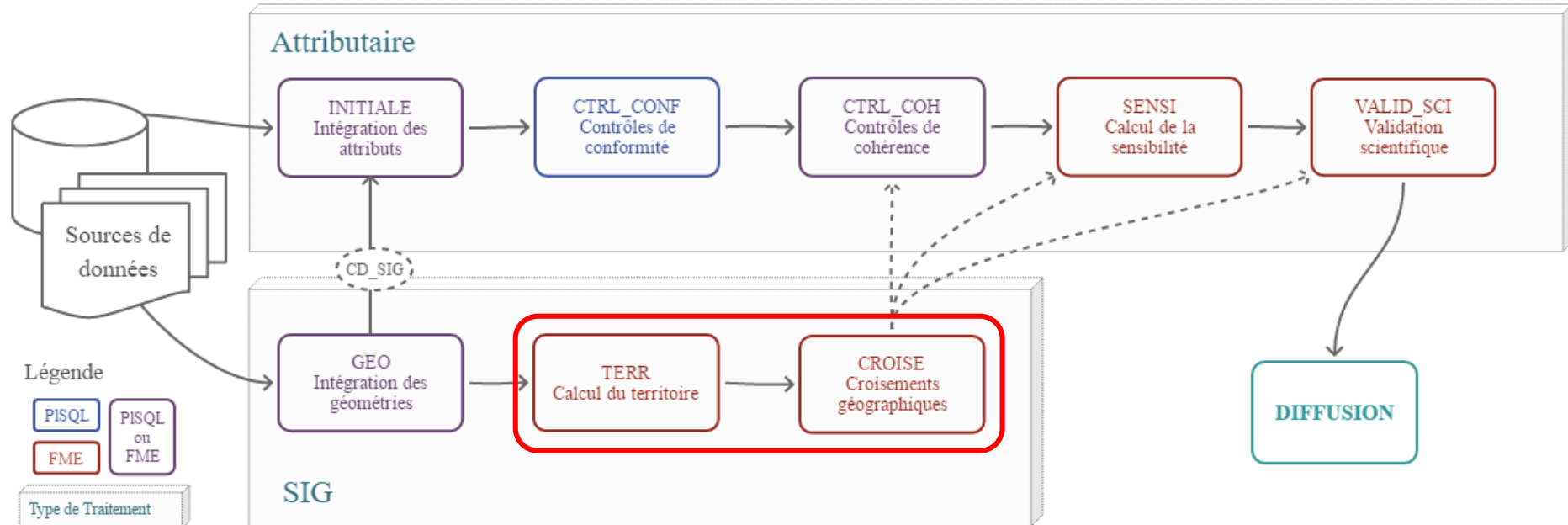
THEMA.PROCEDURES_REF	
* ID_PROC_REF	NUMBER
PROCEDURE	VARCHAR2 (255 BYTE)
PROC_REF	VARCHAR2 (255 BYTE)
PROC_VERSION	VARCHAR2 (20 BYTE)
PROCEDURES_REF_FK (ID_PROC_REF)	
PROCEDURES_REF_FK (ID_PROC_REF)	

THEMA.TYPE VALIDATION	
* CD_VALID	VARCHAR2 (200 BYTE)
* ID	NUMBER
* LIBELLE	VARCHAR2 (255 BYTE)
* TYPE_VALIDATION	VARCHAR2 (200 BYTE)
TYPE_VALIDATION_FK (CD_VALID)	
TYPE_VALIDATION_FK1 (CD_VALID)	

THEMA.VALIDATION SINP	
COMM_VALIDATION	VARCHAR2 (255 BYTE)
DATE_CONTACT	DATE
DATE_CTRL	DATE
ECHILLE_VALIDATION	NUMBER
* ID_SINP_OCCATX	VARCHAR2 (255 BYTE)
NIVEAU_VALIDATION	NUMBER
PERIMETRE_VALIDATION	NUMBER
PROCEDURE	VARCHAR2 (255 BYTE)
PROC_REF	VARCHAR2 (255 BYTE)
PROC_VERSION	VARCHAR2 (200 BYTE)
TYPE_VAL	VARCHAR2 (200 BYTE)
ID_JDD	NUMBER
F * ID_SUIJET_OBS	NUMBER
VALIDATION_SINP_FK (ID_SUIJET_OBS)	
VALIDATION_SINP_UID_IDX (ID_SINP_OCCATX)	
VSINP_ID_SUIJET_OBS (ID_SUIJET_OBS)	

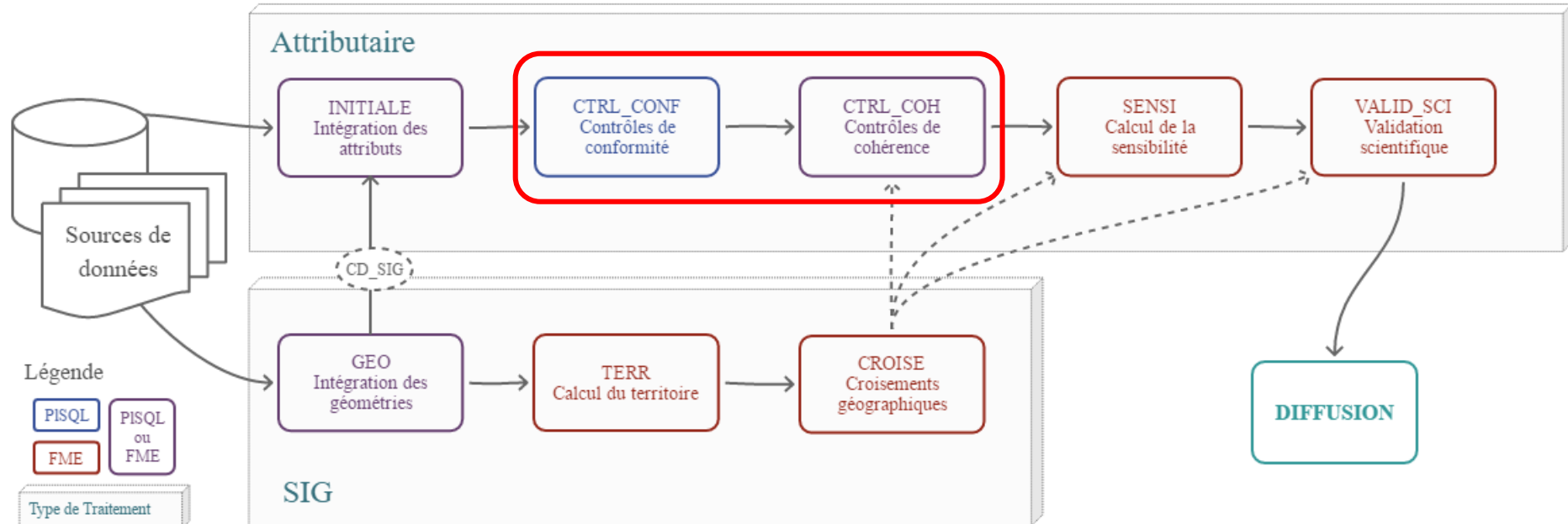


La chaîne de traitement globale



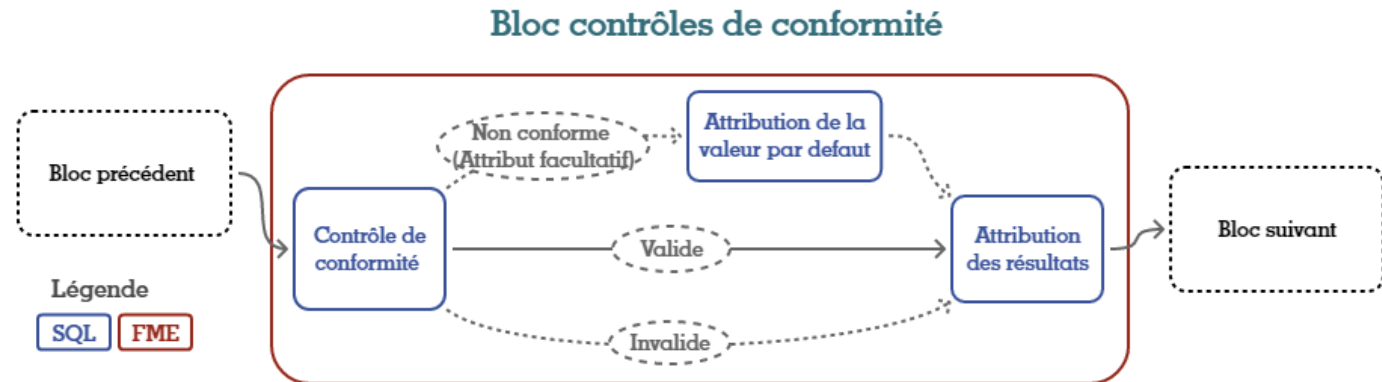
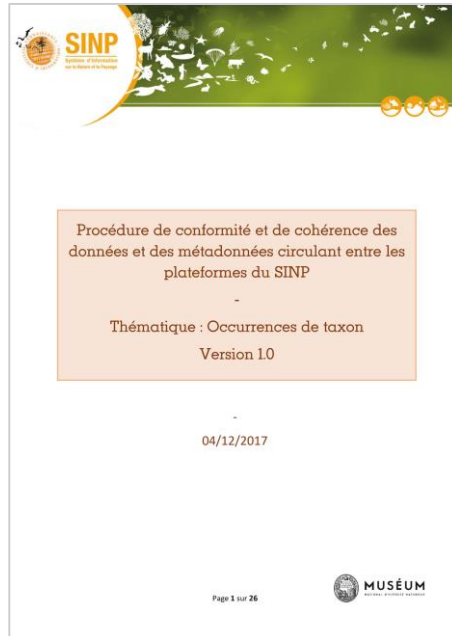


La chaîne de traitement - Conformité / Cohérence





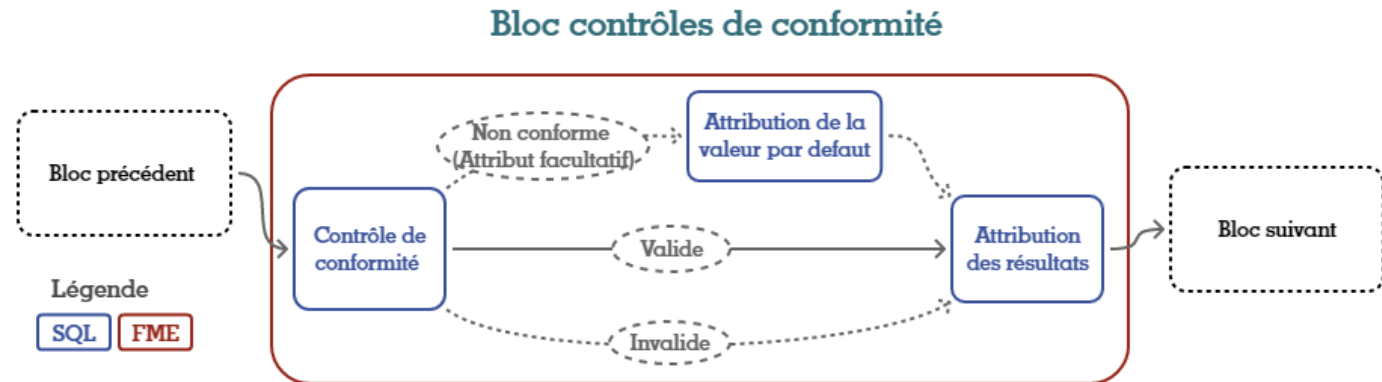
La chaîne de traitement - Conformité / Cohérence



- Compare les données transmises aux normes des standards DEE/INPN



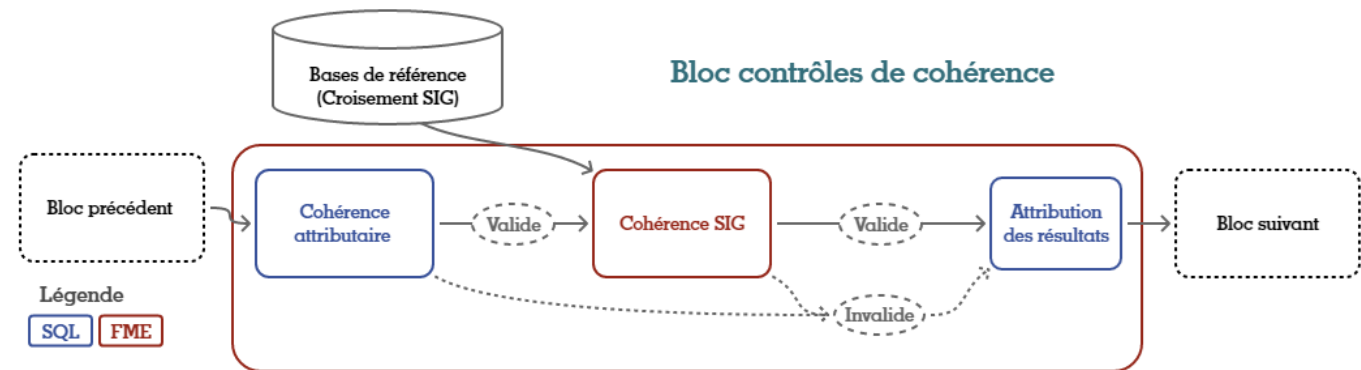
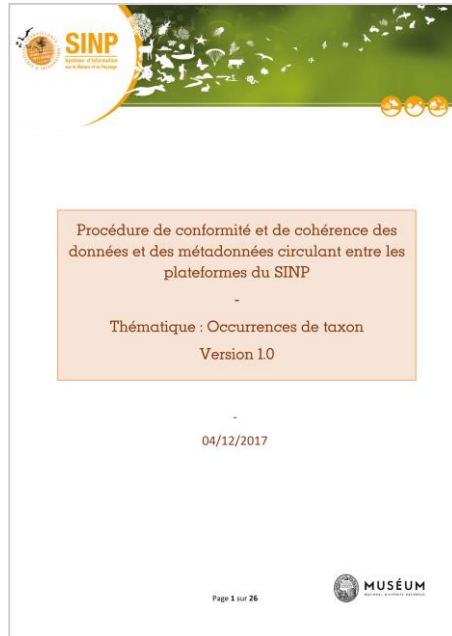
La chaîne de traitement - Conformité / Cohérence



- Compare les données transmises aux normes des standards DEE/INPN



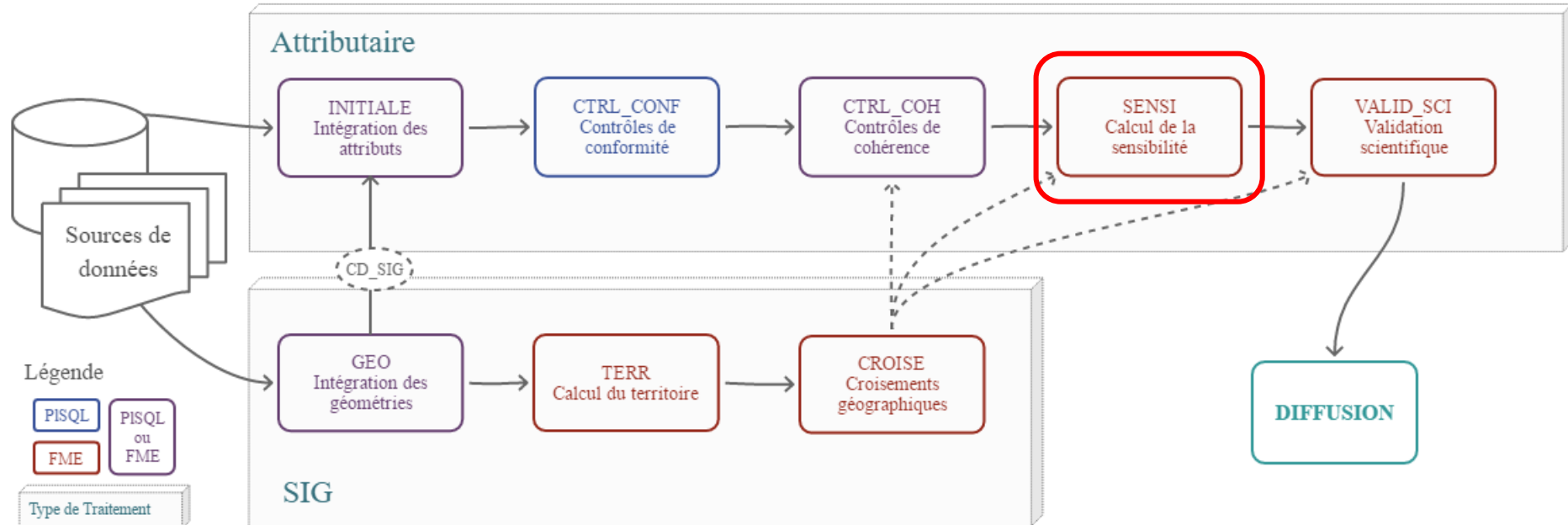
La chaîne de traitement - Conformité / Cohérence



- Compare les données entre elles

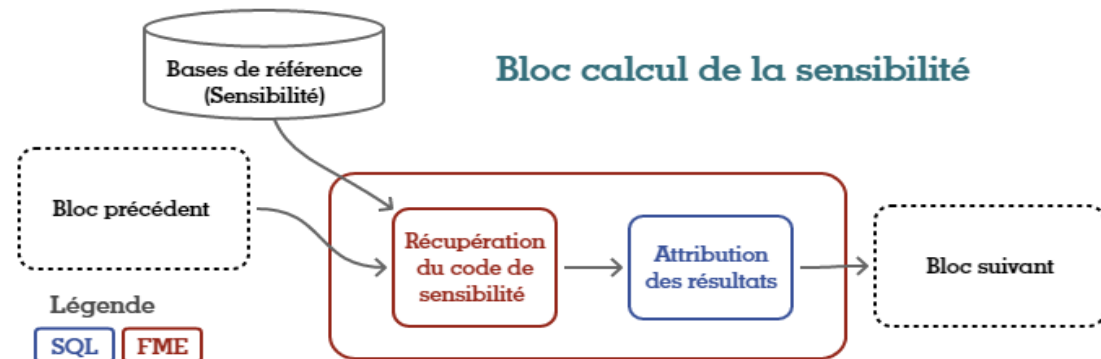
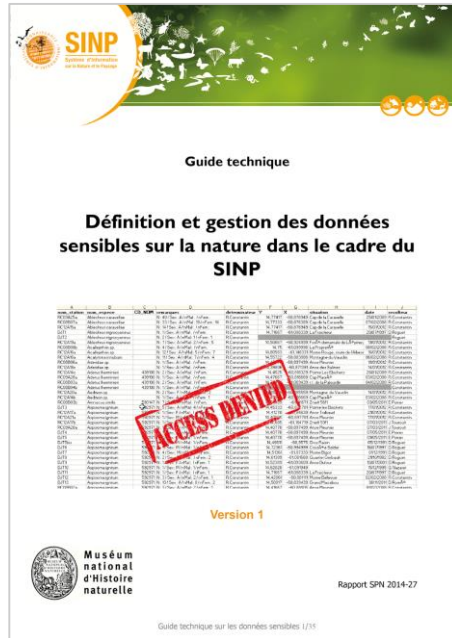


La chaîne de traitement - Sensibilité





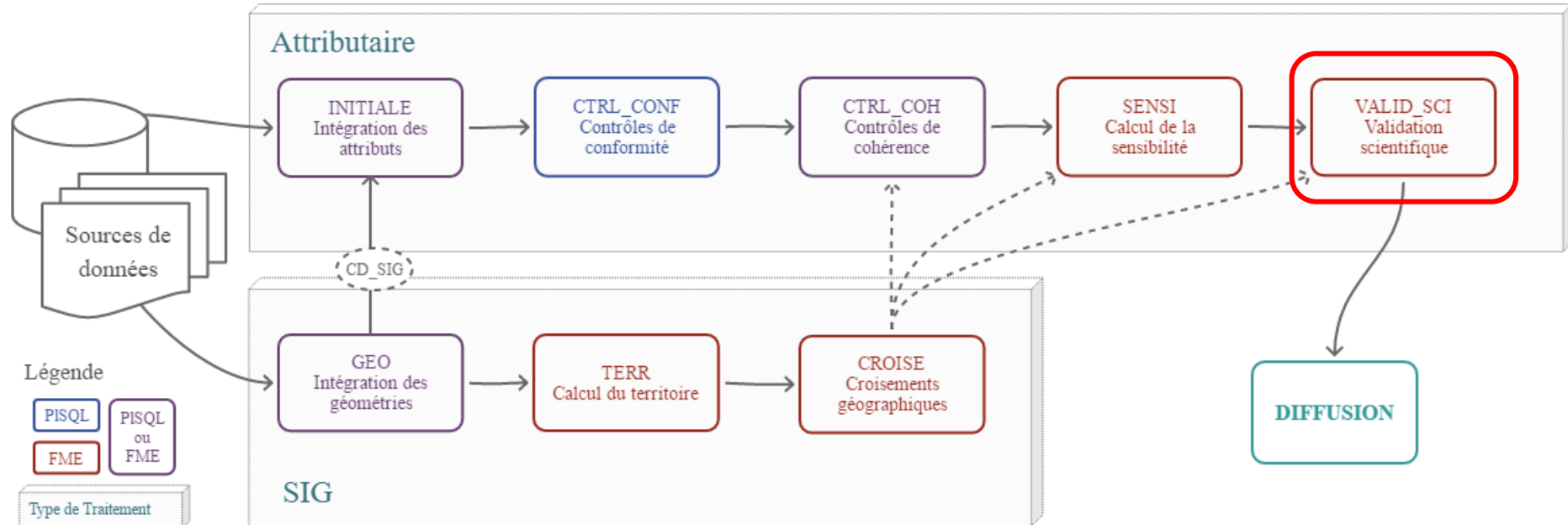
La chaîne de traitement - Sensibilité



- Attribue aux données un niveau de sensibilité selon le référentiel national

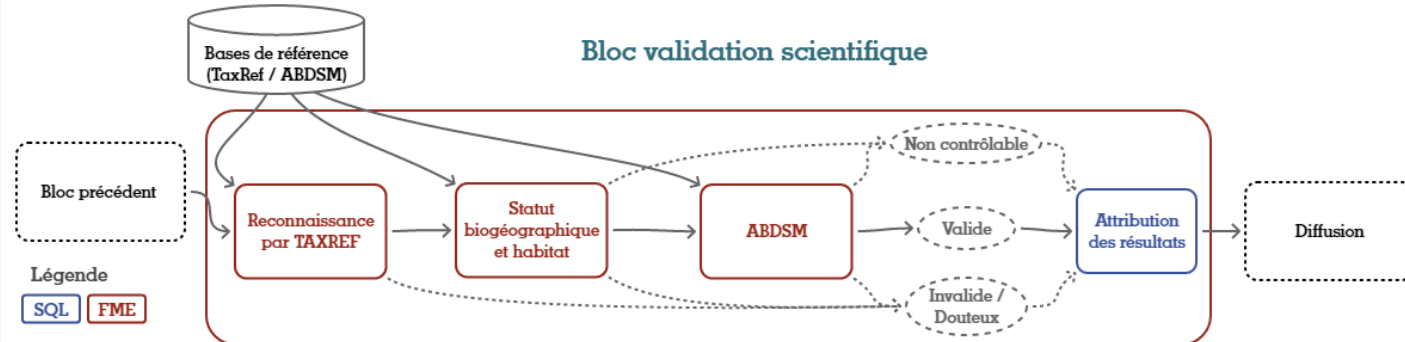
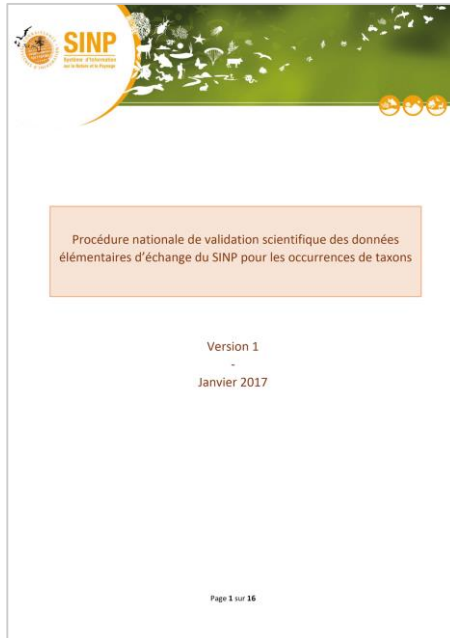


La chaîne de traitement globale





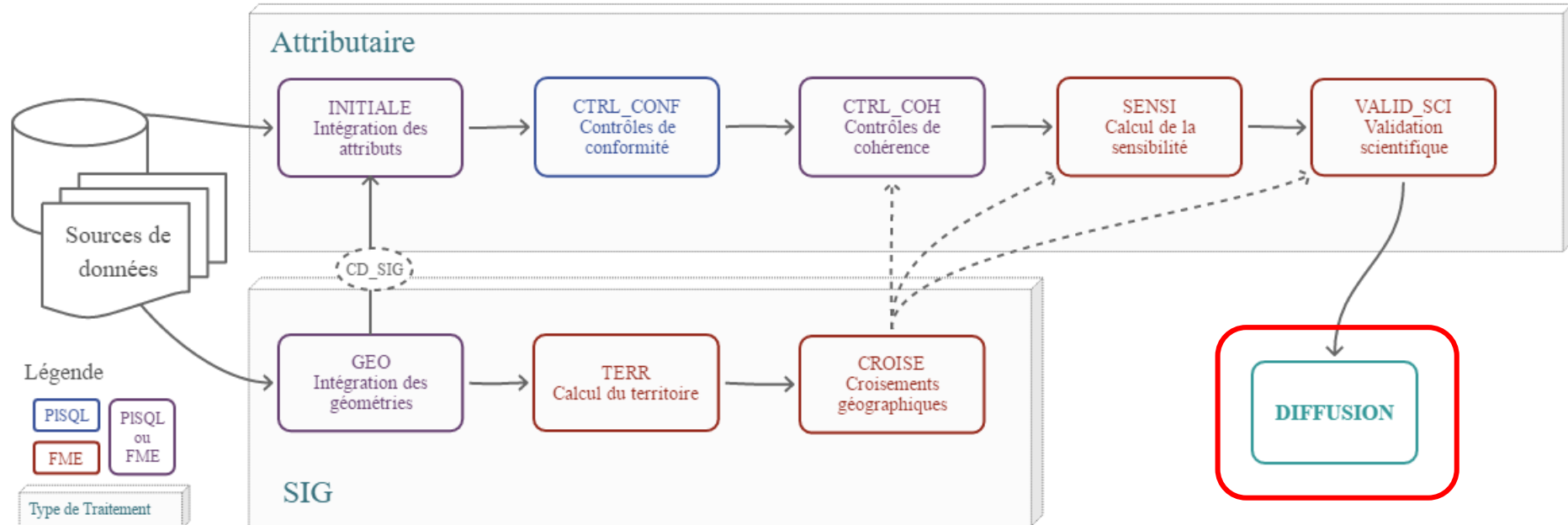
La chaîne de traitement - Validation scientifique



- Attribue un niveau de validité aux données en fonction des référentiels



La chaîne de traitement globale





Exemples de mise en oeuvre - CardObs

- Plus de 340 jeux de données à traiter



CardObs

Saisie, gestion et analyse en ligne
de données naturalistes



Exemples de mise en oeuvre - CardObs

- Plus de 340 jeux de données à traiter
- Mapping de deux modèles de données



CardObs

Saisie, gestion et analyse en ligne
de données naturalistes



Exemples de mise en oeuvre - CardObs

- Plus de 340 jeux de données à traiter
- Mapping de deux modèles de données
- Rapidité de traitement = augmentation de la fréquence de mise à jour possible



CardObs

Saisie, gestion et analyse en ligne
de données naturalistes



Exemples de mise en oeuvre - Plateforme régionale

- Des données répondant au standard DEE





Exemples de mise en oeuvre - Plateforme régionale

- Des données répondant au standard DEE
- Des quantités de données et de jeux de données variables





Exemples de mise en oeuvre - Plateforme régionale

- Des données répondant au standard DEE
- Des quantités de données et de jeux de données variables
- Des données SIG de qualités différentes





Exemples de mise en oeuvre dans un futur proche

- GBIF
 - Mapping DwC / standard INPN



Exemples de mise en oeuvre dans un futur proche

- GBIF
 - Mapping DwC / standard INPN
- Collections du MNHN
 - Prise en compte des particularités de ce type de données



Exemples de mise en oeuvre dans un futur proche

- GBIF
 - Mapping DwC / standard INPN
- Collections du MNHN
 - Prise en compte des particularités de ce type de données
- Etudes d'impact (dépôt légal)
 - Traitement en continu



Conclusion

- Une refonte complète des méthodes utilisées ...



Conclusion

- Une refonte complète des méthodes utilisées ...
- ... Mais une réduction considérable du temps de latence entre la réception et la diffusion des données



Conclusion

- Une refonte complète des méthodes utilisées ...
- ... Mais une réduction considérable du temps de latence entre la réception et la diffusion des données
- Une réflexion sur les processus de mise à jour des données à prévoir



Conclusion

- Une refonte complète des méthodes utilisées ...
- ... Mais une réduction considérable du temps de latence entre la réception et la diffusion des données
- Une réflexion sur les processus de mise à jour des données à prévoir
- Un travail à mener en parallèle de l'évolution des standards et des règles de contrôle



Conclusion

- Une refonte complète des méthodes utilisées ...
- ... Mais une réduction considérable du temps de latence entre la réception et la diffusion des données
- Une réflexion sur les processus de mise à jour des données à prévoir
- Un travail à mener en parallèle de l'évolution des standards et des règles de contrôle
- Une méthode adaptable à d'autres outils



Merci de votre attention