



MaDICS

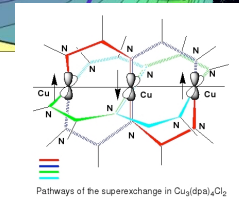
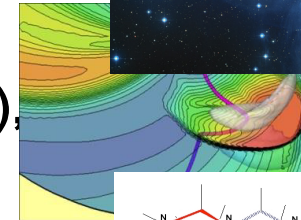
in the nutshell

Paris 6 décembre 2017, **Conseil Scientifique MaDICS**

*Christine Collet, Vincent Claveau
Pierre Gançarski, Aurélien Garivier*

Masses de données (scientifiques)

provenant d'instruments, de simulations numériques, de multiple dispositifs de collecte de données ...
 produites en grande quantité (**Volume**),
 à un rythme intense (**Vélocité**),
 ne pouvant plus être stockées (**Volatilité**),
 de nature très hétérogènes (**Variété**),
 avec des sémantiques variables au cours du temps (**Variabilité**),
 de qualité et précision peu contrôlables (**Véracité**).



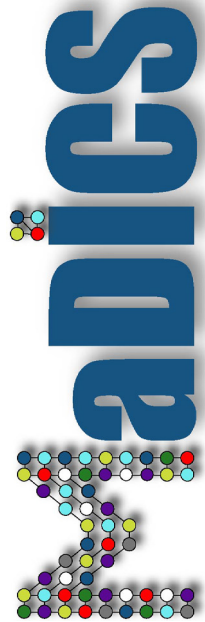
Nouvelle démarche scientifique

De : Croire -> Vérifier (avec des données)
 A : Découvrir (dans les données) => Comprendre



Créer un écosystème - une dynamique de rapprochement entre scientifiques de différentes disciplines ou communautés afin d'élaborer ensemble les nouvelles méthodes et outils pour la gestion, l'exploitation et la valorisation des données scientifiques.

- **Organiser des activités interdisciplinaires** (journées d'animation scientifiques/thématiques, des ateliers, séminaires, hackaton, ...) et des **réponses collectives** à différents appels à projets
- **Contribuer à être un lieu de veille et d'expertise pour les décideurs** identification de compétences, notes de synthèse ou de prospective
- **Favoriser la formation** des spécialistes en Sciences de Données avec une attention particulière aux jeunes chercheurs.



Volume

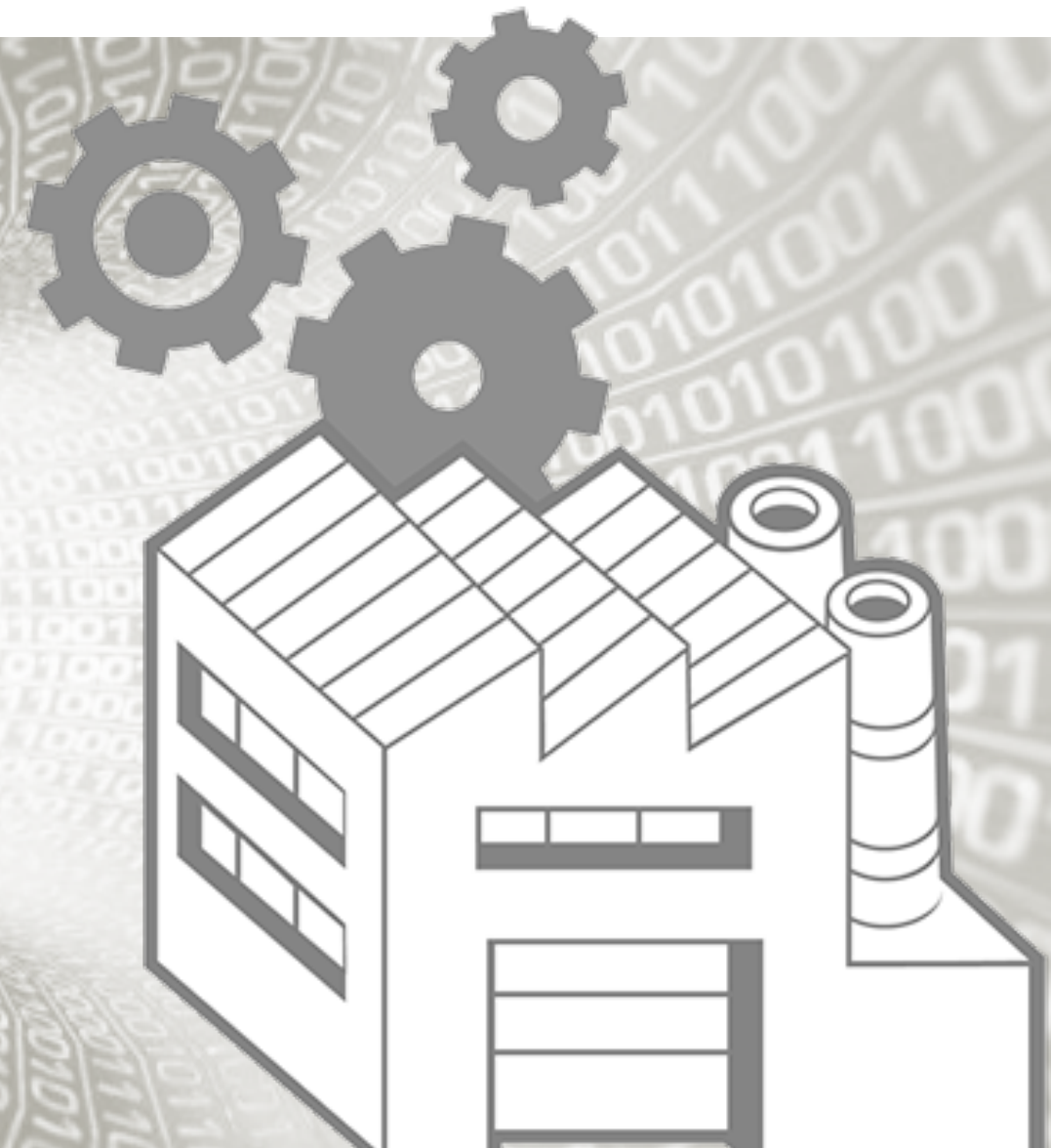
Vélocité

Volatilité

Variété

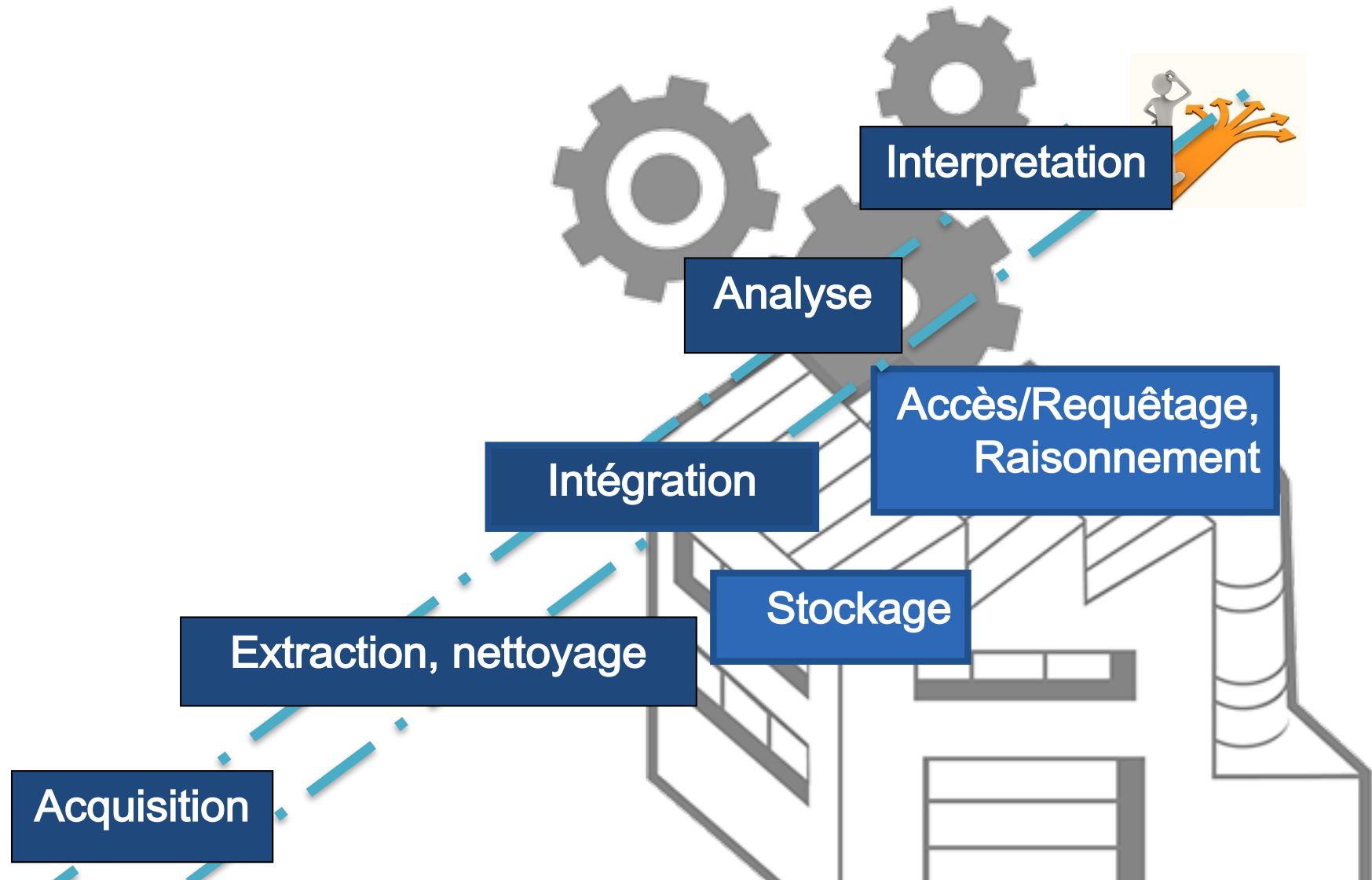
Variabilité

Véracité

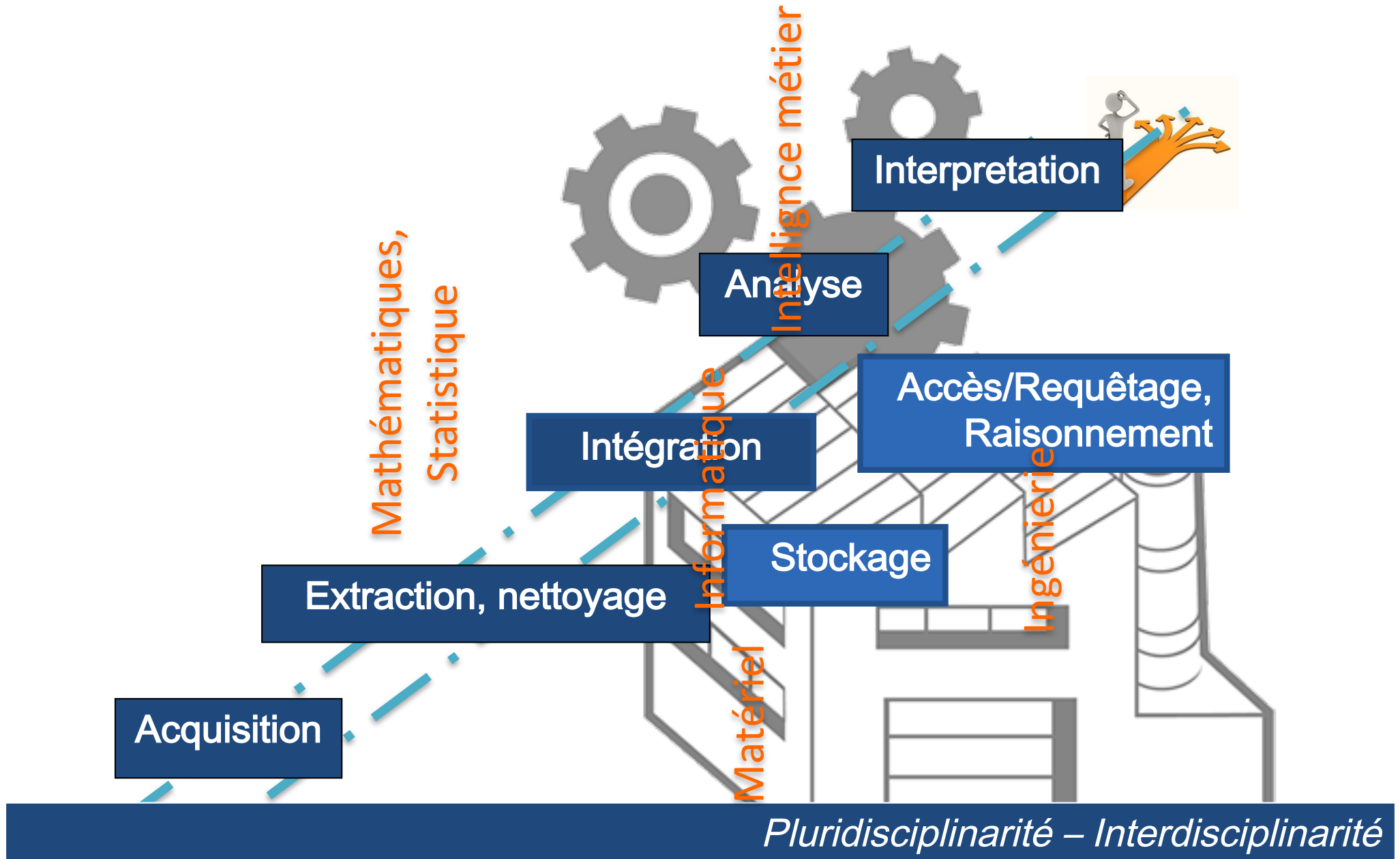


Activités d'animation, de veille, de prospective

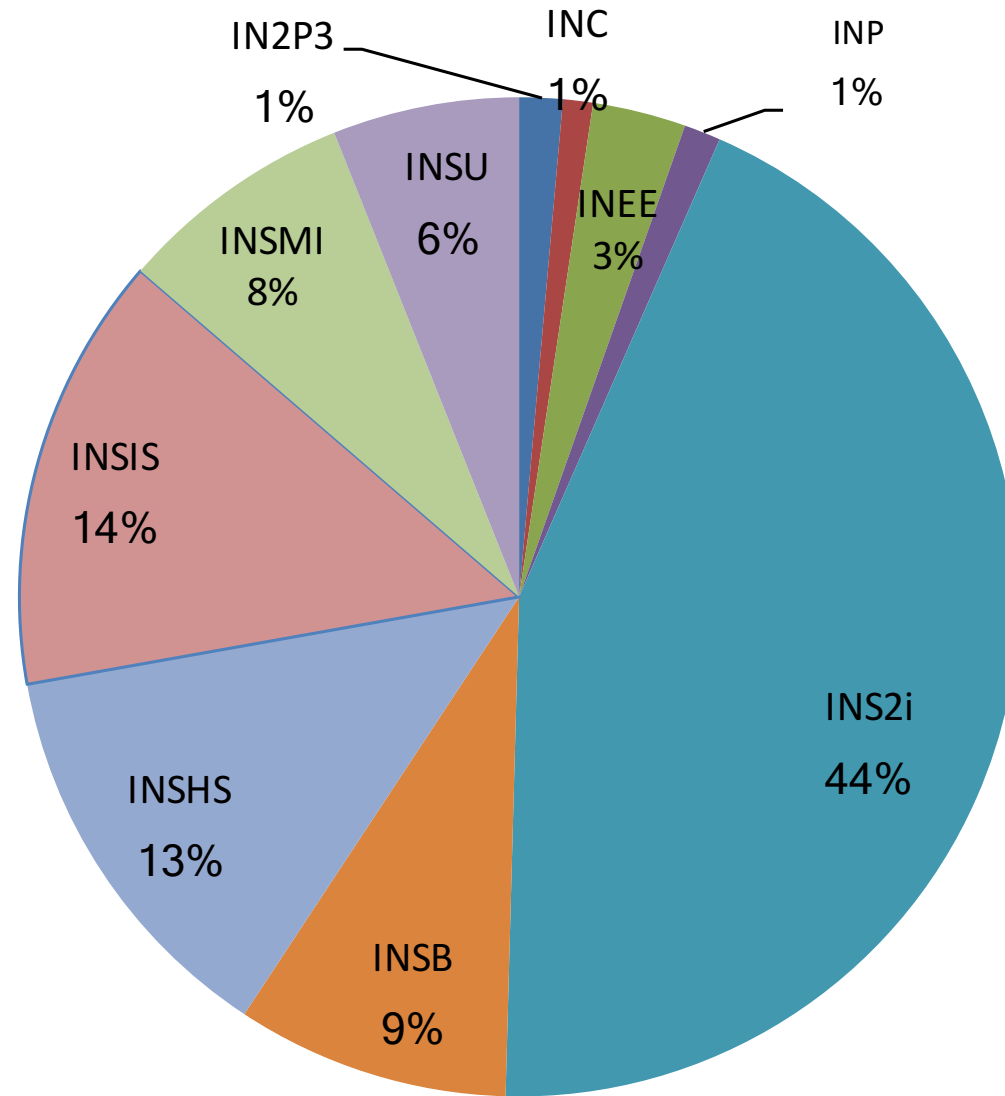
Chaîne de traitement

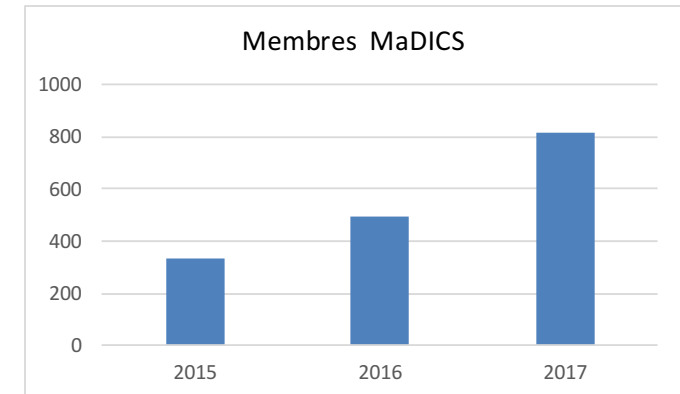


Nouveaux outils algorithmiques et mathématiques



Pluridisciplinarité – Interdisciplinarité





815 Membres (2017)

–504 Chercheurs et Enseignant-chercheurs

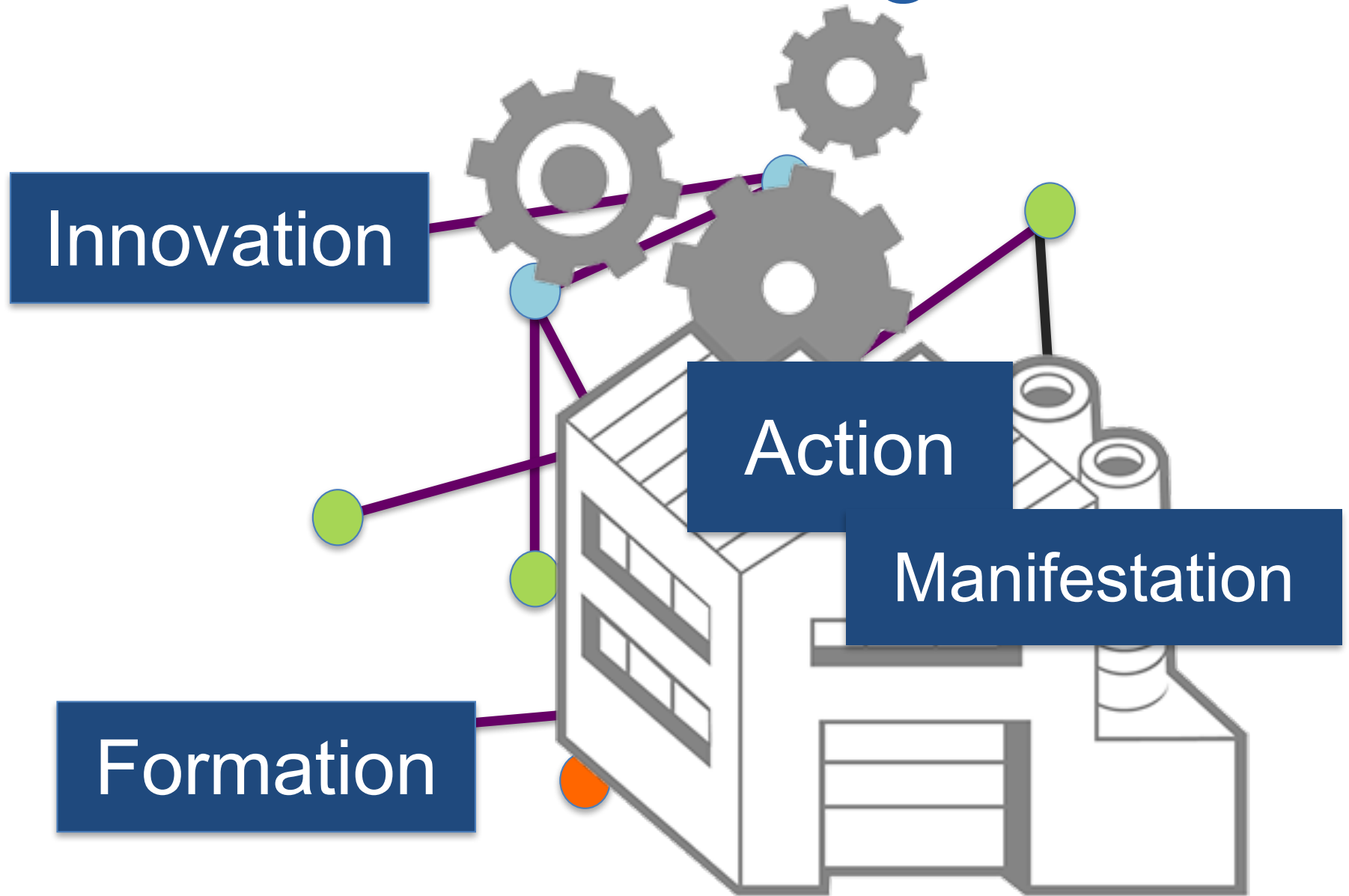
–143 Post-Doc + Doc

–93 Industriels

–75 Autres Instituts (INRA, INRIA, INSERM, INA, IRD, CEA, CHU, ...)

63 % Chercheurs, 17% Post-doc + Doc, 20% Autres+Indus

89% académiques, 11% industriels



Une Action propose un projet d'animation interdisciplinaire coordonnant diverses activités (Manifestations).

Un appel par an

Élément d'animation structurant :

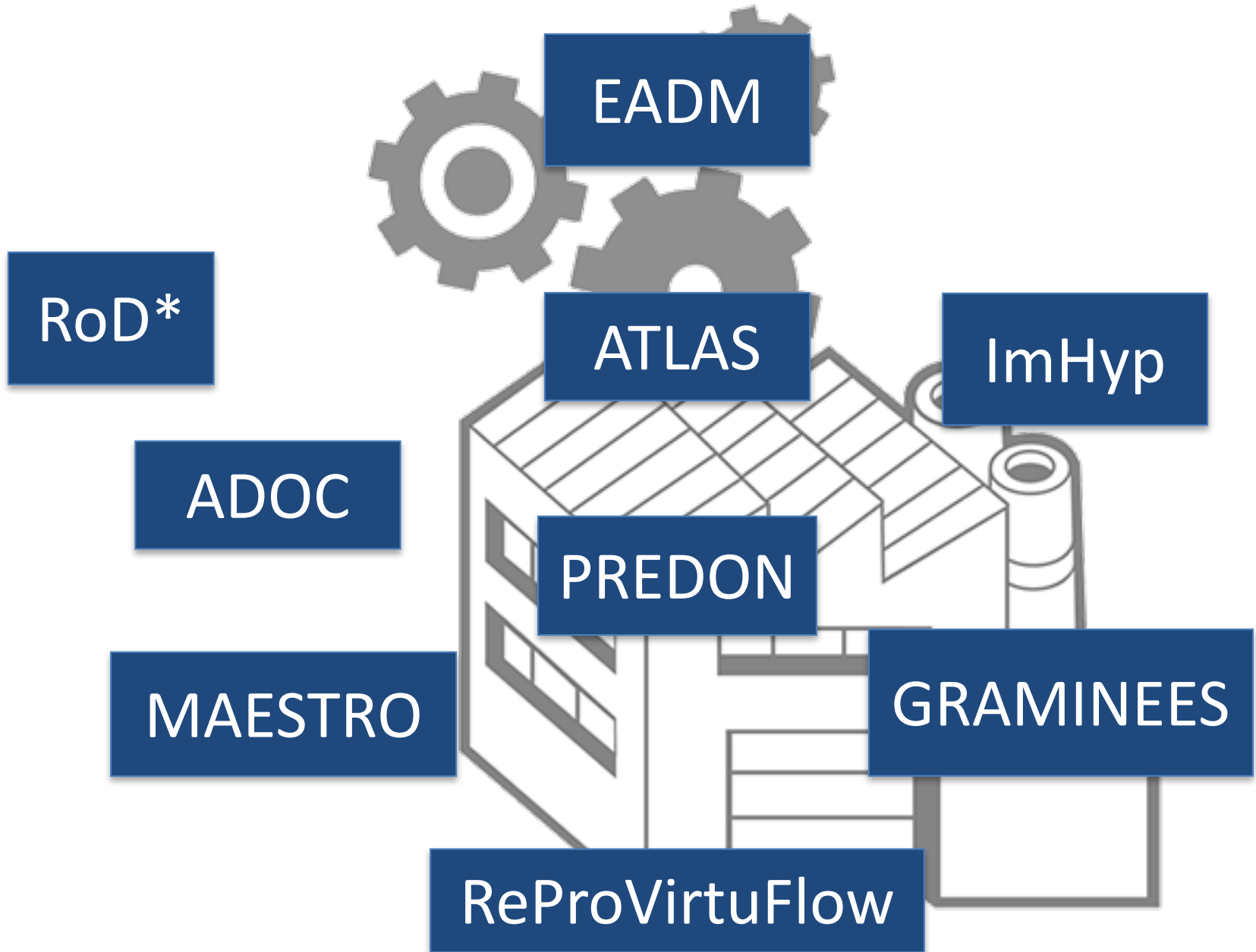
- Nature des données produites et consommées, clairement identifiée ;
- Verrous scientifiques et techniques identifiés
- Durée de vie limitée (2 ans, renouvelable une fois)
- Deux responsables de disciplines différentes

Organisation :

- Le format de fonctionnement d'une action prend la forme qui paraît la plus adéquate aux responsables
- Les Manifestations sont affichées sur le site web MaDICS, et communiquées largement
- Le financement accordé est un montant global, révisable chaque année.
- Un rapport annuel scientifique et financier en fin d'année

Un correspondant attitré au sein de la direction du GDR

Actions en cours



* Action partagée pré-GDR IA

ADOC : ENTrepôt et Analyse de DOcuments

Jérôme Darmont, ERIC, Olivier Teste, INS2i, Natalia Grabar, INSHS/ISCC

ATLAS : Apprentissage, opTimisation Large-échelle et cAlculs distribués

Marianne Clausel, INSMI, Massih Amini, INS2i

EADM : Environmental Acoustic Data mining

Hervé Glotin, LSIS, INS2i, Frédéric Precioso, I3S, INS2i

MAESTRO : MAsses de données En aSTRONomie et astrophysique

Sofian Maabout, INS2i, Christian Surace, INSU

ImHyp : Imagerie Hyperspectrale

Jocelyn Chanusot, INS2i, Mauro Dalla Mura, INS2i, Marie-Françoise Devaux, INRIA

ReProVirtuFlow : Reproductibilité des exp. d'analyse de données

scientifiques - Sarah Cohen-Boulakia, INS2i, Christophe Blanchet, IFB, INSB

PREDON : PREservation des DONnées scientifiques

Cristinel Diaconu, IN2P3, Salima Benbernou, INS2i

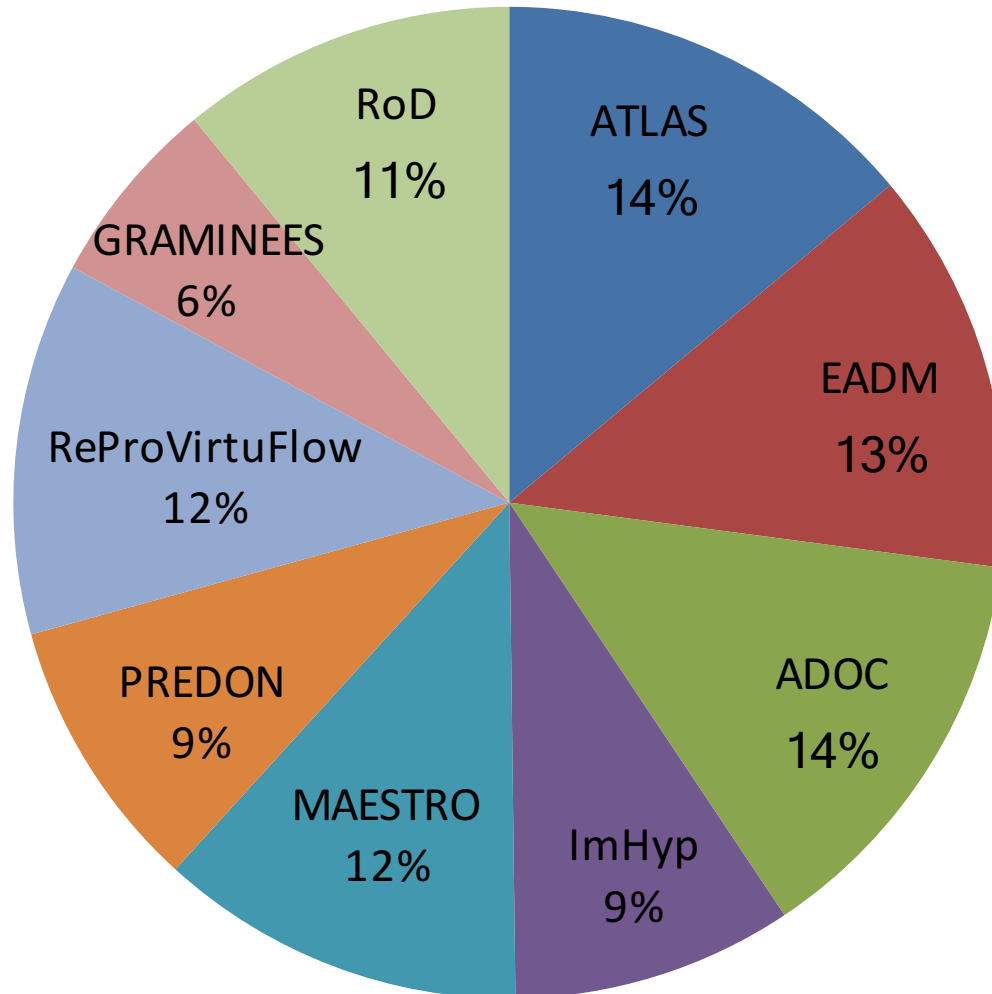
GRAMINEES : GRaph data Mining in Natural, Ecological and Environnemental Sciences - Romain David, IMBE, Dino IENCO, TETIS, IRSTEA

RoD : Raisonner sur les données (Reasoning on Data)

Marie-Laure Mugnier, INS2i, Marie-Christine Rousset, INS2i, Véronique Bellon, IRSTEA / SupAgro, Olivier Palombi, CHU Grenoble-Alpes

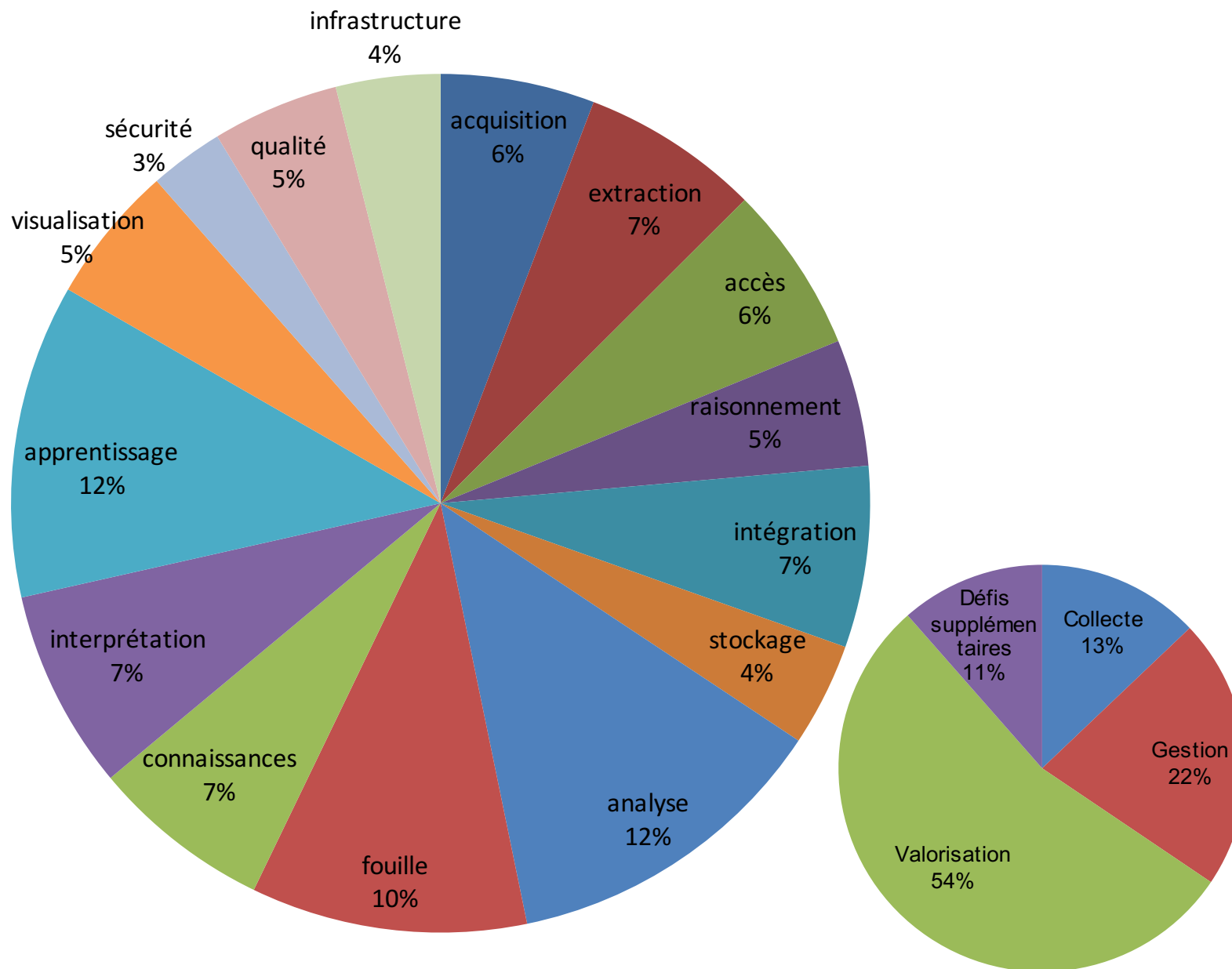
- **Catalyseur** : lancement/organisation de workshops, conférences, challenges, réponses à projets, sondages pour rapprochement de communautés, numéro spéciaux
- **Résultats** : publications, mise à disposition de jeux de données et workflow
- **Visibilité en croissance**
 - rayonnement et communication
 - liens entre Actions, avec d'autres structures

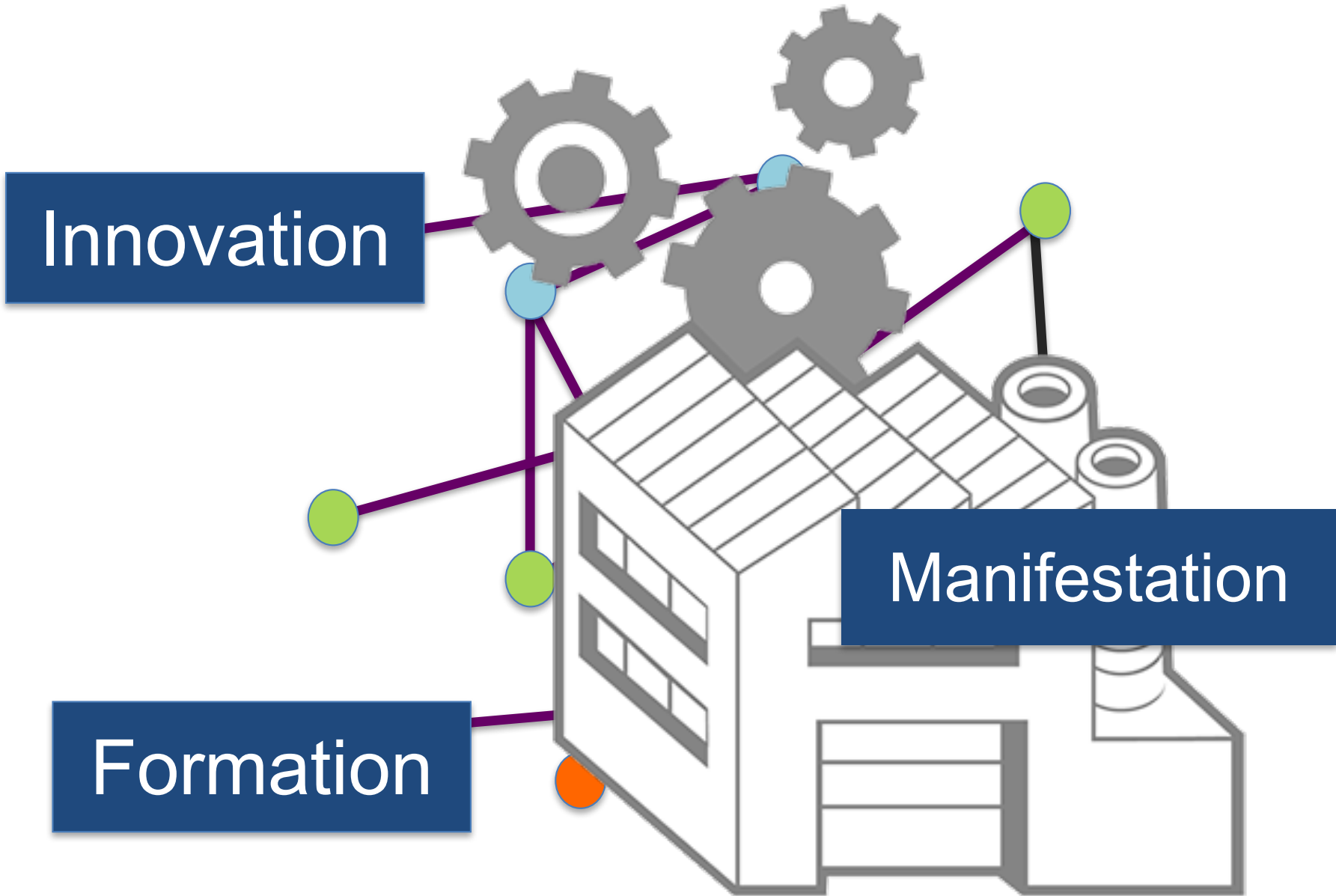
MaDICSiens par Action



55% membres sont rattachés à des Actions

par domaines (mot-clés)





Journée d'animation

exposé tutoriel par des chercheurs confirmés
et exposé plus pointus de doctorants.
Participation d'industriels souhaitée.

Ecole

école thématique/d'été ou articulation avec les écoles
nationales et européennes existantes ;

Groupe de travail, de réflexion

Forum d'échanges

Data fest, Hackathon

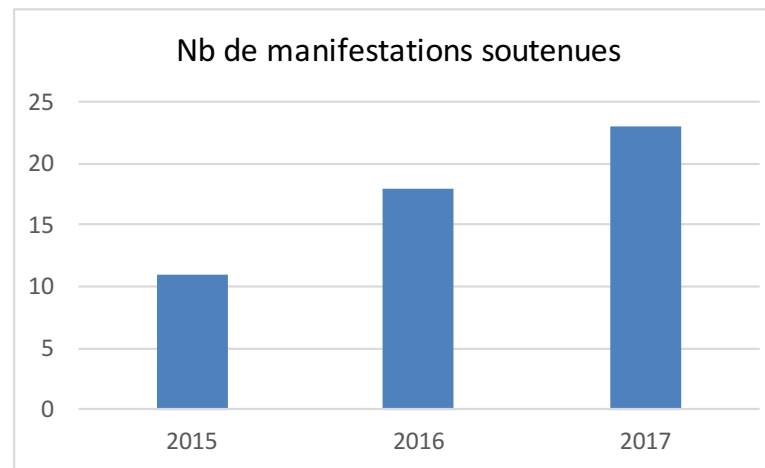
....

des activités au sein des Actions

... hors Actions en lien avec les problématiques MaDICS

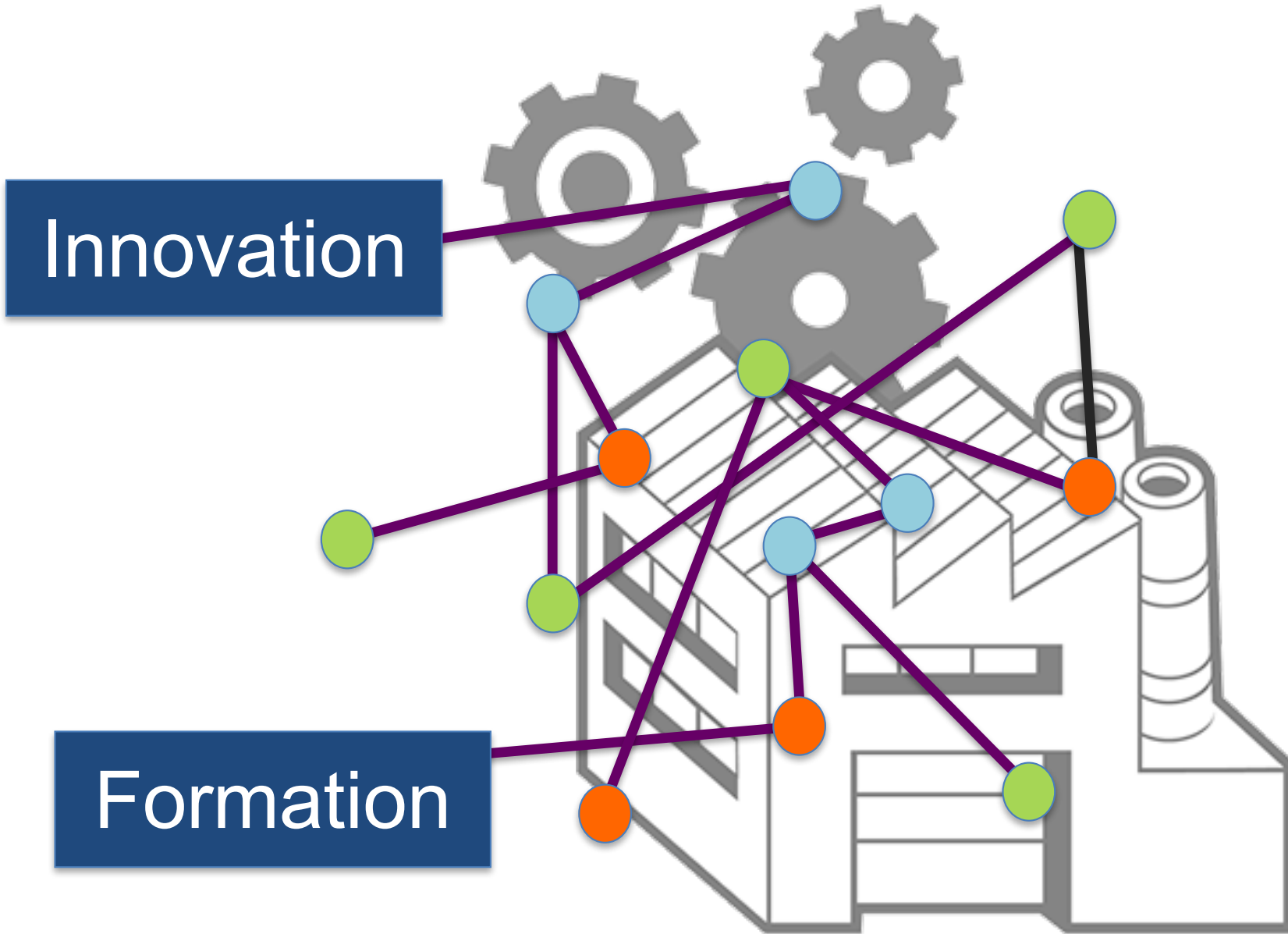
Manifestations MaDICS majeures

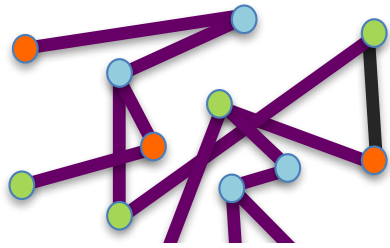
- **Assemblée Constitutive, Lyon, les 24 et 25 juin 2015**
- **Innovatives Big Data (DIRE CNRS), Paris, le 13 octobre 2016**
- **Rencontre annuelle, Paris, le 14 octobre 2016**
- **Journées Scientifiques MaDICS , Marseille, les 22 et 23 juin 2017**
- **Ateliers (montage Actions) / conférences d'ouverture**
 - Qualité des données
 - Masses de Données et santé
 - Infrastructures logicielles multi-échelles
 - Masses de données et intermédiation
 - Bolckchain,
 - Gouvernance des données (questions juridiques)



- Flux significatif de manifestations : 25/35 Manifestations soutenues/labélisées en moyenne par an
- Modèle de soutien de manifestations visant à intégrer au maximum les communautés et à maximiser la visibilité du GDR
- Exigences en cohérence avec ce modèle (transparence, ouverture, faible financement, pas de manifestation récurrente)

Dynamique en progression





- Pour soutenir les **activités de formation** autour des sciences des données et des données massives
 - ressources : liste des Masters, MOOCs, etc.
 - écoles : écoles labellisées MaDICS

En préparation :

- (a) Ecole thématique RESCOM des GDRs MaDICS et RSD (pôle ResCom) sur le thème « **Apprentissage et fouille de données sur les réseaux** »
- (b) Labélisation de l'Ecole thématique BigDataSpeech sur les **Nouvelles Technologies pour l'Exploration de Corpus de Parole**

Pour encourager les **activités des doctorants et jeunes chercheurs** en sciences des données

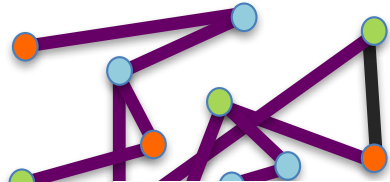
-**aides à la mobilité** : bourses pour favoriser les échanges inter-disciplinaires,

-**aides pour participation à des Manifestations MaDICS** : bourses prenant en charge – partiellement – certains frais

- **posters et démonstrations** (sur appel / rencontre annuelle, manifestation spécifique)

Contrainte : être membre du GDR, appartenir à un laboratoire français

Sélection : adéquation des projets et/ou travaux aux thématiques MaDICS
le nombre de demandes par laboratoire et le budget disponible.



Pour faciliter et intensifier les échanges entre le monde de la recherche académique et les entreprises

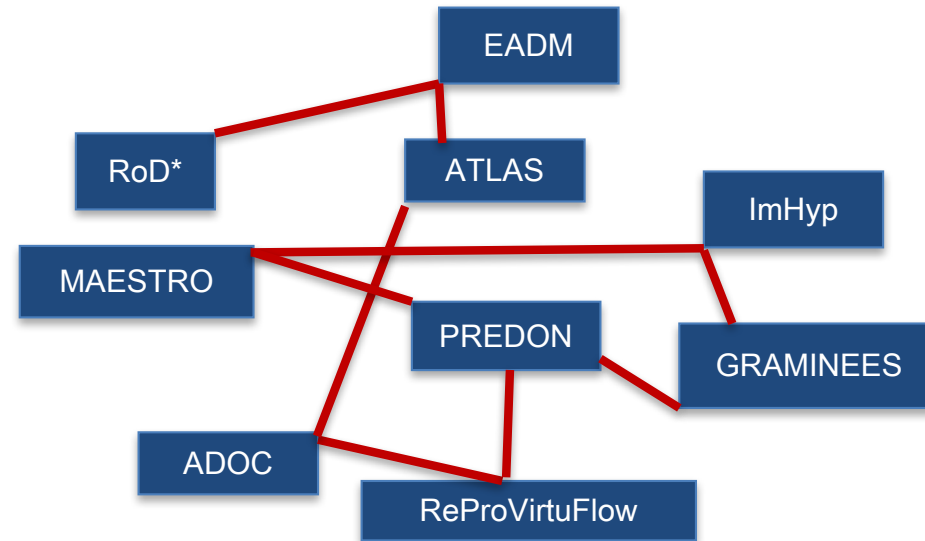
- offrir aux entreprises une vision de compétences académiques
- offrir aux laboratoires un accès à des problématiques industrielles
- mise à disposition de jeux de données, des problèmes
- permettre un accès à l'information, aux résultats

au moins un des responsables du réseau est un industriel

Pour faciliter et intensifier les échanges entre le monde de

- Innovatives Big Data (octobre 2016)
- Convention d'adhésion Club de partenaires en cours de finalisation
- Plateforme d'intermédiation :
 - données ouvertes, privées, logiciels / algo libres
 - dépôt d'annonces,
 - cartographie de compétences

- Internes : entre ACTIONS



- Externes :

- GDR : manifestations communes, liens via les Actions

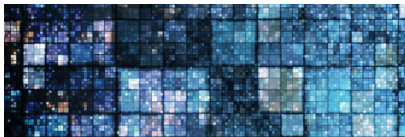


- Centres de recherche / observatoires / grands projets 

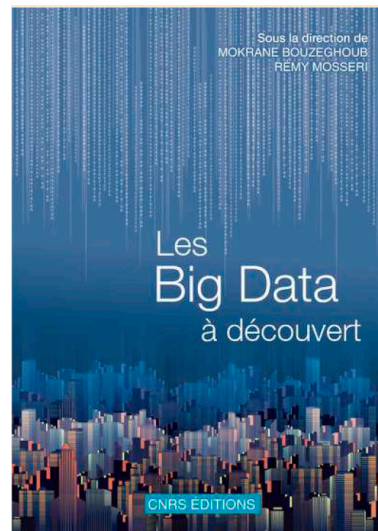


- Grands instruments (TIGIR) / Platerformes (PerSCIDO)
- Associations (AFIA, ...)

- Une dizaine de communications annuelles sur MaDICS ou en rapport avec ses thématiques (sans compter celles faites par les Actions)
- Participation à l'ouvrage collectif et pluridisciplinaire « *Big Data à découvert* »



Livre blanc
Approches contemporaines en
hébergement et gestion de
données



Site Web

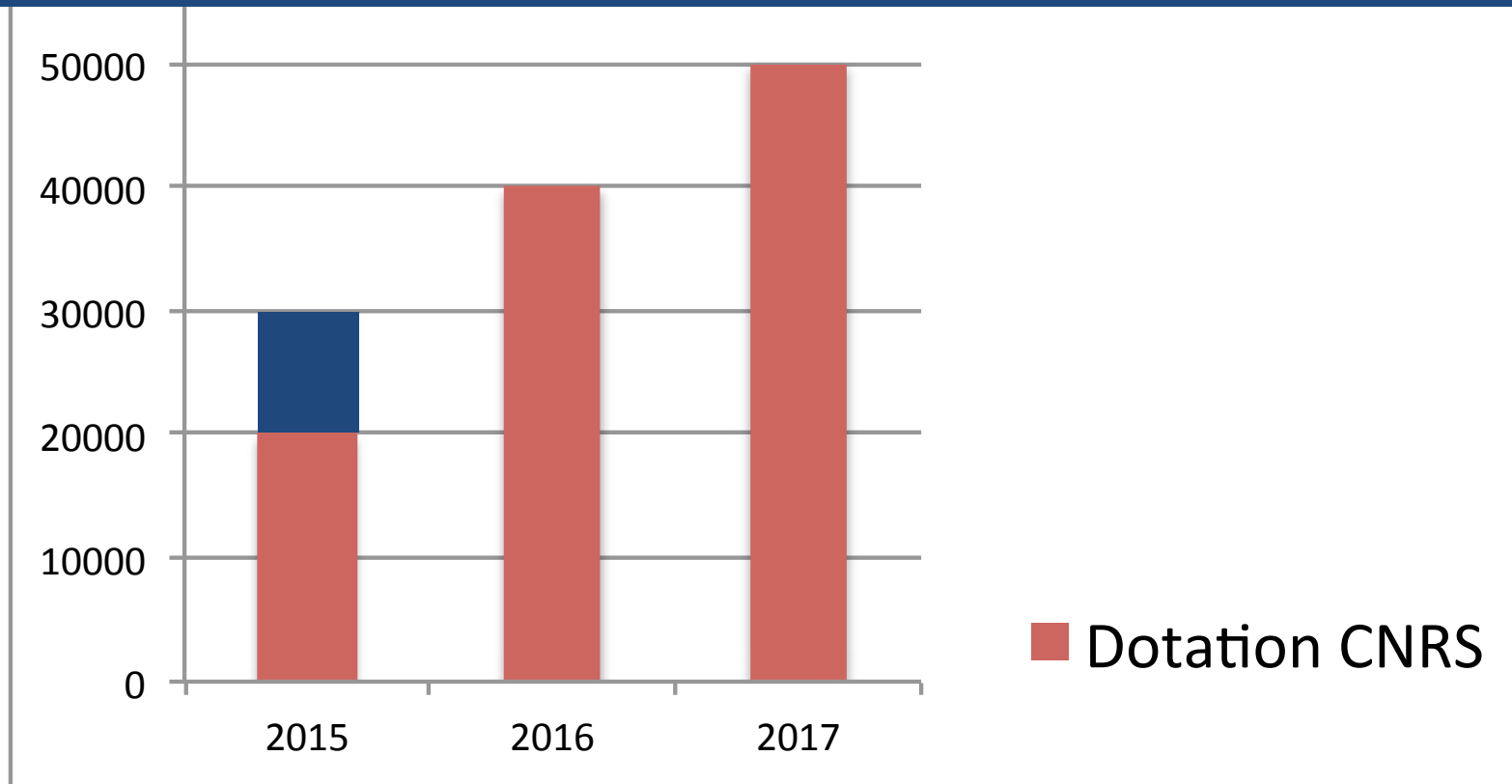
- une gestion des profils des membres leur permettant d'affiner leur profil
- des formulaires de demandes de soutien, d'aides, de diffusion, d'offres d'emploi
- un outil d'aide à la gestion des manifestations pour les organisateurs
- un site web pour chaque Action et son calendrier associé
- des pages de présentation des Manifestations
- des pages de présentation du Réseau Formation et son espace dédié aux doctorants

Associé à

- Listes de diffusion, un bulletin (hebdomadaire)
- Un compte twitter

 [@GDR_MaDICS](https://twitter.com/GDR_MaDICS)

+25%



- **Mise en place de 2 nouvelles Actions en 2017 et poursuite des 7 déjà engagées en 2016**
- **Bonne reconnaissance, dynamique des Actions**
- **Identification d'ateliers pour montage de nouvelles Actions sur des thématiques non couvertes par les Actions**
- **Nb de Manifestations MaDICS en progression ; Rencontre annuelle en perte de vitesse**
- **Collaborations à gérer ; Communications à mieux cibler**
- **Réseau formation en place ; Réseau innovation à implanter**
- **Modèle d'organisation adaptatif**



Merci pour votre attention