



**HAL**  
open science

## La Plateforme universitaire de données d'Aix-Marseille Université (PUD-AMU) et PROGEDO : un dispositif national d'accompagnement au cœur des universités.

Clément De Belsunce

### ► To cite this version:

Clément De Belsunce. La Plateforme universitaire de données d'Aix-Marseille Université (PUD-AMU) et PROGEDO : un dispositif national d'accompagnement au cœur des universités.. De la multiplicité d'acteurs au guichet unique ?, Mar 2022, Aix-en-Provence, France. hal-03671595

**HAL Id: hal-03671595**

**<https://hal.science/hal-03671595>**

Submitted on 18 May 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Maison Méditerranéenne  
des Sciences de l'Homme  
Aix-Marseille Université

**De la multiplicité d'acteurs au guichet unique ?**

**La PUD-AMU et PROGEDO : un dispositif national  
d'accompagnement au cœur des universités**

**Clément De Belsunce, IE PUD-AMU**

# IR\* PROGEDO

- Née il y a 10 ans du constat qu'il existait de nombreuses données et enquêtes SHS, mais très largement sous-exploitées par la communauté des chercheurs.
- On est arrivé à la situation où parfois les utilisateurs ne connaissent pas les données existantes
- PROGEDO a été créé pour pallier ce constat en développant la culture des données en SHS en France

→ **Mettre en place la politique des données quantitatives et issues d'enquêtes en sciences humaines et sociales en France**

# OUVRIR LA SCIENCE !

*Politique Française des données en SHS*

**HN** Huma-Num  
la TGR des humanités numériques

**OpenEdition**

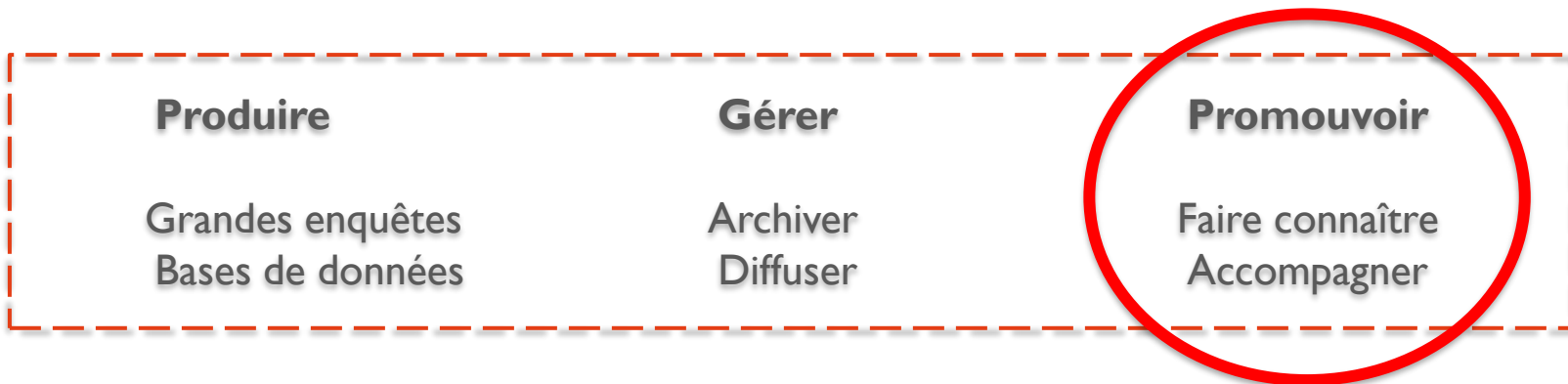
**Rn MSH**  
Réseau national

 European Commission  
**ERIC**  
European Research  
Infrastructure Consortium

**QUETELET  
PROGEDO**  
DIFFUSION 

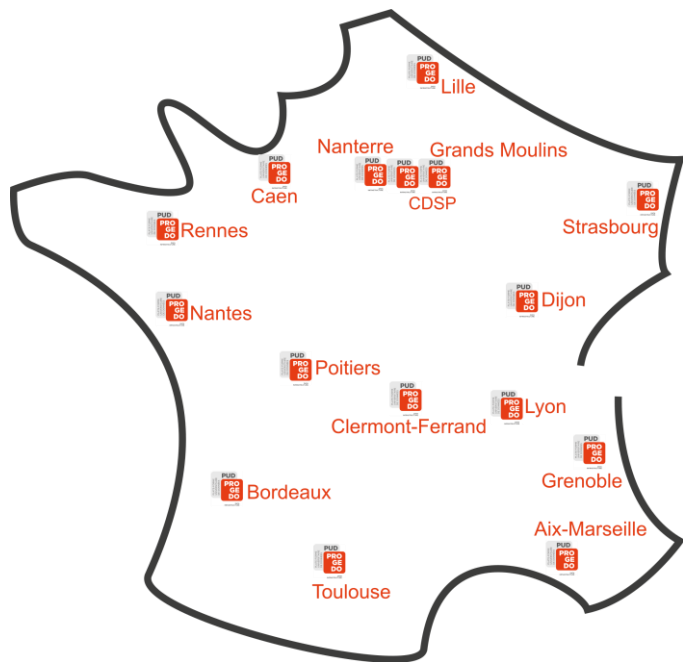
**MATE-SHS**

# IR\* PROGEDO



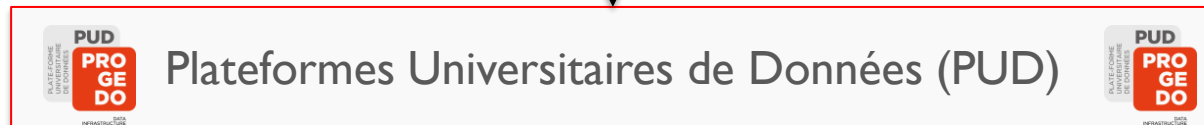
- Assurer la production des volets français des grandes enquêtes internationales
- Assurer la gestion des enquêtes quantitatives en SHS
- Assurer l'archivage et la diffusion des données issues de la statistique publique
- Assurer l'accès aux bases de données quantitatives en SHS
- Assurer la promotion des données quantitatives en SHS pour développer leur utilisation

# TGIR PROGEDO



**PRO**duction  
et **GE**stion  
des **DO**nnées

DATA  
INFRASTRUCTURE



# PUD-AMU

(1<sup>er</sup> avril 2018)

## Référente scientifique

Isabelle Blöss-Widmer

## Ingénieurs d'études

Clément De Belsunce

 [pud-amu@univ-amu.fr](mailto:pud-amu@univ-amu.fr)

 <https://pud.mmsh.univ-aix.fr/>



## Localisation

Maison Méditerranéenne des Sciences de l'Homme  
Bureau 107, bâtiment C.

# La PUD-AMU au service de quels publics ?

CHERCHEURS  
ENSEIGNANTS-CHERCHEURS

Articles, montage de projets ...  
Fournir des jeux de données  
quantitatives supports aux cours ...

PERSONNELS ADMINISTRATIFS  
ET TECHNIQUES

Formations ...

DOCTORANTS

Thèse, articles, formations ...

ETUDIANTS DE MASTER

Mémoires, projet de thèse ...



# La PUD-AMU, un guichet unique ?

- **Guichet unique : Oui**
- **Mais on parlera surtout de point d'entrée !**
  
- **Au sein des MSH, au plus proche des laboratoires.**
- **Un réseau national dans le périmètre local universitaire pour développer au mieux la culture des données en SHS.**

# Les missions de la PUD-AMU

- **Accompagnement individuel**
- Accompagnement collectif
- Promotions des bonnes pratiques autour des données en SHS
- Organisation de manifestations scientifiques

# Accompagnement individuel

## Pour connaître les données existantes

- Identification des bases : choix des sources, des instituts diffuseurs/producteurs
- Données existantes en relation avec un thème de recherche
- Comprendre les métadonnées



## Pour accéder aux données



## Pour traiter les données



## Pour interpréter les résultats



# Plateforme Universitaire de Données - AMU



# Où et comment chercher ?



Documents ▾



Fr ☰

## Mobilité professionnelle Fr

### Fiche du sujet

Référentiel

Rameau

Identifiant

http://data.bnf.fr/ark:/12148/cb11950828g

### Synonyme

Occupational mobility Job mobility

Mobility Occupational Movilidad laboral

Movilidad de mano de obra

Movilidad ocupacional Movilidad profesional

Trabajadores-Movilidad

### Discipline

Sociologie (139) - Economies et finances (96) - Histoire (43) - Droit (35) - Education (26) - Géographie (22) - Gestion et management (16) - Psychologie (8) - Science politique (6) - Anthropologie sociale et ethnologie (6) - Musique, musicologie et arts de la scène (3) - Linguistique (3) - Sciences de l'information et de la communication (3) - Architecture, aménagement de l'espace (2) - Démographie (1)

### Type de document

Articles (250) - Données d'enquêtes (141) - Livres et chapitres d'ouvrages (95) - Textes imprimés (68) - Mémoires, Thèses et HDR (50) - Colloques et conférences (37) - Autres (24) - Documents audiovisuels (12) - Billets de blog (12) - Recensions (10) - Rapports (9) - Page Web (4) - Prépublication (2) - Séminaires (1) - Présentations (1) - Périodiques (1) - Photos et images (1)

### Auteur

Institut National de la Statistique et des Études

672 documents

- Les pulsations de la mobilité en entreprise. Entre reconfiguration d'un marché interne et constitution d'un régime d'épreuve à la mobilité professionnelle** Textes imprimés   
Christophe Claisse (3 janv. 2)
- Mobilité sociale : enquête formation qualification professionnelle de 1985 / Ministère de l'économie, des finances et de la privatisation, Institut national de la statistique et des études économiques ; [réd. par] Michel Gollac, Pierre Laulhé, Jeanine Soleilhavoup** Textes imprimés   
Institut national de la statistique et des études économiques (France). Auteur du texte (1988)
- Enquête Emploi en continu - 2016** Données d'enquêtes   
Institut National de la Statistique et des Études Économiques (10 avr. 2018)
- L'effet de l'âge sur la distribution des salaires** Articles   
Sarah Le Duigou (2012)
- L'adaptation de la relation de travail pour motif économique** Mémoires, Thèses et HDR   
Pierre Lopes (9 déc. 2017)
- Développer l'employabilité et favoriser la mobilité professionnelle par l'étude et la prise en compte du contexte de travail** Colloques et conférences   
Anne Pignault et al. (janv. 2005)
- Enquête Emploi - 1968** Données d'enquêtes   
Institut National de la Statistique et des Études Économiques (9 avr. 2018)
- Navette domicile-travail à grande vitesse : situation d'exception, arrangement traditionnel** Articles   
Xavière Lanéelle (2006)
- Se rencontrer à la frontière : marchands et inquisition dans l'Italie** Livres et chapitres d'ouvrages

# Accompagnement individuel

## Pour connaître les données existantes

- Identification des bases : choix des sources, des instituts diffuseurs/producteurs
- Données existantes en relation avec un thème de recherche
- Comprendre les métadonnées



## Pour accéder aux données

- Exploiter les nouveaux outils numériques de recherche de données
- Démarches/procédures d'accès



## Pour traiter les données



## Pour interpréter les résultats



# Comment accéder aux données ?



ADISP Commandier des données

DESCRIPTION TABLEAU ANALYSE

ADISP

- Éducation et formation
- Travail et emploi
  - Enquêtes Emploi
    - Série 1962-1967
    - Série 1968-1974
    - Série 1975-1981
    - Série 1982-1989
    - Série 1990-2002
    - Enquête Emploi en continu 2003-2016 - version standard
    - Enquête Emploi en continu 2003-2018 - version FPR
      - Enquête Emploi en continu (version FPR) - 2003
      - Enquête Emploi en continu (version FPR) - 2004
      - Enquête Emploi en continu (version FPR) - 2005
      - Enquête Emploi en continu (version FPR) - 2006
      - Enquête Emploi en continu (version FPR) - 2007
      - Enquête Emploi en continu (version FPR) - 2008
      - Enquête Emploi en continu (version FPR) - 2009
      - Enquête Emploi en continu (version FPR) - 2010
      - Enquête Emploi en continu (version FPR) - 2011
      - Enquête Emploi en continu (version FPR) - 2012
      - Enquête Emploi en continu (version FPR) - 2013
      - Enquête Emploi en continu (version FPR) - 2014
      - Enquête Emploi en continu (version FPR) - 2015
      - Enquête Emploi en continu (version FPR) - 2016
      - Enquête Emploi en continu (version FPR) - 2017
      - Enquête Emploi en continu (version FPR) - 2018
    - Métadonnées
    - Description des variables
      - Base individus
        - Caractéristiques individuelles
          - Age détaillé en années révolues au 31 décembre
          - Age en années révolues au 31 décembre (3 postes, premier type de regroupement)
          - Age en années révolues au 31 décembre (3 postes, second type de regroupement)

Jeux de données: Enquête Emploi en continu (version FPR) - 2018

**Variable AG: Age détaillé en années révolues au 31 décembre**

Valeurs	Catégories	N	
100	100 ans	29	0,0%
101	101 ans	30	0,0%
102	102 ans	20	0,0%
103	103 ans	20	0,0%
104	104 ans	11	0,0%
105	105 ans	4	0,0%
106	106 ans	2	0,0%
107		4	0,0%
15	15 ans	3553	0,8%
16	16 ans	6788	1,6%
17	17 ans	6920	1,6%
18	18 ans	7034	1,7%
19	19 ans	6267	1,5%
20	20 ans	6035	1,4%
21	21 ans	5516	1,3%
22	22 ans	5072	1,2%
23	23 ans	4779	1,1%
24	24 ans	4463	1,1%
25	25 ans	4198	1,0%
26	26 ans	4476	1,1%
27	27 ans	4470	1,1%
28	28 ans	4645	1,1%
29	29 ans	4623	1,1%
30	30 ans	4611	1,1%
31	31 ans	4915	1,2%
32	32 ans	5338	1,3%
33	33 ans	5248	1,2%

# Comment accéder aux données ?

The screenshot displays the 'Réseau Quetelet' search interface. The search term 'mobilité' is entered in the search bar. The results are sorted by 'score de pertinence' and show a list of questions and variables related to mobility. The interface includes filters for 'Producteur', 'Série d'enquêtes', and 'Enquête'. The search results are displayed in a list format, with each entry showing the question text, the number of modalities, and the source of the data.

Produit par : <http://bdq.quetelet.progedo.fr>



# Accompagnement individuel

## Pour connaître les données existantes

- Identification des bases : choix des sources, des instituts diffuseurs/producteurs
- Données existantes en relation avec un thème de recherche
- Comprendre les métadonnées



## Pour accéder aux données

- Exploiter les nouveaux outils numériques de recherche de données
- Démarches/procédures d'accès

## Pour traiter les données

- Choix des méthodes adaptées
- Utilisation des logiciels informatiques pour le traitement des données



## Pour interpréter les résultats

# Quelle méthode choisir ?

## Choisir une méthode d'analyse multivariée

Les différentes méthodes d'analyse multivariée permettent de répondre à des problématiques variées. Le choix d'une méthode dépend de l'objectif initial, des types de variables manipulées mais aussi, de la forme des résultats obtenus qui peuvent être plus ou moins faciles à présenter et à expliquer.

<b>OBJECTIF</b>	Résumer l'information en minimisant la déperdition et repérer des dimensions "cachées"	Variables Numériques et/ou ordinales	→	<b>Analyse en Composantes Principales (ACP)</b>
		Deux variables qualitatives	→	<b>Analyse Factorielle des Correspondances (AFC)</b>
		Trois variables qualitatives ou plus	→	<b>Analyse des Correspondances Multiples (ACM)</b>
	Constituer des groupes d'individus aussi similaires que possible	Le nombre de groupes est fixé	→	<b>Analyse typologique</b>
		Le nombre de groupes n'est pas fixé	→	<b>Classification hiérarchique</b>
	Expliquer une variable par plusieurs autres	Variable à expliquer numérique ou ordinale	→	<b>Régression multiple</b>
		Variable à expliquer qualitative et variables explicatives numériques	→	<b>Analyse discriminante</b>
		Variable à expliquer qualitative et variables explicatives qualitatives	→	<b>Segmentation</b>

# Quelles fonctions utiliser ?

```

1 install.packages("tidyverse")
2 install.packages("haven")
3 install.packages("FactoMineR")
4 install.packages("missmda")
5 install.packages("Factoshiny")
6 install.packages("FactoInvestigate")
7 install.packages("openxlsx")
8
9 library(tidyverse)
10 library(haven)
11 library(FactoMineR)
12 library(missmda)
13 library(Factoshiny)
14 library(FactoInvestigate)
15 library(openxlsx)
16
17 setwd("C:/Users/debe/sunce/Downloads")
18
19 donnee1 <- read.xlsx("milieu.xlsx", sheet = 1, startRow = 1, colNames = TRUE,
20                   rowNames = FALSE, detectDates = FALSE, skipEmptyRows = TRUE,
21                   skipEmptyCols = TRUE, rows = NULL, cols = NULL, check.names = FALSE,
22                   namedRegion = NULL, na.strings = "NA", fillMergedCells = FALSE)
23
24 res.mca1 = MCA(donnee1)
25
26 Investigate(res.mca1)
27
28 ## le fichier origine ne marche pas##
29
30 donnee2 <- read.xlsx("origine2.xlsx")
31 res.mca2 = MCA(donnee2)
32 plot.MCA(res.mca2)
33 Investigate(res.mca2)
34
35 #####
36
37
38 donnee3 <- read.xlsx("ambiance.xlsx", sheet = 1, startRow = 1, colNames = TRUE,
39                   rowNames = FALSE, detectDates = FALSE, skipEmptyRows = TRUE,
40                   skipEmptyCols = TRUE, rows = NULL, cols = NULL, check.names = FALSE,
41                   namedRegion = NULL, na.strings = "NA", fillMergedCells = FALSE)
42
43 res.mca3 = MCA(donnee3)
44
45 Investigate(res.mca3)

```

2423 | (Top Level) | R Script

Console Terminal Jobs

```

C:/Users/debe/sunce/Downloads/
> setwd("C:/Users/debe/sunce/Downloads")
>
+ donnee1 <- read.xlsx("milieu.xlsx", sheet = 1, startRow = 1, colNames = TRUE,
+                   rowNames = FALSE, detectDates = FALSE, skipEmptyRows = TRUE,
+                   skipEmptyCols = TRUE, rows = NULL, cols = NULL, check.names = FALSE,
+                   namedRegion = NULL, na.strings = "NA", fillMergedCells = FALSE)
+
+ res.mca1 = MCA(donnee1)
+ |

```

# Accompagnement individuel

## Pour connaître les données existantes

- Identification des bases : choix des sources, des instituts diffuseurs/producteurs
- Données existantes en relation avec un thème de recherche
- Comprendre les métadonnées



## Pour accéder aux données

- Exploiter les nouveaux outils numériques de recherche de données
- Démarches/procédures d'accès

## Pour traiter les données

- Choix des méthodes adaptées
- Utilisation des logiciels informatiques pour le traitement des données



## Pour interpréter les résultats

- Interprétations des résultats pour répondre aux questions de recherche
- Valorisation graphique et mise en forme des résultats

# Comment interpréter ses résultats ?

## 2. Distribution de l'inertie

L'inertie des axes factoriels indique d'une part si les variables sont structurées et suggère d'autre part le nombre judicieux de composantes principales à étudier.

Les 2 premiers axes de l'analyse expriment **13.66%** de l'inertie totale du jeu de données ; cela signifie que 13.66% de la variabilité totale du nuage des individus (ou des variables) est représentée dans ce plan. C'est un pourcentage très faible, et le premier plan ne représente donc seulement qu'une part de la variabilité contenue dans l'ensemble du jeu de données actif. Cette valeur est supérieure à la valeur référence de **10.45%**, la variabilité expliquée par ce plan est donc significative (cette inertie de référence est le quantile 0.95 de la distribution des pourcentages d'inertie obtenus en simulant 680 jeux de données aléatoires de dimensions comparables sur la base d'une distribution uniforme).

Du fait de ces observations, il serait alors certainement nécessaire de considérer également les dimensions supérieures ou égales à la troisième dans l'analyse.

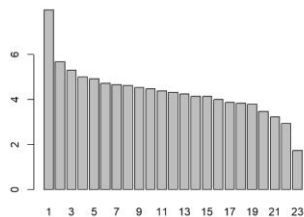


Figure 2 - Decomposition of the total inertia on the components of the analyse

Une estimation du nombre pertinent d'axes à interpréter suggère de restreindre l'analyse à la description des 5 premiers axes. Ces composantes révèlent un taux d'inertie supérieur à celle du quantile 0.95 de distributions aléatoires (28.86% contre 25.19%). Cette observation suggère que seuls ces axes sont porteurs d'une véritable information. En conséquence, la description de l'analyse sera restreinte à ces seuls axes.

## 3. Description du plan 1:2

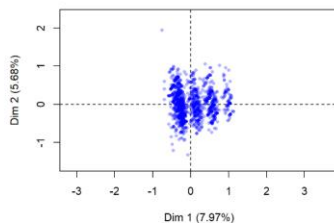
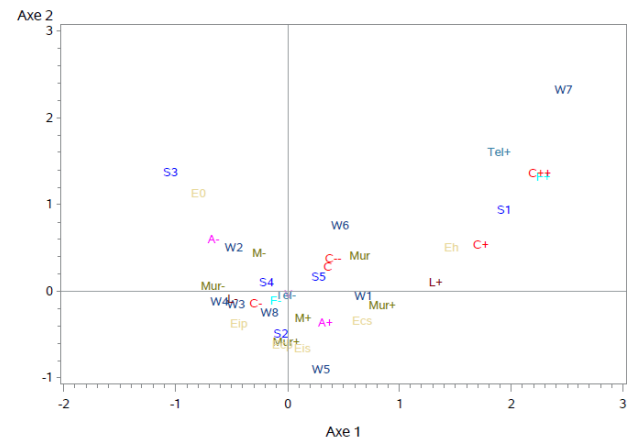


Figure 3.1 - Graphe des individus (ACM) Les individus libellés sont ceux ayant la plus grande contribution à la construction du plan.



# Les missions de la PUD-AMU

- Accompagnement individuel
- **Accompagnement collectif**
- Promotions des bonnes pratiques autour des données en SHS
- Organisation de manifestations scientifiques

# Accompagnement collectif

- **Interventions :**
  - Outils d'aide à la recherche de données
  - Initiation au traitement statistique
  - Initiation à la gestion d'une base de données
- **Formations :**
  - Logiciel de traitement statistique R (3 niveaux)
  - Logiciel tableur Excel (2 niveaux)
- **Projets**
- **Collaboration avec les groupes de travail des laboratoires.**

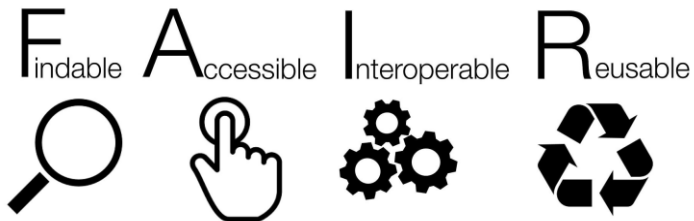
# Les missions de la PUD-AMU

- Accompagnement individuel
- Accompagnement collectif
- **Promotions des bonnes pratiques autour des données en SHS**
- Organisation de manifestations scientifiques



# Promotions des bonnes pratiques autour des données en SHS

- Ouvrir la science
- **Relais informatif dans le domaine des données**
- Sensibilisation à la culture des données
- **Aiguillage vers les services traitant des sujets précis (PGD, archivage, stockage...)**



# Les missions de la PUD-AMU

- Accompagnement individuel
- Accompagnement collectif
- Promotions des bonnes pratiques autour des données en SHS
- **Organisation de manifestations scientifiques**

# Organisation de manifestations scientifiques

- **Semaine Data SHS** : 05 au 09 décembre 2022
- **Conférence + Master Class**
- **Montre-moi ta donnée**
- **Présentation de ressources (enquêtes, cours, logiciels...)**
- **Interventions de spécialistes sur un sujet demandé**



Maison Méditerranéenne  
des Sciences de l'Homme

Aix★Marseille Université

**Merci de votre attention !**

**[pud-amu@univ-amu.fr](mailto:pud-amu@univ-amu.fr)**

**<http://pud.mmsh.univ-aix.fr>**