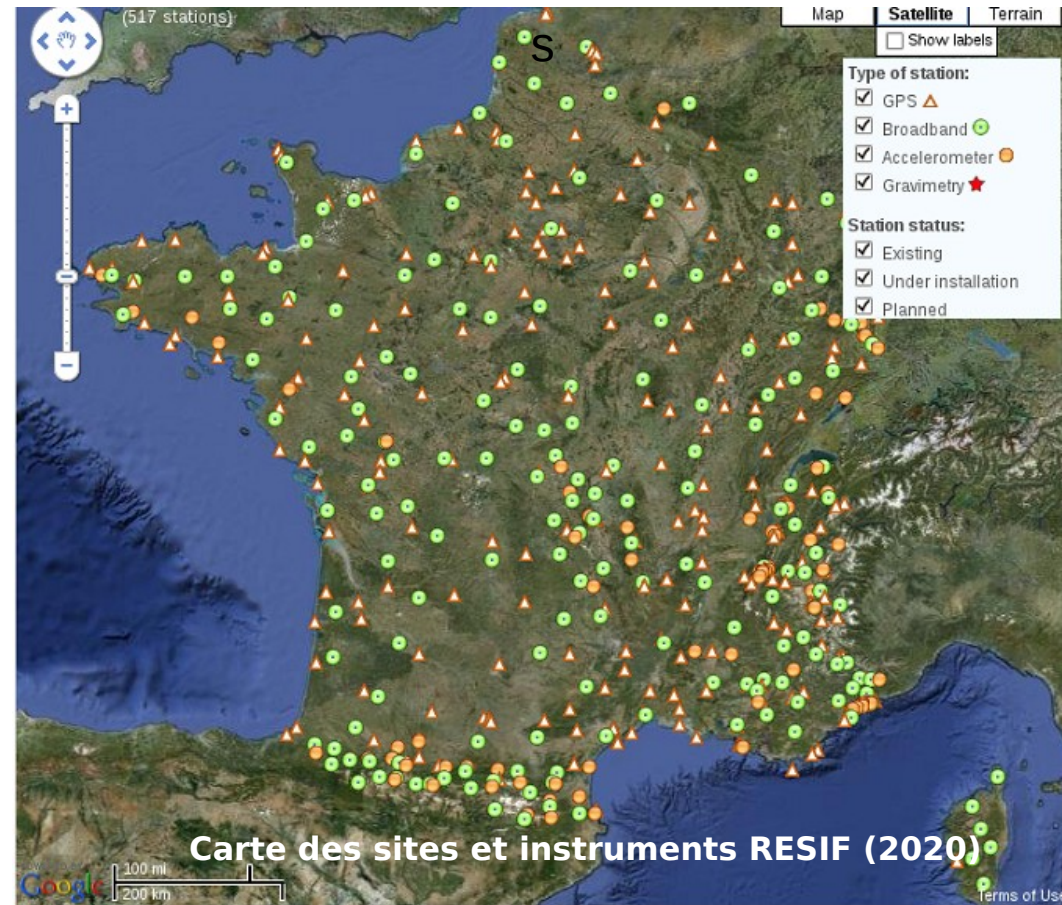


Gestion des données sismologiques de l'infrastructure de recherche RESIF (Réseau Sismologique et géodésique Français)

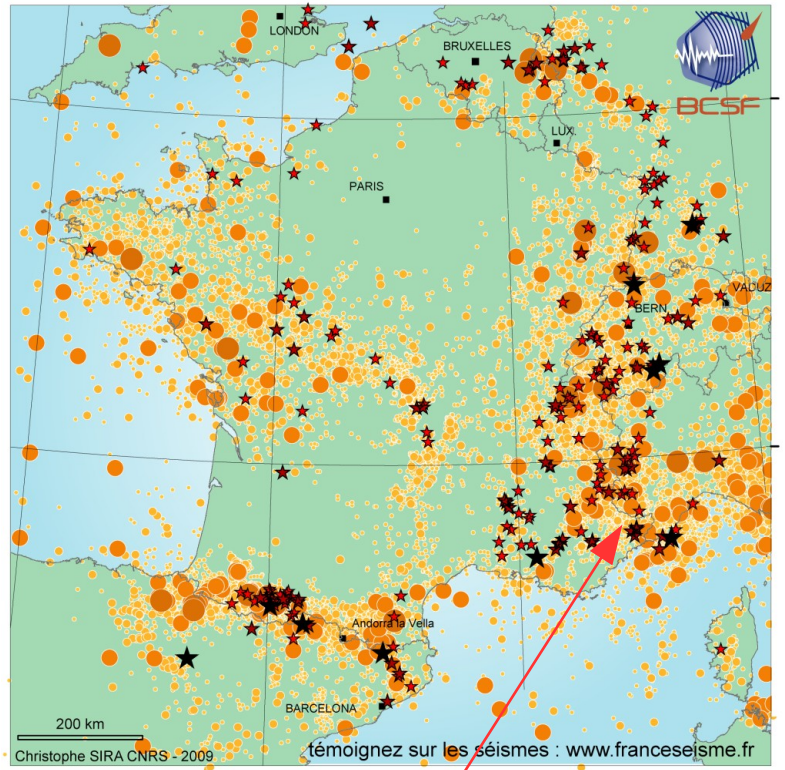
G. Arneodo, H. Pedersen, C. Péquegnat, P. Volcke, D. Wolyniec (ISerre & OSUG)
RESIF-SI working group



- Antenne intégrée d'observation de la déformation de la terre (sismologie, *géodésie, gravimétrie*)
- Très Grande Infrastructure de Recherche (inscription 2009)
- Contribution majeure de la France au projet européen EPOS (European Plate Observation System)
- Construction sur 10 ans, budget consolidé de 60 Mo, Equipex 2012 (9,3 Mo)
- Tutelles : MESR, MEDDE
- Consortium de 19 membres dont l'Université Grenoble Alpes et le CNRS-INSU
- S'appuie sur 11 Observatoires
- Projet majoritairement instrumental
- Données libres d'utilisation et gratuites



RESIF et la recherche. L'une des applications majeures : aléa et risque sismique



Séismes historiques connus depuis 1000 ans d'intensité supérieure ou égale à VIII (Sisfrance)

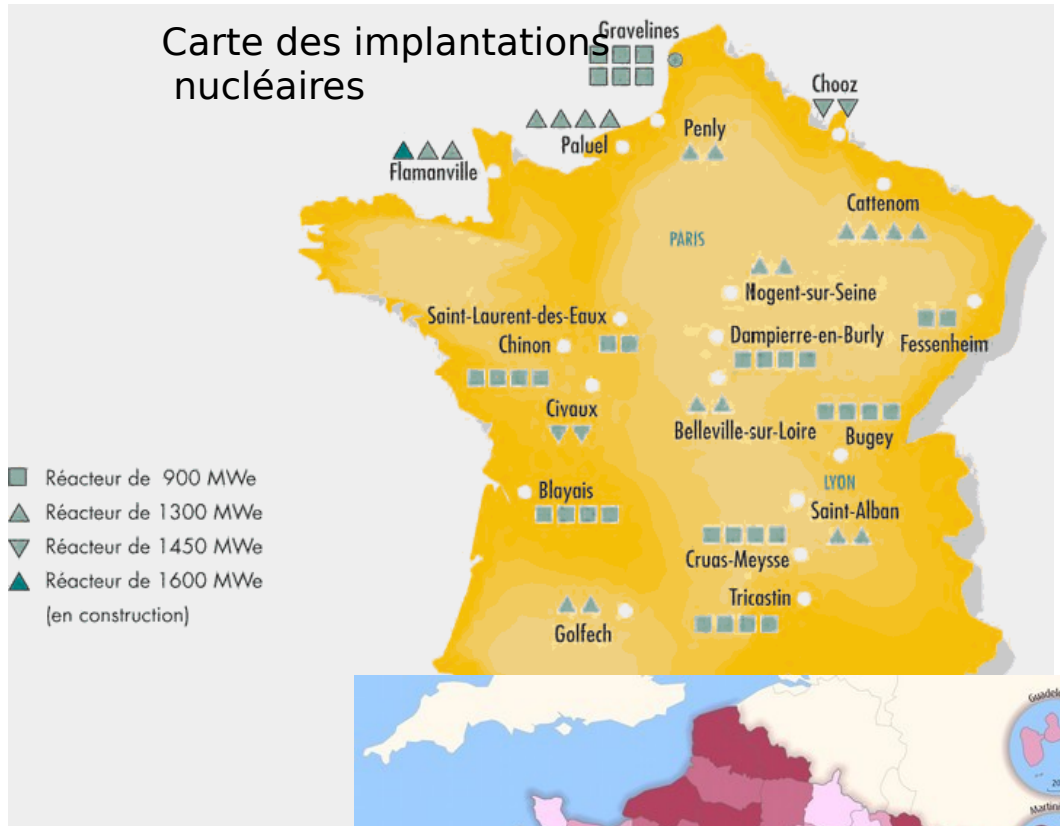
- ★ IX - dommages généralisés aux constructions
- ★ VIII - destructions partielles de bâtiments
- ★ VII - dommages aux constructions

Séismes par magnitude enregistré instrumentalement de 1964 à 2008 (LDG-ReNaSS)

- supérieure ou égal à 5
- 4 - 4,9
- 3 - 3,9
- 2 - 2,9



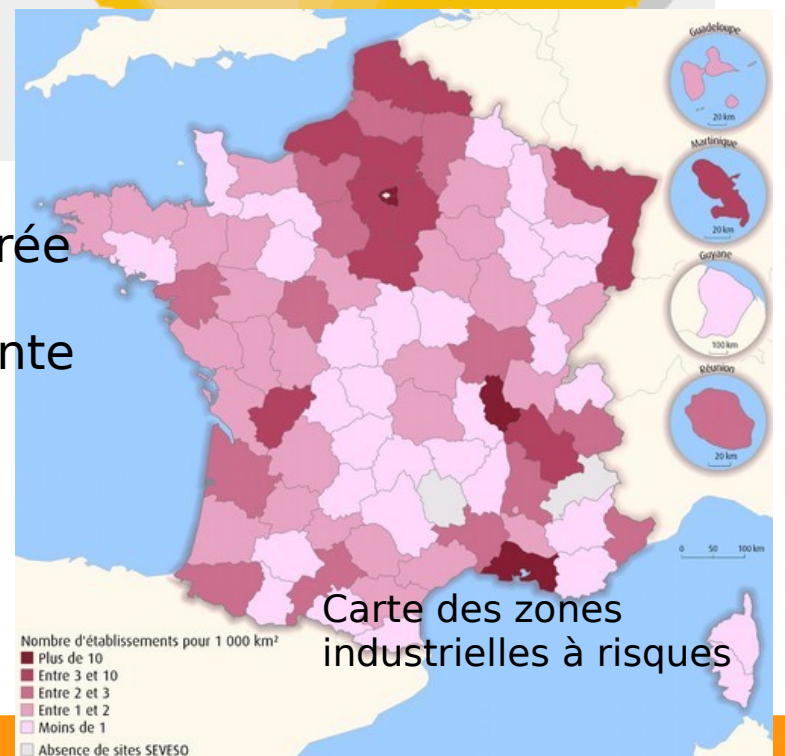
Séisme de Lambesq 1909 (M 6.2)



- Réacteur de 900 MWe
- ▲ Réacteur de 1300 MWe
- ▼ Réacteur de 1450 MWe
- ▲ Réacteur de 1600 MWe (en construction)

Source: Ministère de l'Environnement

- Sismicité faible à modérée
- Vulnérabilité importante

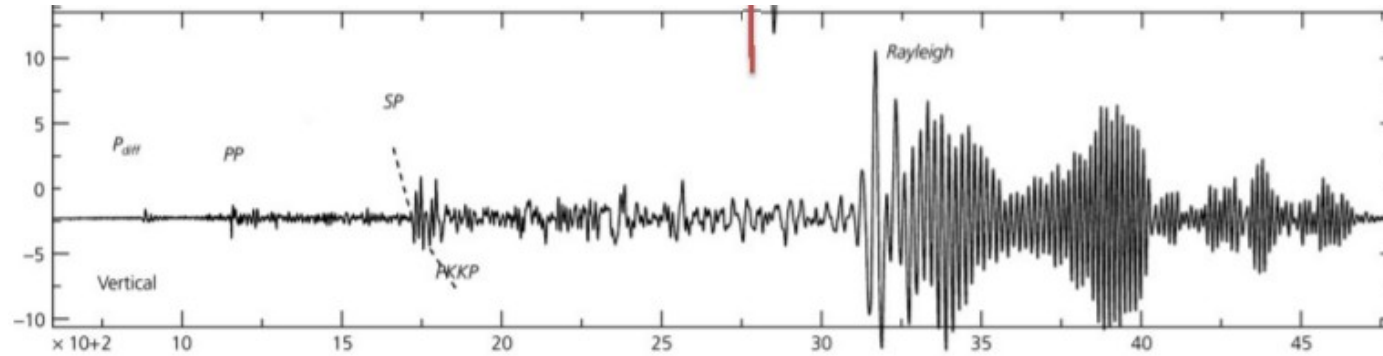


- Nombre d'établissements pour 1 000 km²
- Plus de 10
 - Entre 3 et 10
 - Entre 2 et 3
 - Entre 1 et 2
 - Moins de 1
 - Absence de sites SEVESO

Carte des zones industrielles à risques

Données sismologiques et réseaux

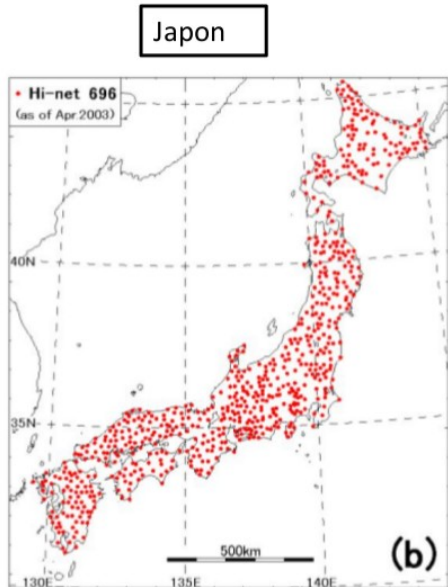
Données primaires : formes d'ondes (séries temporelles, échantillonnage constant [1p/sec à 1000p/sec])



From Stein and Wysession

Métadonnées : données administratives, spatio-temporelles, données de site (géologie, bâti), réponse instrumentale complète d'une chaîne d'acquisition

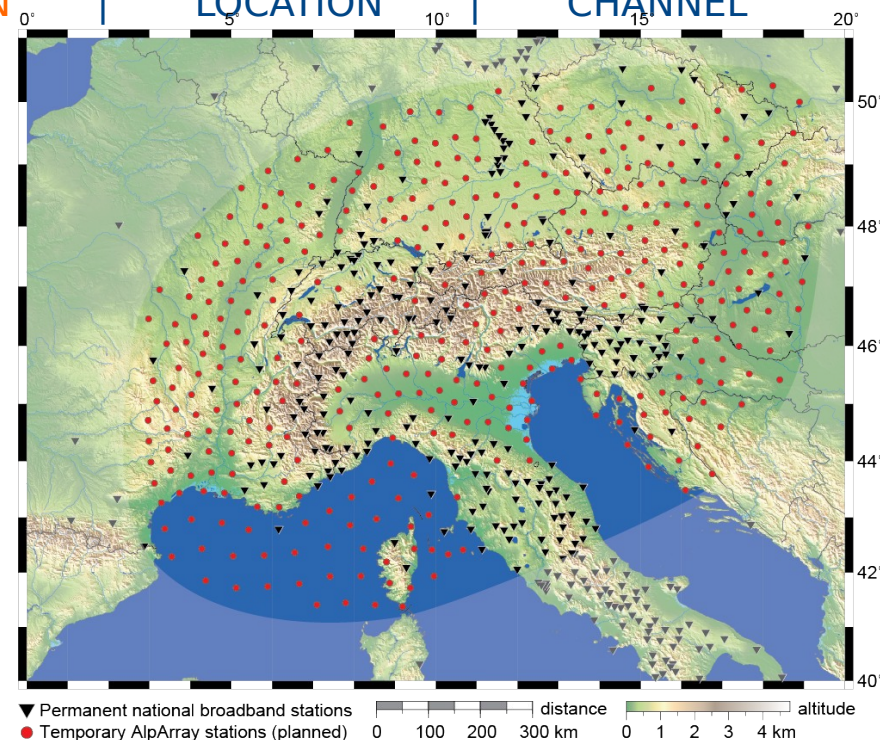
Identité des données : **NETWORK** | **STATION** | **LOCATION** | **CHANNEL**



- initiative d'échelle européenne (>15 pays impliqués);
- réseau sismologique large-bande dense, couverture homogène et dense (40-60 km) des Alpes, leurs avant-pays, leurs terminaisons

~350 stations permanentes + ~350 stations temporaires à terre + 33 OBS LB (2 ans)

AlpArray (EU) : 2015-2019 **RESIF**



Source : Brenguier F.

Missions : mise à disposition des données de qualité (validées, métadonnées complètes), respects des standards internationaux (formats, services de bases), délais de mise à disposition de qq sec. À qq jours. Interconnexion aux fédérations de centres de données

L' Université Grenoble Alpes est fortement impliquée :

- Via l' OSUG et les SNO¹ (Réseau Accélérométrique Permanent) et SISMOB²
- L' UJF porteur du projet 'Centre de données national pour la sismologie (2012-2016)' (avec ISTerre et l'OSUG)
- RESIF et la DSI de l'UJF initiateurs du projet SUMMER (2013, P. Volcke, J. Le Tanou) sur lequel les données RESIF sont déposées.
- Mandat renouvelé à l'UGA en 2016(-2020) pour l'exploitation du centre de données RESIF

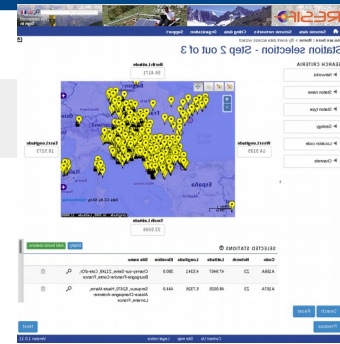
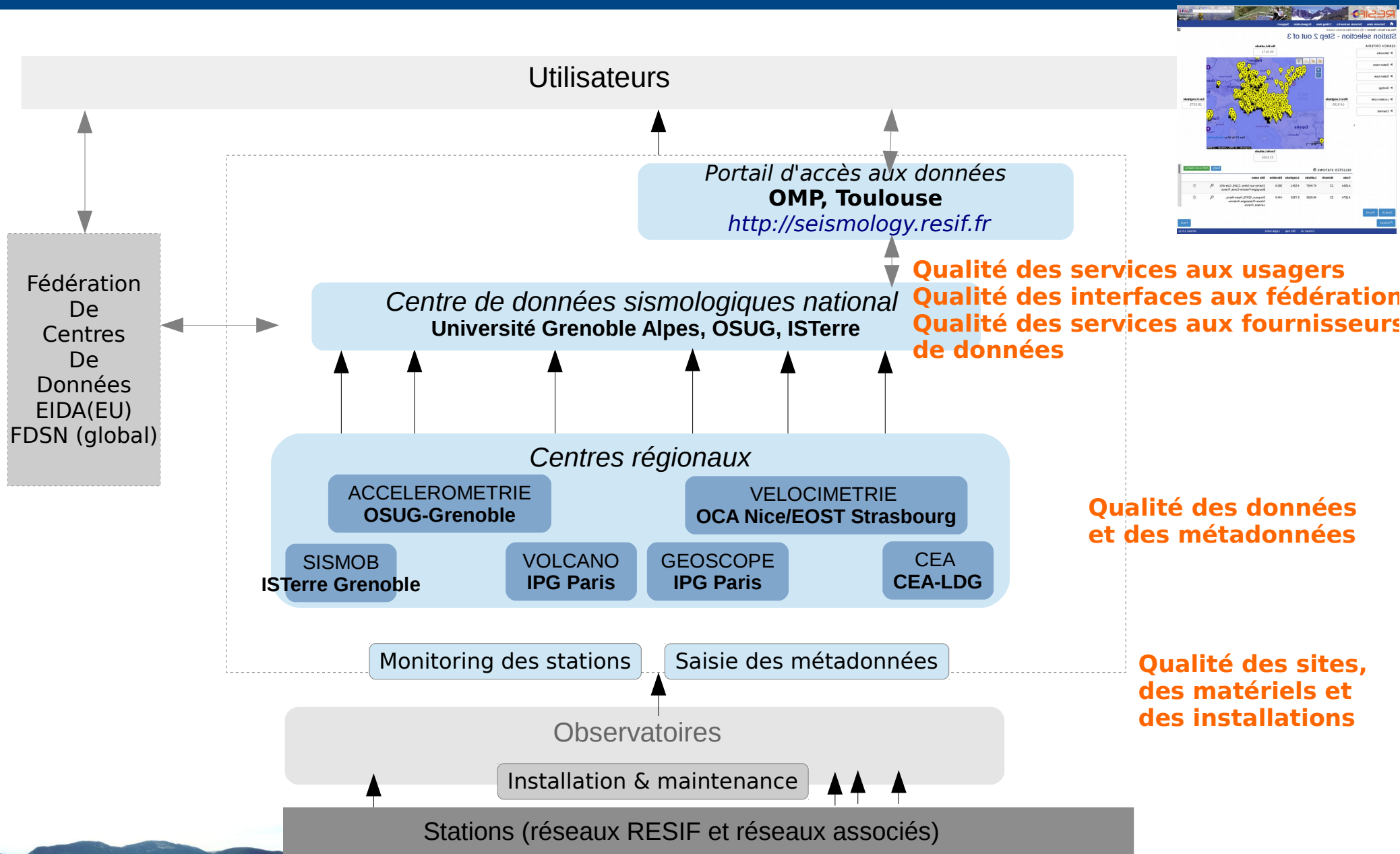
Aujourd'hui (T + 5 ans)

- Services opérationnels, interfacés aux canaux internationaux d'échange de données (archive virtuelle européenne EIDA, fédération mondiale des sismographes numériques (FDSN))
- Volumétrie : 55 Tb de données collectées (> 1982) et 35 Tb de données servies, 11 réseaux permanents, ~ 50 réseaux mobiles, 2000 points de mesure
- Les données d'observation et de recherche sont normalisées et centralisées, elles sont accessibles à partir d'un point de diffusion unique
- Plusieurs centaines d'utilisateurs (France, EU, US), montée en charge attendue (densification des sources de données + fédération moteurs de requêtes)

1 Service National d'Observation en Terre Solide

2 Parc national d'instruments sismologiques terrestres mobile INSU

Le système d'information de RESIF



Briques de bases pour la distribution des données RESIF

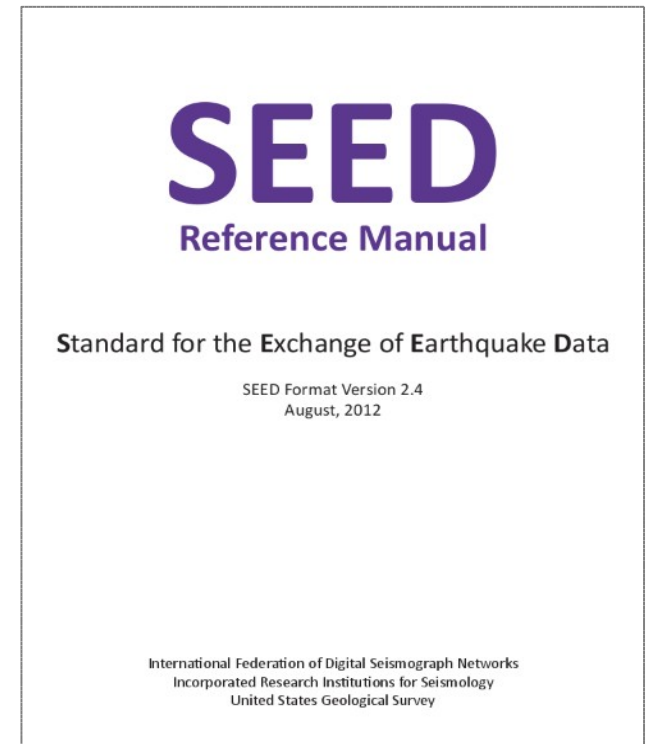
- Format des données et des métadonnées (normalisés)
- Services de données (normalisés) :
 - Service de données en temps quasi réel (streaming)
 - Accès aux données par 2 WEBServices :
 - Accès aux formes d'ondes : dataselect
 - Accès aux métadonnées : station
- Services de données en cours de normalisation (spec EU, spec US)
 - Service d'accès aux informations sur la qualité des données : wfcatalog
 - Divergences (EU/US) mineures sur les informations calculées et les modes de calcul
 - Services pour le routage : affichage du catalogue des données et services exposés :
 - Divergence (EU/US) majeure sur la façon de construire un annuaire des données : 'harvesting' (US) et déclaratif (EU)
- Service 'datacentre dépendant'
 - Accès rsync à des ensembles de données
- Un portail <http://seismology.resif.fr>

Briques de base : format des données et des métadonnées



International Federation of Digital Seismograph Networks

- Format défini par la FDSN - organisme en charge de la spécification des formats, protocoles et interfaces pour l'échange des données et des métadonnées géophysiques.
- Format historique (1987 !) orienté exclusivement vers la géophysique, sous-efficace d'un point de vue informatique
- ... compliqué, redondance partielle des métadonnées dans les données. Une évolution xml pour les métadonnées.
- ... partage effectif des données depuis les années 1990
- ... aujourd'hui, les données sismologique publiques mondiales sont en SEED
- ... les évolutions vers d'autres formats et modèles (HDF5, OGC, sensorML....) sont en cours de discussion et difficiles
- ... l'interopérabilité est du domaine d'EPOS



http://www.fdsn.org/seed_manual/SEEDManual_V2.4.pdf

Echange des données en temps quasi réel

- utilisé par les services d'alerte tsunami, *early warning*, etc. France : Cenalt <http://www.info-tsunami.fr/>
- protocole métier, mode connecté, supporte des déconnexions/reconnexions successives sans perte de données
- serveurs et clients 'domain public', serveur implémenté aux stations

```
slinktool -S 'FR_RUSF' -A resif.rusf.miniseed rtserve.resif.fr:18000
```

```
slinktool -S 'FR_RUSF' -A iris.rusf.miniseed rtserver.iris.edu:18000
```

- Pas d'annuaire mondial des services de données temps réel, « tout le monde » peut tout collecter/diffuser, données NON VALIDÉES
- Toutes les données « temps réel » diffusées par RESIF sont garantie 'avec' métadonnées

Documentation et exemples :

http://seismology.resif.fr/#CMSConsultPlace:NEAR_REALTIME_DATA

Briques de base : l' accès aux données par webservice FDSN

Accès aux métadonnées : webservice station

- Critères d'identité (network, station, location, channel), de localisation, de durée

<http://ws.resif.fr/fdsnws/station/1/query?starttime=2011-03-11T05:00:00&endtime=2011-03-12T05:00:00&level=response>

<http://ws.resif.fr/fdsnws/station/1/query?level=station>

Accès aux données : webservice dataselect

- Critères d'identité et de durée

<http://ws.resif.fr/fdsnws/dataselect/1/query?starttime=2011-03-11T05:00:00&endtime=2011-03-12T05:00:00&quality=B>

<http://service.iris.edu/fdsnws/dataselect/1/query?starttime=2011-03-11T05:00:00&endtime=2011-03-12T05:00:00&quality=B>

RESIF : 1ère implémentation européenne des webservice station et dataselect

Documentation et exemples :

http://seismology.resif.fr/#CMSConsultPlace:FDSN_WEBSERVICES

Accès aux informations de qualité sur les données : webservice WFCatalog

- Critères d'identité (network, station, location, channel), de localisation, de durée

<http://ws.resif.fr/wfcatalog/alpha/query?start=2015-06-01&end=2017-06-02&net=Z3>

Catalogue de services

http://ws.resif.fr/eida_routing.xml

- 7 posters AGU, EGU , de nombreuses présentations
- 2 actions nationales de formation (avec le CNRS) « la mesure en sismologie » dédiées à la formation des ingénieurs à la récupération et à la manipulation des données
- D.O.I attachés aux différents réseaux d'observations et de recherche
 - Voir http://seismology.resif.fr/#CMSConsultPlace:DOI_INVENTORY
 - Exemple : <http://data.datacite.org/10.15778/RESIF.FR>
- Référencement Registry of Research Data Repositories
<http://www.re3data.org/search?query=RESIF>

- Ca paraît simple (et en fait, ça l'est). On n'a rien fait de très compliqué, informatiquement parlant
- Cependant, si on reprend les demandes faites aux ingénieurs des observatoires et laboratoires concernés (~ 10 à 14 ingénieurs) il y a 6 ans :
 - 1) Imaginer un système réparti tenant compte de l'existant (SNO, Réseaux et Observatoires) ;
 - 2) Rédiger collaborativement les cahiers des charges des différents éléments ;
 - 3) Jouer la concurrence et candidater sur les différents lots ;
 - 4) Après attribution des lots, travailler collaborativement dans la coordination, la concertation, l'accompagnement, chacun dans son contexte (OSU et laboratoires), lesquels sont « parfois » en concurrence.
- ... les verrous à lever et les difficultés n'étaient pas tant d'ordre technique qu'organisationnel et humains.
- ces aspects sont difficilement évaluable en début de projet, et ont impacté très fortement le calendrier des livraisons

RESIF et GRICAD data axe1

- Réflexions en cours autour de la problématique des licences à associer aux données

RESIF et GRICAD data axe4

- Des pistes techniques à explorer à la frontière de l'ingénierie « données » et « calcul », par exemple intégration d'API spécialisées dans les codes de calcul (ex. IRODS.), en lien avec le centre de données OSUG
- Projet MASTODONS i-WORMS (GRICAD, RESIF, ISTERRE)

RESIF et GRICAD data axe2

- Traçabilité des données : réflexion en cours dans le cadre du projet EUDAT (associer un PID à chaque requête, assurer la reproductibilité des requêtes)
- Versionner des données (?), versionner les métadonnées et être capable de les réassocier aux requêtes
- Certifier les données ?
- Métriques à partir des DOI, produire des H-Index pour les réseaux et les données
- Data paper

Etc....

Merci de votre attention ! ... questions ???