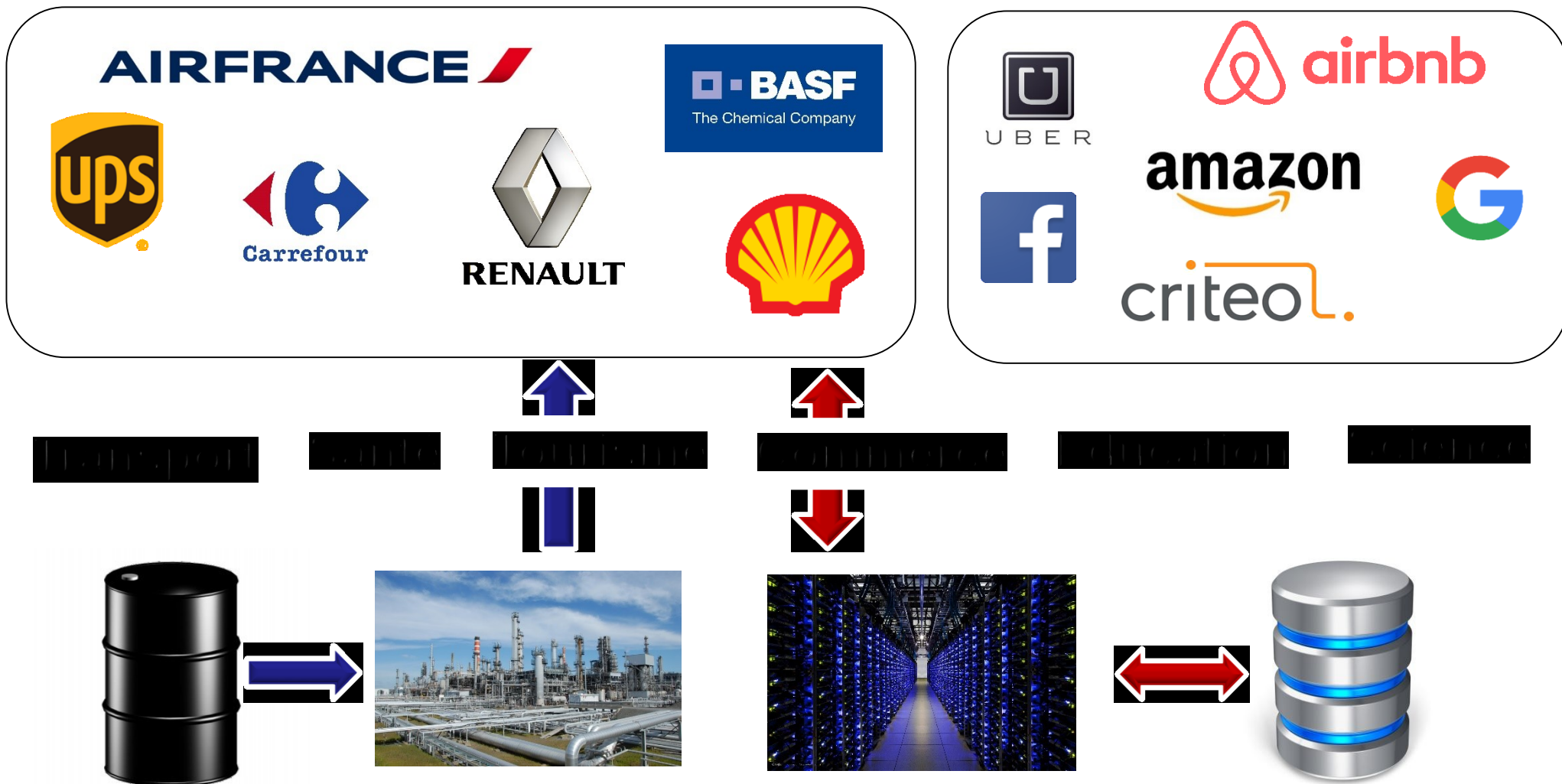




# Données = Pétrole du 21e siècle

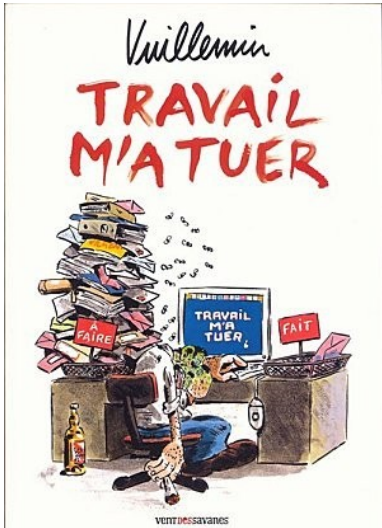


# Données – D trole du 21e si cle

## Ce que vous allez en faire:

Recommandations de produits, publicit ,  
Int gration GAFAM,  
R duction de la consommation d' nergie,  
D tection de fraudes,  
IA robotique, R volution du jeu de Go vid o,  
Optimisation de m dicaments,  
D tection de galaxies, ...

# Il faut aimer



un peu

- les maths : probas, stats, logique, ...

Compute the standard deviation for 15, 15, 19, 21, 13, 13  
1st compute mean,  $\bar{x} = \frac{\sum x}{n} = \frac{15+15+19+21+13+13}{6} = 16$   
2nd Remember the formula for standard dev.  
 $S = \sqrt{\frac{\sum (x-\bar{x})^2}{n-1}}$   
3rd Lay out table using order of operation

x	x - $\bar{x}$	(x - $\bar{x}$ ) <sup>2</sup>
13	13 - 16 = -3	(-3) <sup>2</sup> = 9
13	13 - 16 = -3	(-3) <sup>2</sup> = 9
15	15 - 16 = -1	(-1) <sup>2</sup> = 1
15	15 - 16 = -1	(-1) <sup>2</sup> = 1
19	19 - 16 = 3	(3) <sup>2</sup> = 9
21	21 - 16 = 5	(5) <sup>2</sup> = 25
$\sum (x - \bar{x}) = 0$ always 0		$\sum (x - \bar{x})^2 = 54$

$S = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}} = \sqrt{\frac{54}{5}} = \sqrt{10.80} = 3.29$   
3.286

à la folie

- apprendre
- travailler
- réfléchir

beaucoup

- programmer
- modéliser

```
def add(x):  
    return x+5  
  
def dotwrite(mat):  
    nodename = getNodeName()  
    label=symbol.sym_name.get(id(mat[0]),mat[0])  
    print ' %s [label=%s' % (nodename, label),  
    if isinstance(mat[1], str):  
        if mat[1].strip():  
            print '%s]' % mat[1]  
    else:  
        print '']  
    else:  
        print '[]'  
    children = []  
    for i in range(len(mat[1])):  
        children.append(dotwrite(child))  
    print ' %s -> {' % nodename  
    for i in range(len(children)):  
        print '%s' % children[i]
```

# DAC : Profils de compétence

## Tronc commun (S5)

Modèles avancés de bases de données (MLBDA)

Logique et représentation des connaissances (LRC)

Modèles et Algo Proba/Stats pour l'Informatique (MAPSI)

## Apprentissage

Machine Learning (ML)

Recherche d'Information et Traitement Automatique du Langage Naturel (RITAL)

Advanced MACHine Learning and Deep Learning (AMAL)

Reinforcement Learning and advanced Deep Learning (RLD)

Research in Data Science and Methodology (REDS)

## Intelligence Artificielle

Intelligence Artificielle et Manipulation Symbolique de l'Information (IAMSI)

Recherche d'Information et Traitement Automatique du Langage Naturel (RITAL)

Linked Open Data, Apprentissage Symbolique (LODAS)

Computational Intelligence (CI)

## Bases de Données

Stockage et Accès aux Mégadonnées (SAM)

Business Intelligence and User Modelling (BIUM)

Linked Open Data, Apprentissage Symbolique (LODAS)

Bases de Données Large-Echelle (BDLE)

# Science des données et des connaissances

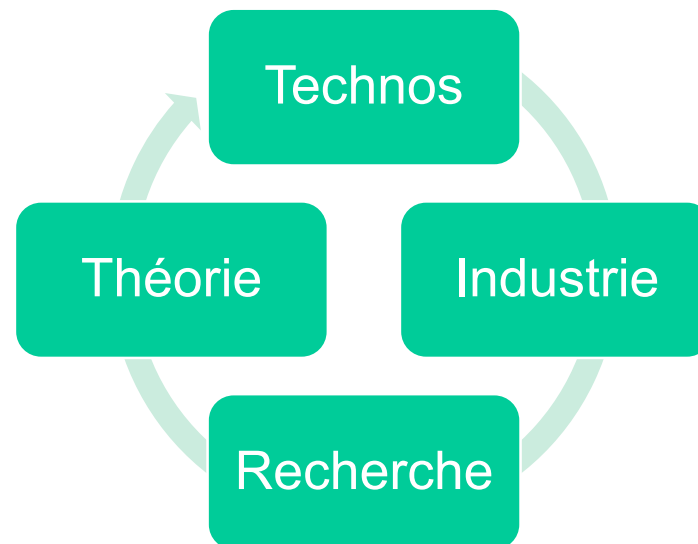
## Domaines

- BD Large-Echelle, noSQL, Map-Reduce,
- Apprentissage, Deep Learning,
- Web Sémantique, Intelligence Computationnelle, ...

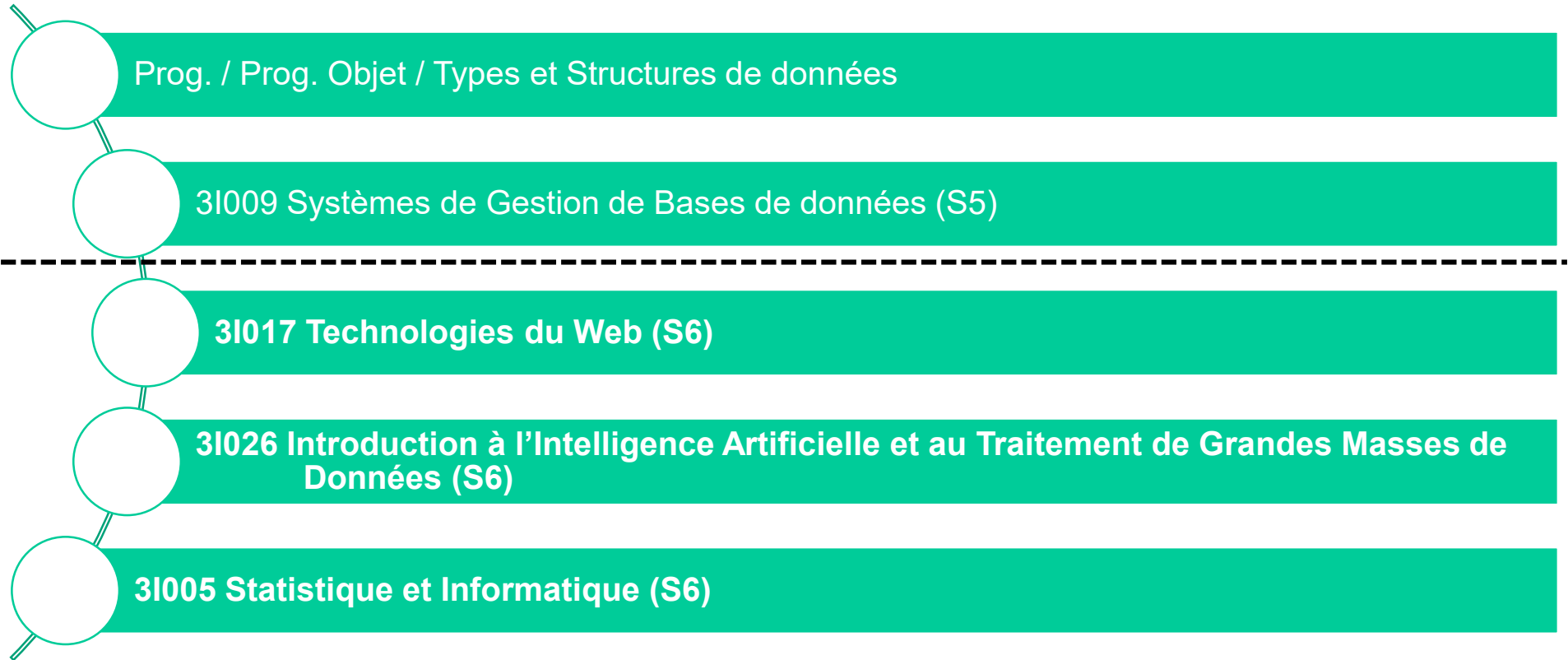


## Compétences

- Formalisation de problèmes
- Identification de solutions
- Choix des technologies
- Réalisation de logiciels



# UEs de « découverte » en L3



# DAC est aussi

Intervenants et  
séminaires  
industriels

Connections  
avec différents  
« Meetup »

Participations  
aux challenges  
(Kaggle, ...)

Réseau d'industrie  
et d'anciens  
étudiants

Enseignants de  
plusieurs équipes de  
recherche du LIP6

<http://dac.lip6.fr/master>