

Projet de Développement Logiciel

(Master 1 – MIAGE)

<https://github.com/acherm/PDL2021>

Mathieu Acher

@acherm

Maître de Conférences

mathieu.acher@irisa.fr

PDL (avec le COVID)

- 100% distanciel
 - salles de TP disponibles
- Teams
 - Cours/Questions/Diffusion de messages informatifs
- <https://github.com/acherm/PDL2021>
 - Slides, ressources, consignes: à consulter régulièrement
- Le développement logiciel en vrai
 - De plus en plus distribué et à distance
 - Plein d'outils pour supporter le travail à distance
- **COMMUNIQUER**
 - plus que jamais: eg si problème de matériel ou d'organisation
 - avec le client ou via votre groupe de projet

PDL: objectifs pédagogiques

- Pratique et (re-)visite de votre **cursus**
 - Modélisation/UML, Programmation OO, test, design patterns, etc
 - Outils: git, IDE, Maven, intégration continue, documentation
 - Méthodes: travail en groupe, dates limites
- Une **expérience** de la difficulté du développement logiciel
 - indispensable pour votre future vie professionnelle
- **Contribution** à un projet réel, open source

Projet

- “Polygot Wikipedia Matrix Extractor”
- Reprise d’un projet de l’année dernière
 - chaque groupe reprendra le git/Github d’un groupe de 2018-2019 (qui lui-même avait repris un projet de 2017-2018)
 - le comprendre, le faire fonctionner, le ré-utiliser, l’améliorer au niveau de la documentation, du code, et de la suite de tests
- Ecrire une “meilleure” solution dans un autre langage
 - A priori Python ou JavaScript
- Ecrire des outils pour fabriquer des jeux de données de test



WIKIPEDIA
The free encyclopedia

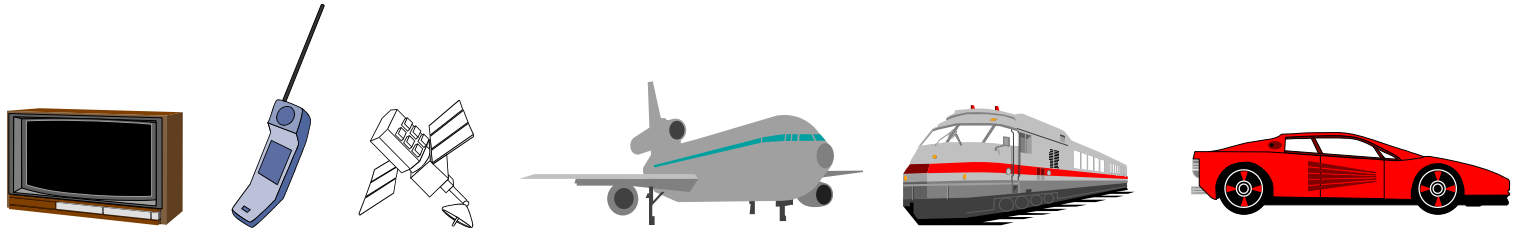
A screenshot of a Wikipedia table, likely from the 'List of countries by area' article. The table has multiple columns and rows, with some cells containing text and others containing numbers or symbols. The table is partially obscured by the Wikipedia logo and text.

Développer du logiciel \sim =

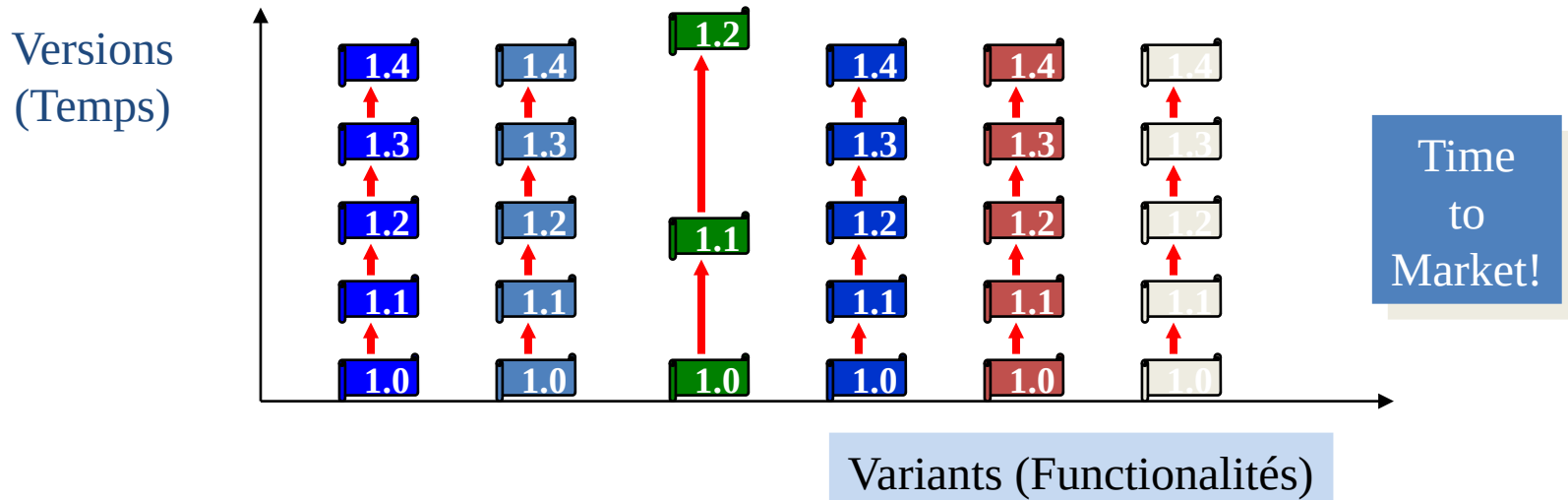
« **Multi-Person Construction
Of
Multi-Versions Programs** »

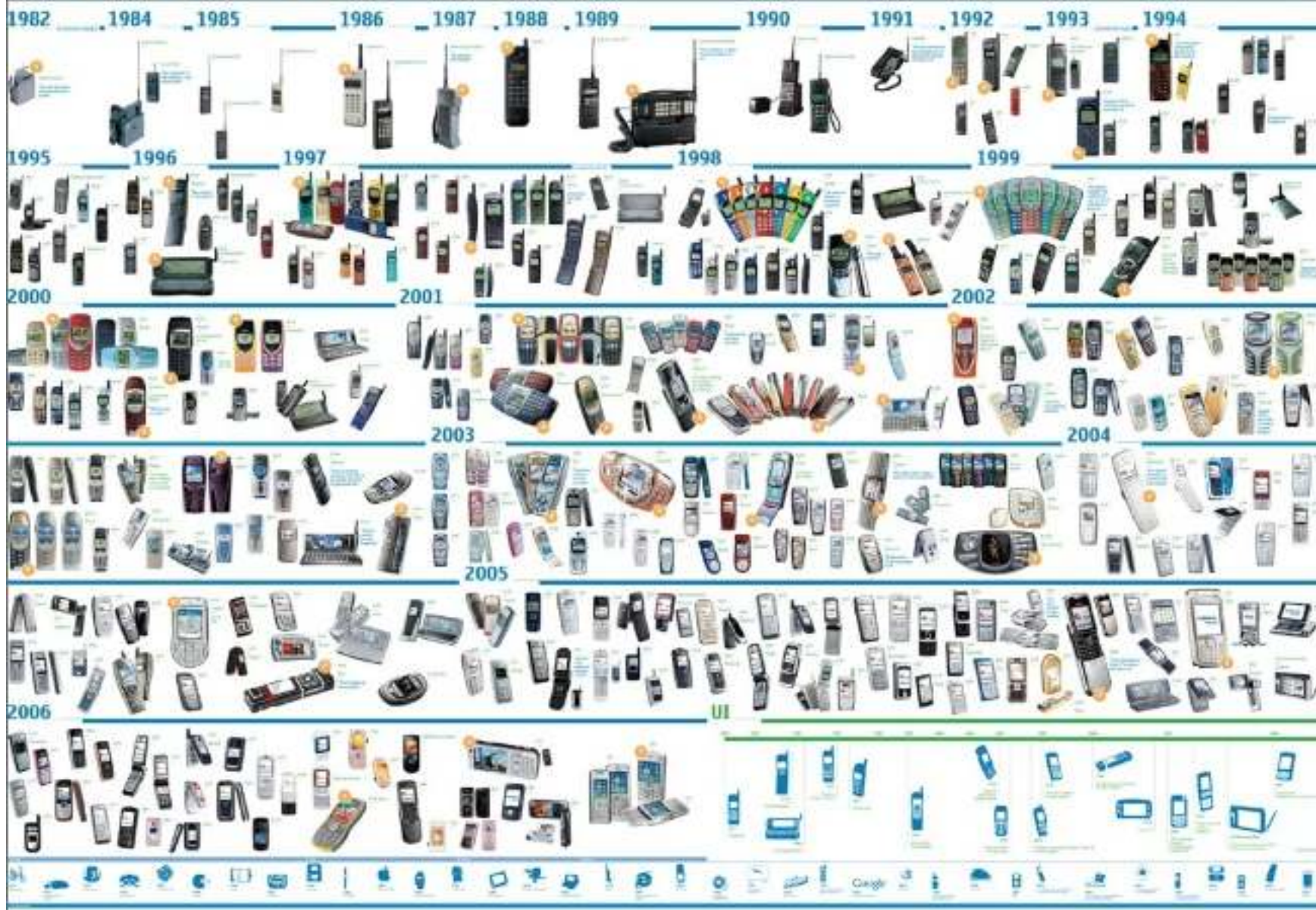
David Parnas, 2014

Ingénierie du logiciel



- De plus en plus complexe
 - Systèmes distribués
 - Qualité de service: performance, sécurité, sûreté, utilisabilité, etc.
- Explosion des fonctionnalités
 - Lignes de produits (espace/temps)





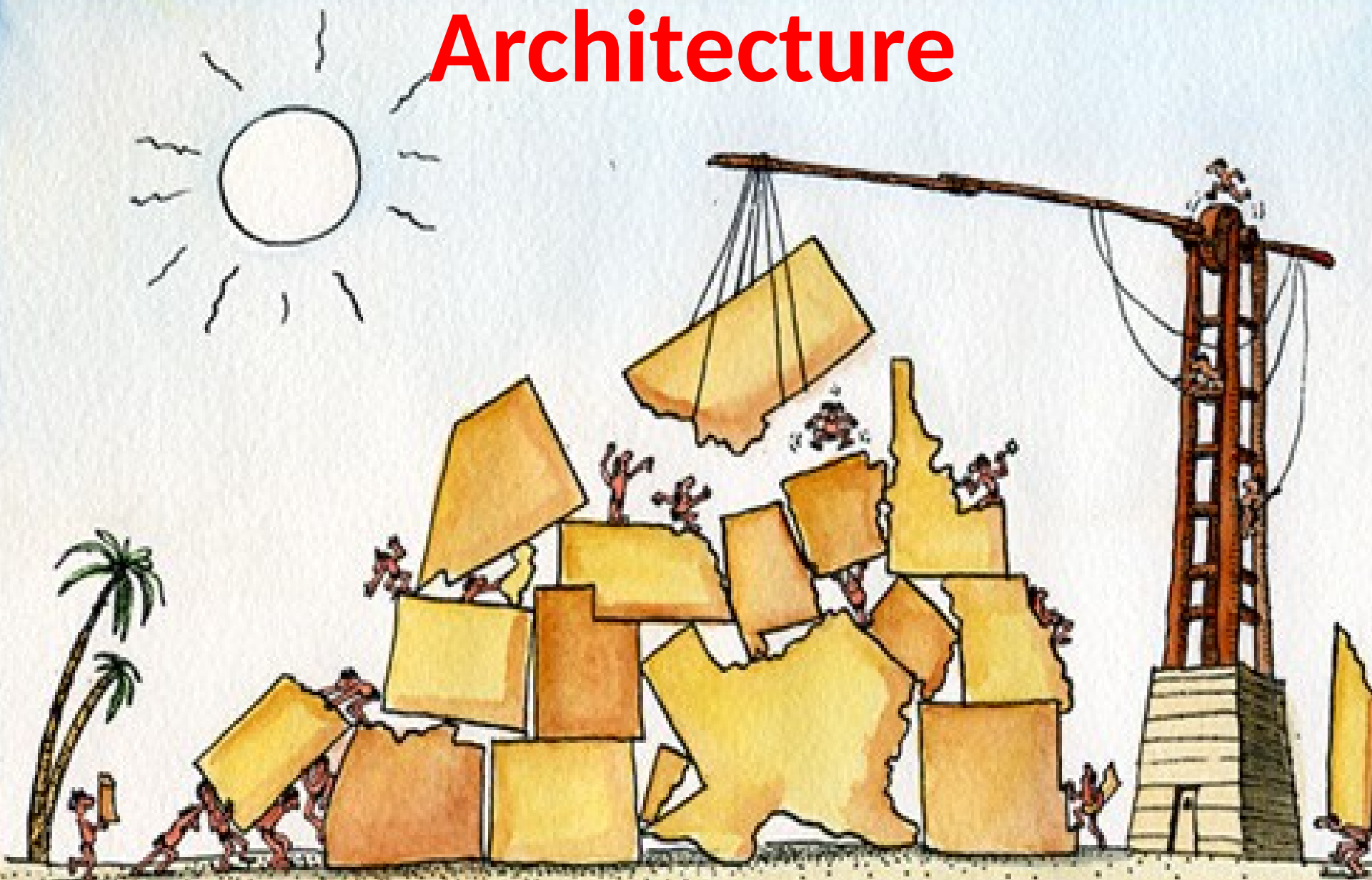


Travail d'équipe

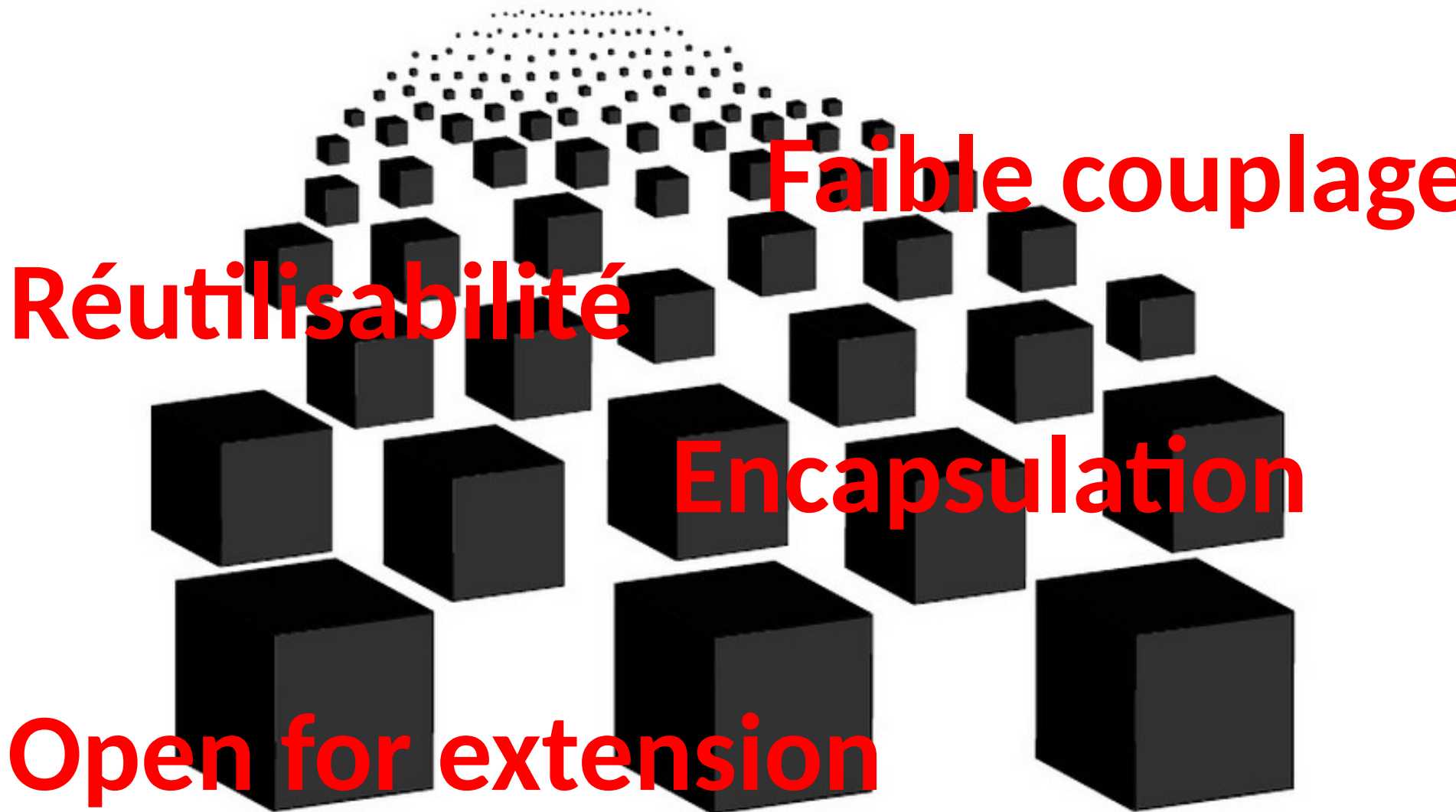
- Organisation
 - Partage des tâches
 - Planification
 - Communication
- Code idéalement...
 - Bien conçu, modulaire, documenté
 - Maintenable, compréhensible
 - Testable
- Outils
 - Collaboratifs (e.g., système de versions)



Architecture

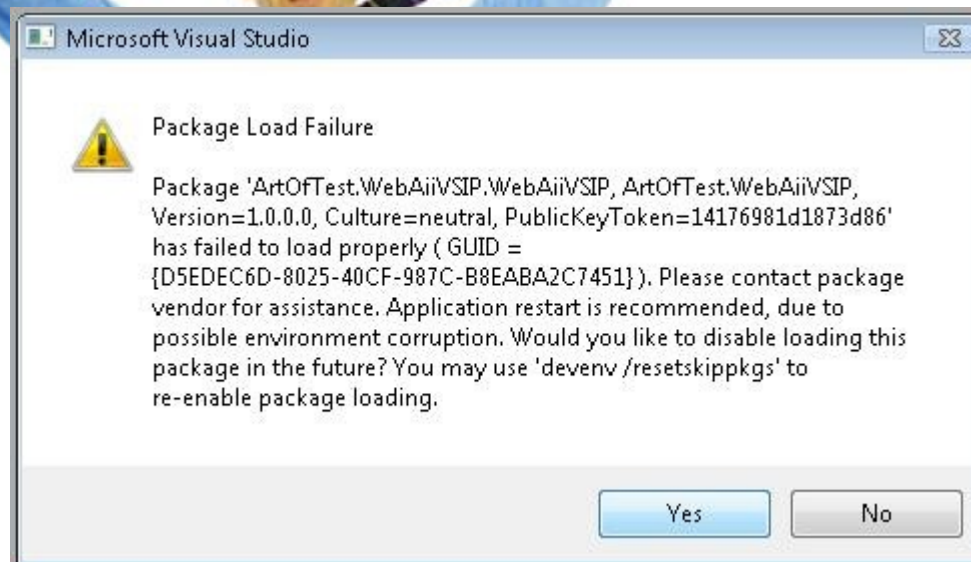
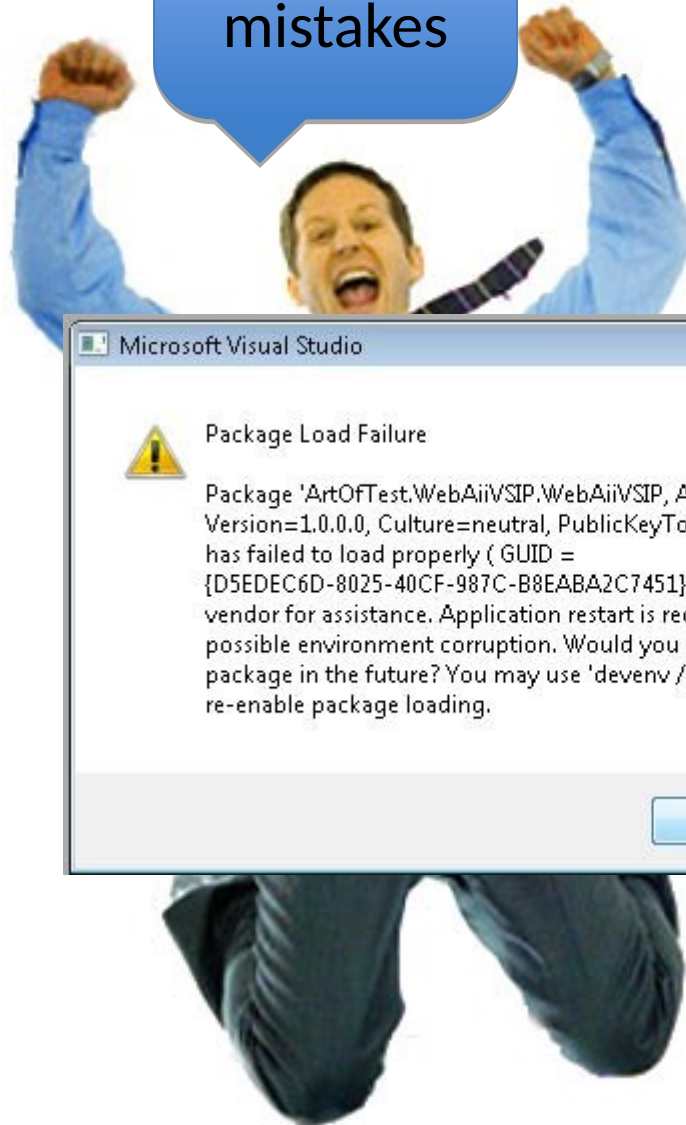


Idéalement:
« modular black boxes »

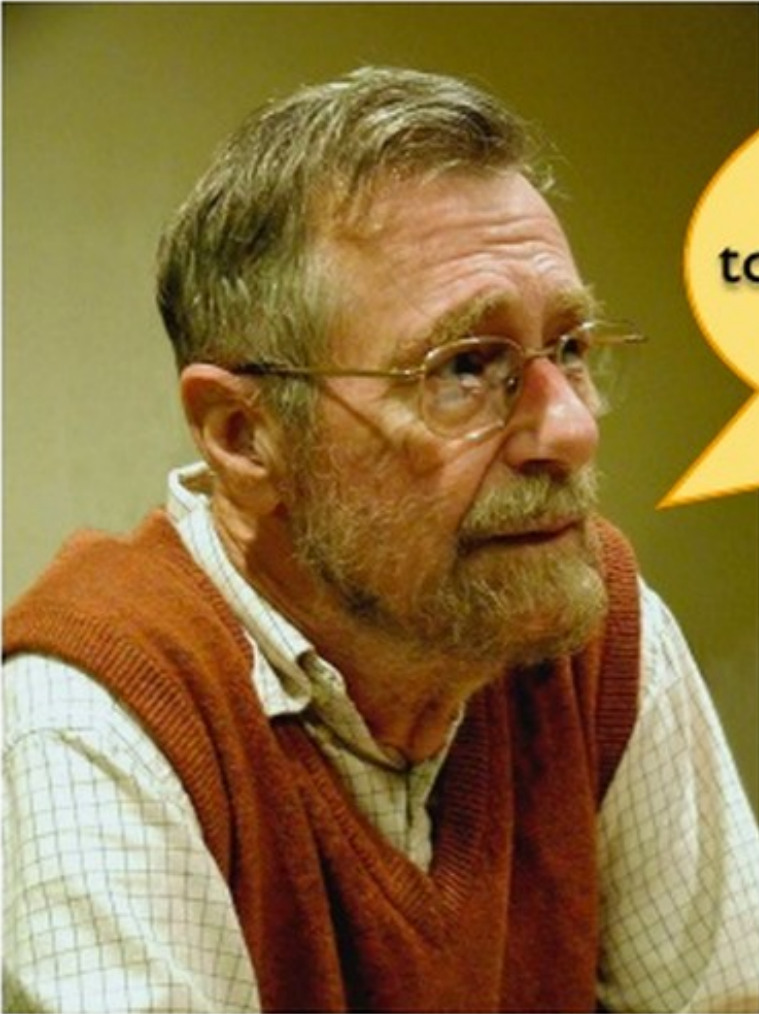


I don't
make
mistakes

Testing

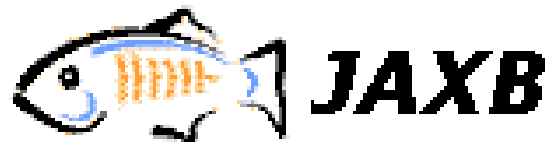


Dijkstra



Program testing can be used
to show the presence of bugs, but
never to show their absence!

Software Integration



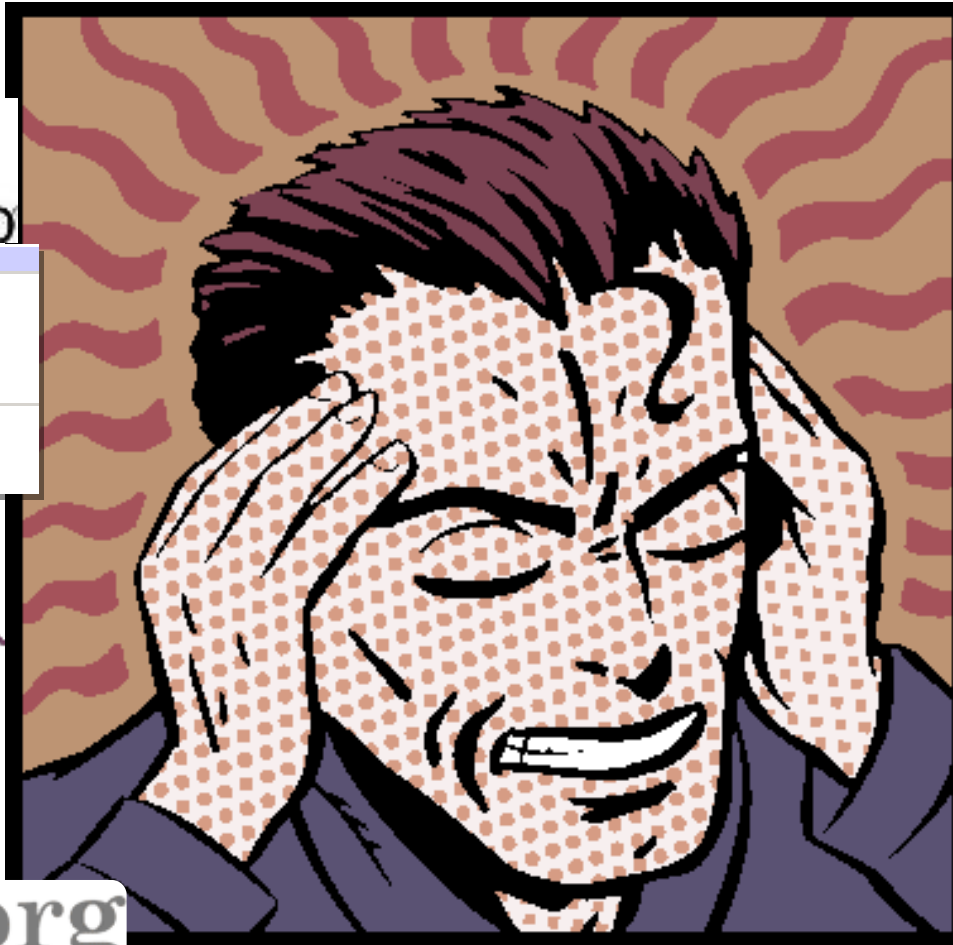
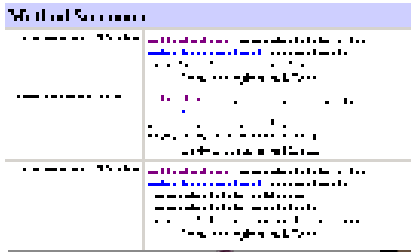
google-guice

Guice (pronounced 'juice') is a lightweight dependency injection framework for Java 5 and above, brought to you by Google.



PostgreSQL

Développement Logiciel





How the customer explained it



How the Project Leader understood it



How the Analyst designed it



How the Programmer wrote it



How the Business Consultant described it



How the project was documented



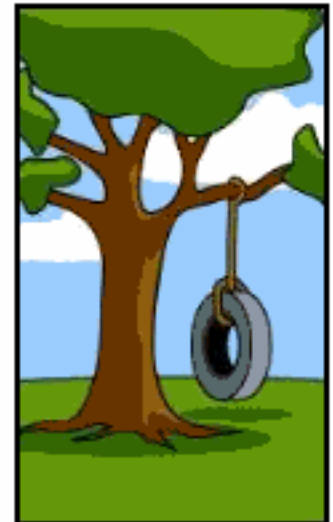
What operations installed



How the customer was billed



How it was supported



What the customer really needed

PDL: Objectifs

- Analyse, conception, réalisation, test, par la pratique
 - (Re)visite de votre cursus (UML, Programmation OO, etc.)
- Gestion de projets
 - Sur un exemple « joué » mais bien réel où des résultats sont attendus
 - Projet en groupe
- Préparation pour votre future vie professionnelle
 - Capacité à rédiger des spécifications
 - Capacité à appréhender de nouvelles technologies
 - Comprendre la difficulté du génie logiciel





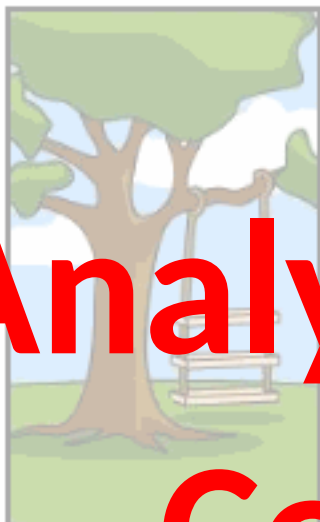
PDL en pratique?

Analyse

Conception

Réalisation

Validation



How the customer explained it



How the project manager understood it



How the Analyst designed it



How the Programmer wrote it



How the Business Consultant described it



How the project was documented



What operations installed



How the customer was billed



How it was supported



What the customer really needed

4 rendus

- (EX) Eliciter des exigences, comprendre un projet, éditer des documents techniques
- (SP1+SP2) Implémenter dans un laps de temps prédéfini
- (PR) Présentation pendant 30' (20' + 10'): expliquer et défendre un travail, synthèse, bilan

Le projet



Classament	País	Comptador	Punts	Preferències	A. Força	MOB	Key	MOB	Key	MOB	Key
1			1118	1118	1118	1118	1118	1118	1118	1118	1118
2			1069	1069	1069	1069	1069	1069	1069	1069	1069
3			1005	1005	1005	1005	1005	1005	1005	1005	1005
4			1000	1000	1000	1000	1000	1000	1000	1000	1000
5			1000	1000	1000	1000	1000	1000	1000	1000	1000
6			1000	1000	1000	1000	1000	1000	1000	1000	1000
7			1000	1000	1000	1000	1000	1000	1000	1000	1000
8			1000	1000	1000	1000	1000	1000	1000	1000	1000
9			1000	1000	1000	1000	1000	1000	1000	1000	1000



Model	Image processor	Sensor format						
D3X	EXPEED	Full-frame	CMOS	Sony	24.5	51	1005	
D2Xs	-	APS-C	CMOS	Sony	12.4	11	1005	
D2X	-	APS-C	CMOS	Sony	12.4			
D1X	-	APS-C	CCD	Sony	5.3			
D1	-	APS-C	CCD	Sony	2.66			



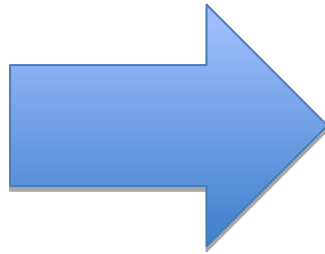
D5	E			
D4S	E			
D4	E			
D3S	E			
D3	E			
D2Hs				

Model	Image processor	Sensor format						
D5	EXPEED	Full-frame	CMOS	Sony	24.5	51	1005	
D4S	-	APS-C	CMOS	Sony	12.4	11	1005	
D4	-	APS-C	CMOS	Sony	12.4			
D3S	-	APS-C	CCD	Sony	5.3			
D3	-	APS-C	CCD	Sony	2.66			

51	1005
11	1005



WikipediaMatrix



CSV
(Comma Separated
Values)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

WikipediaMatrix

https://en.wikipedia.org/wiki/Comparison_of_Canon_EOS_digital_cameras

Article Talk Read Edit View history Search Wikipedia



Wiki Loves Monuments: Photograph a monument, help Wikipedia and win!



Comparison of Canon EOS digital cameras

From Wikipedia, the free encyclopedia

The following tables provide general information as well as a comparison of technical specifications for a number of [Canon EOS digital cameras](#).

General information [[edit](#)]

Model	Image processor	Sensor format	Megapixels	Min ISO	Max ISO	Autofocus points	Viewfinder magnification, coverage	Display size, dots (ratio)	Touch screen	Live view	Max FPS	Storage	Release date	Weight (kg)	Dimensions, WxHxD (mm)	Video	Main Battery
1Ds	DIGIC	Full-frame CMOS	11.4	100	1250	45	0.75x, 100%	2.0" 120k	No	No	3	CF	2002Q4	1.265	156 x 158 x 80	-	NP-E
1Ds Mk II	DIGIC II	Full-frame CMOS	16.7	50	3200	45	0.75x, 100%	2.0" 230k	No	No	4.5	CF, SD	2004Q4	1.215	156 x 158 x 80	-	NP-E
1Ds Mk III	Dual DIGIC III	Full-frame CMOS	21.1	50	3200	45	0.75x, 100%	3.0" 230k	No	Yes	5.0	CF, SD	2007Q4	1.210	156 x 160 x 80	-	LP-E
1D	DIGIC	APS-H CCD	4	100	3200	45	0.72x, 100%	2.0" 120k	No	No	8.0	CF	2001Q4	1.250	156 x 158 x 80	-	NP-E
1D Mk II	DIGIC II	APS-H CMOS	8.2	50	3200	45	0.72x, 100%	2.0" 230k	No	No	8.5	CF, SD	2004Q2	1.220	156 x 158 x 80	-	NP-E
1D Mk II N	DIGIC II	APS-H CMOS	8.2	50	3200	45	0.72x, 100%	2.5" 230k	No	No	8.5	CF, SD	2005Q3	1.225	156 x 158 x 80	-	NP-E
1D Mk III	Dual DIGIC III	APS-H CMOS	10.1	50	6400	45	0.75x, 100%	3.0" 230k	No	Yes	10	CF, SD	2007Q1	1.155	156 x 157 x 80	-	LP-E
1D Mk IV	Dual DIGIC 4	APS-H CMOS	16.1	50	102,400	45	0.75x, 100%	3.0" 820k	No	Yes	10	CF, SD	2008Q4	1.160	156 x 157 x 80	1080p30	LP-E

WikipediaMatrix

https://en.wikipedia.org/wiki/Comparison_of_Canon_EOS_digital_cameras

Comparison of Canon EOS digital cameras

Model

Model	Image sensor	Image processor	Sensor format	Megapixels	Min ISO	Max ISO	Autofocus points	Viewfinder	magnification,	coverage	Display	size, dots	(ratio)	Touch screen	Live view	Max FPS	Storage
EOS 5D Mark IV	30.5 MP	DIGIC 6	Full frame	30.5	100	51200	45	Optical	0.76x	100%	3.2"	2359k	3:2	Yes	Yes	12	CFexpress, SD
EOS 5D Mark III	22.3 MP	DIGIC 5+	Full frame	22.3	100	12800	39	Optical	0.7x	100%	3.2"	2304k	3:2	Yes	Yes	10	CFexpress, SD
EOS 5D Mark II	21 MP	DIGIC 5	Full frame	21	100	12800	39	Optical	0.7x	100%	3.2"	2304k	3:2	Yes	Yes	10	CFexpress, SD
EOS 5D	12.8 MP	DIGIC 4	Full frame	12.8	100	12800	39	Optical	0.7x	100%	3.2"	2304k	3:2	Yes	Yes	10	CFexpress, SD
EOS 5D Mark I	9.9 MP	DIGIC 2	Full frame	9.9	100	12800	39	Optical	0.7x	100%	3.2"	2304k	3:2	Yes	Yes	10	CFexpress, SD

```
==General information==
{{ class="wikitable sortable" style="text-align: center; font-size: 85%; width: auto;"
|
| Model
| Image processor
| Sensor format
| Megapixels
| Min ISO
| Max ISO
| Autofocus points
| Viewfinder
magnification,
coverage
| Display
size, dots
(ratio)
| Touch
screen
| Live view
| Max FPS
| Storage
```

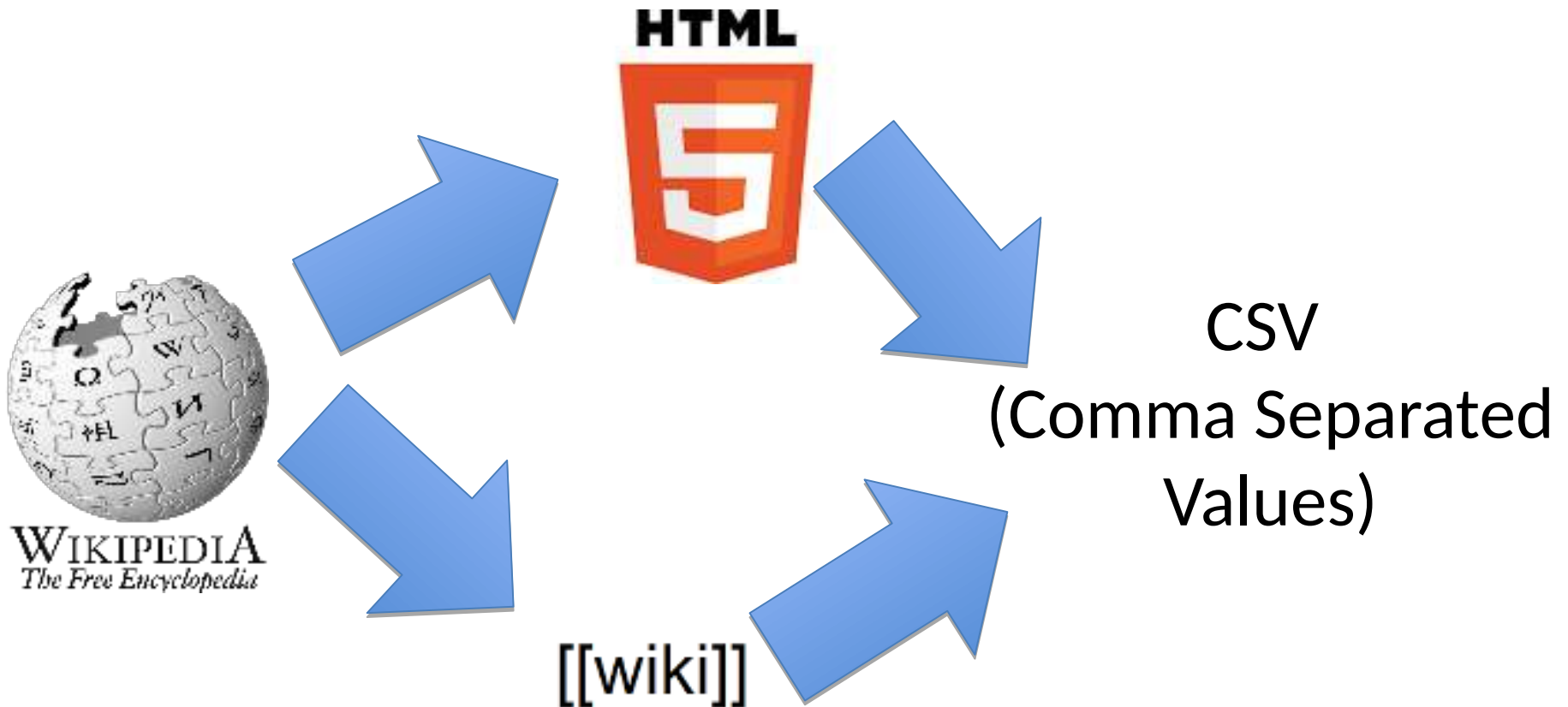
Wikitext

[[wiki]]

```
<tbody><tr>
<th>Model
</th>
<th>Image processor
</th>
<th>Sensor format
</th>
<th>Megapixels
</th>
<th>Min ISO
</th>
<th>Max ISO
</th>
<th>Autofocus points
</th>
<th>Viewfinder
<p>magnification,
</p><p>coverage
</p>
</th>
<th>Display
<p>size, dots
</p><p>(ratio)
</p>
</th>
<th>Touch
<p>screen
</p>
</th>
<th>Live view
</th>
<th>Max FPS
</th>
<th>Storage
</th>
```



WikipediaMatrix

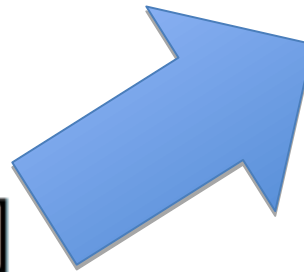


WikipediaMatrix

HTML



Comparison_between_Esperanto_and_Ido
Comparison_between_Esperanto_and_Interlingua
Comparison_between_Esperanto_and_Novial
Comparison_between_Ido_and_Interlingua
Comparison_between_Ido_and_Novial
Comparison_between_U.S._states_and_countries_by_GDP_(PPP)
Comparison_of_ALGOL_68_and_C++
Comparison_of_Afrikaans_and_Dutch
Comparison_of_Android_e-book_reader_software
Comparison_of_Asian_national_space_programs
Comparison_of_Axis_8_Allies_games
Comparison_of_C_Sharp_and_Visual_Basic_.NET
Comparison_of_Chernobyl_and_other_radioactivity_releases
Comparison_of_Exchange_ActiveSync_clients
Comparison_of_Hokkien_writing_systems
Comparison_of_Home_Owners'_and_Civic_Associations
Comparison_of_IOC,_FIFA,_and_ISO_3166_country_codes
Comparison_of_Java_and_C++
Comparison_of_Linux_distributions
Comparison_of_MD_and_DO_in_the_United_States
Comparison_of_Norwegian_Bokmål_and_Standard_Danish
Comparison_of_Portuguese_and_Spanish
Comparison_of_Symbian_devices
Comparison_of_United_States_presidential_candidates,_2008
Comparison_of_World_War_I_tanks
Comparison_of_browser_synchronizers
Comparison_of_business_integration_software
Comparison_of_consumer_brain-computer_interfaces
Comparison_of_domestic_robots
Comparison_of_e-book_formats
Comparison_of_e-book_readers
Comparison_of_file_hosting_services
Comparison_of_layout_engines_(Cascading_Style_Sheets)
Comparison_of_layout_engines_(MathML)



]]

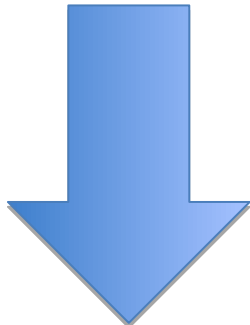
CSV
(Comma Separated
Values)

Plus de 300 pages Wikipedia!

WikipediaMatrix: The Truth



extracteur (groupe 1)

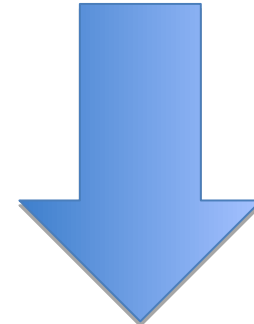


CSV1

Qui a raison?

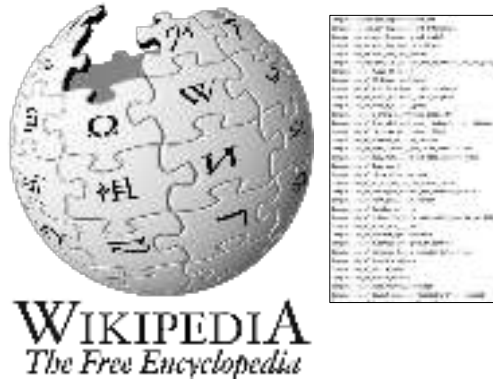


extracteur (groupe 2)



CSV2

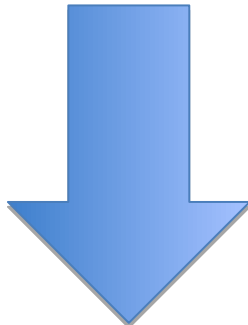
WikipediaMatrix: The Truth



[[wiki]]



extracteurs (groupe 1)



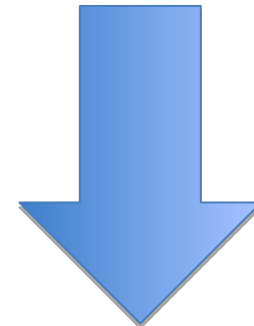
CSV1



[[wiki]]



extracteurs (groupe 2)

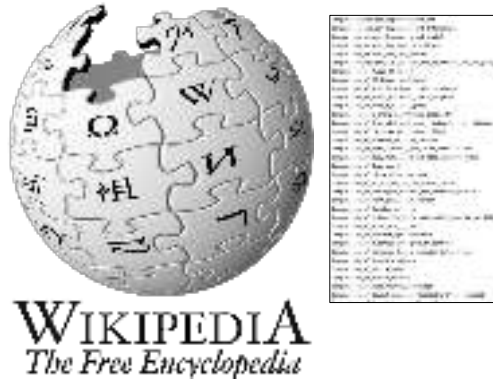


CSV2

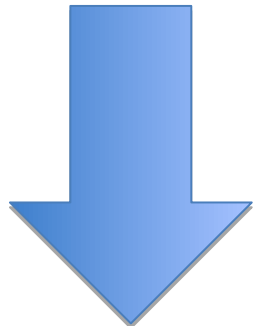
Qui a raison?



WikipediaMatrix: The Truth



(groupe 1)

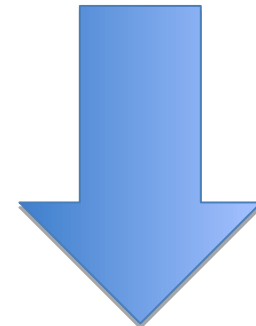


CSV1

...



(groupe 8)



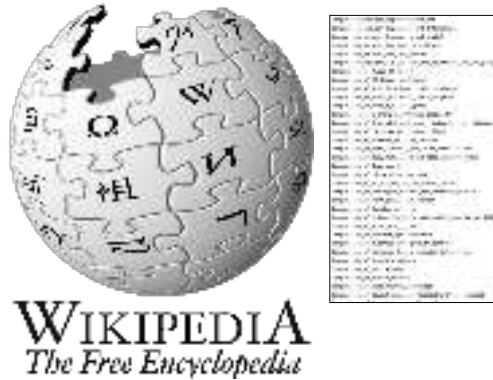
CSV8

Qui a raison?

...

...

WikipediaMatrix: The Truth



[[wiki]]

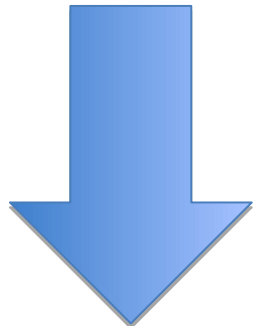
...



[[wiki]]

(groupe 1)

(groupe 8)



CSV1

...

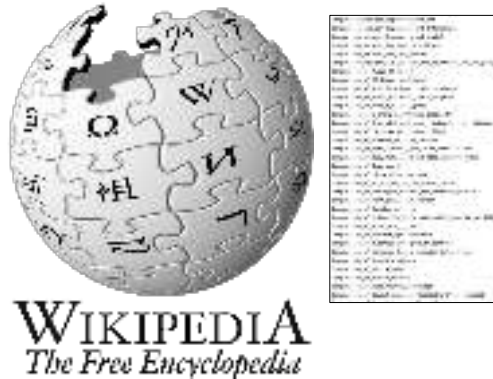
...



CSV8

**[http://
blog.mathieuacher.com/
WikipediaMatrixChallenge/](http://blog.mathieuacher.com/WikipediaMatrixChallenge/)**

WikipediaMatrix: The Truth



(groupe 1)



...



(groupe 8)



**Prise en main, audit, et
amélioration d'un projet de l'année
dernière**

WikipediaMatrix: The Truth



(groupe 1)

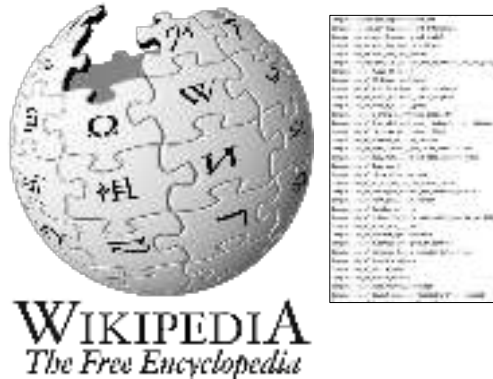
...



(groupe 8)

**Mise en place d'outils automatiques
pour évaluer automatiquement les
extracteurs**

WikipediaMatrix: The Truth



(groupe 1)

...

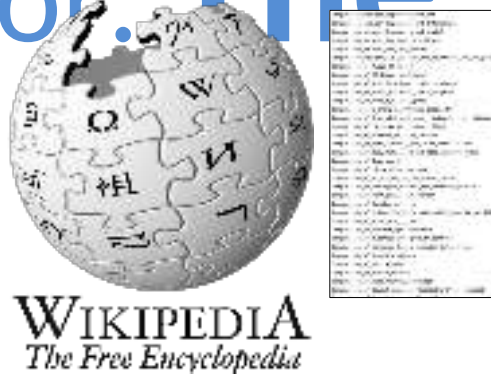


(groupe 8)

Construction d'une vérité terrain (jeu de données sur les résultats attendus) et d'une suite de test

Polyglot Wikipedia Matrix

Extractor: The Truth



(groupe 1)

...



(groupe 8)

Construction d'une vérité terrain (jeu de données sur les résultats attendus) et d'une suite de test avec un outil (SP1)

Polyglot Wikipedia Matrix Extractor



WIKIPEDIA
The Free Encyclopedia



(groupe ??)



(groupe ??)

**Développement de nouveaux extracteurs
(uniquement à partir de HTML), soit en
JavaScript, soit en Python + évaluation de ces
nouveaux extracteurs (SP2)**

Organisation et évaluation

Projet

- “Data engineering”
 - Data extraction/transformation
 - Data analysis/science
- En pratique les activités sont très similaires aux années précédentes:
 - Eliciter et valider des exigences
 - Etre capable d’écrire des documents techniques
 - Développement pour traiter et transformer des données
 - Comprendre des APIs existantes et les utiliser efficacement
- Spécificité cette année
 - thématique autour de données tabulaires
 - reprise de projet
 - multi-langage (polyglote)
 - test

A rendre

- Eliciter des exigences (EX)
 - édition d'un README.md, DESIGN.md, et INSTALL.md sur votre dépôt Github
- Sprint (SP1 + SP2)
 - code source (incluant des tests et des instructions) et jeu de données
 - commits sur github
- Présentation (PR)
 - mi-janvier: préparation de slides et présentation collective de 20' + 10' de questions (en français)

Eliciter des exigences (EX)

- README.md
 - permet à tout visiteur du projet de comprendre rapidement le but du projet, les fonctionnalités supportées, les fonctionnalités à développer dans le futur, ainsi que la licence du projet, les technologies utilisées, les participants et le contexte du projet, etc.
- DESIGN.md
 - reprend les grands éléments d'architecture incluant des modèles statiques (organisation des packages, descriptions des classes principales et de leurs responsabilités, etc.) ainsi que des modèles dynamiques (flux des événements, scénarios nominaux et exceptionnels, etc.)
 - L'objectif de ce document est qu'un potentiel contributeur externe puisse comprendre l'architecture du projet pour pouvoir éventuellement reprendre le code et l'étendre/l'améliorer
- INSTALL.md
 - décrit comment on construit le projet à partir du code source, comment on exécute les suites de tests, comment on exécute le logiciel

EX (bis)

- Dialoguer avec le « client » pour éliciter les exigences
- C'est aussi un travail technique: il faut installer des logiciels, comprendre et réutiliser du code et des données, prototyper, etc., sinon
 - impossible de comprendre le « contexte » et l'existant
 - impossible de dialoguer avec le client
- Un bon document: complet, non ambiguë, bien structuré, lisible; ré-exploitable par d'autres personnes!
- Il est ~~possible~~ normal que le projet que vous reprenez soit imparfait: à vous de l'améliorer!
 - Méthode (dans cet ordre): (1) création de "Github issues" (2) tests automatiques (3) correctifs

Soutenance (PR)

- 20' de présentation
 - Rappel du contexte
 - EX
 - Description de l'implémentation (SP1 + SP2)
 - Retour d'expérience
- 10' de questions par le jury

Séances

- X séances
 - Présentiel non obligatoire
 - Mettre rapidement en place les outils pour travailler collectivement à distance
 - Point chaque semaine (en visio et via l'avancée sur Github)
- TP et TD
 - Pas de distinctions
 - Créneaux pour (1) avancer (2) discuter avec le client
 - **Autonomie** collective
- Cours magistraux: adaptatifs (wait & see)
 - Outils, Méthodologie de tests, Web

Evaluation

- EX (5 points)
 - ~ mi-octobre
- SP = SP1 + SP2 (10 points)
 - ~ mi-novembre et mi-décembre
- PR (5 points)
 - ~ mi-janvier



Projet

- Groupe
 - Outils de versioning (git, github)
 - Outils collaboratifs (Slack)
- Répartissez-vous les rôles
 - Autrement: impossible de rendre en temps et en heure
- Résultats attendus
 - Très fortes contraintes sur les dates de rendus (cela fait partie intégrante de l'exercice)

TODOs

- **Sondage (cf Teams)**
- **Inscription sur github.com**
- **Lecture de**
<http://blog.mathieuacher.com/WikipediaMatrixChallenge/>
- **Lecture de:**
 - <https://annuel2.framapad.org/p/r.47fadb4c2ba65628e2a7ef94c0c5fbea>
 - <https://annuel2.framapad.org/p/r.35d94f38fd3cbcd4e5664cab60aad1f7>
- **Constitution des groupes:**
<https://tinyurl.com/PDL2021-groupes>
- **Une concertation à l'échelle de la promotion est nécessaire; à vous de jouer!**

PDL: le “projet”

- 8 groupes maximum
- Groupe de 4 personnes
 - Exceptions possibles (5 max); m’en parler ASAP

Date limite de composition des groupes: mercredi 9 septembre à 14h

Si vous n’avez pas de groupe, notez quand même votre prénom/nom

Prochaine séance

- Point sur la constitution des groupes
- Créneau pour lire les documents/articles, notamment:
<http://blog.mathieuacher.com/WikipediaMatrixChallenge/>
- Séance de questions/réponses sur l'article
 - Interrogation de certains étudiants
 - Questions de l'audience pour éclaircir certains points

