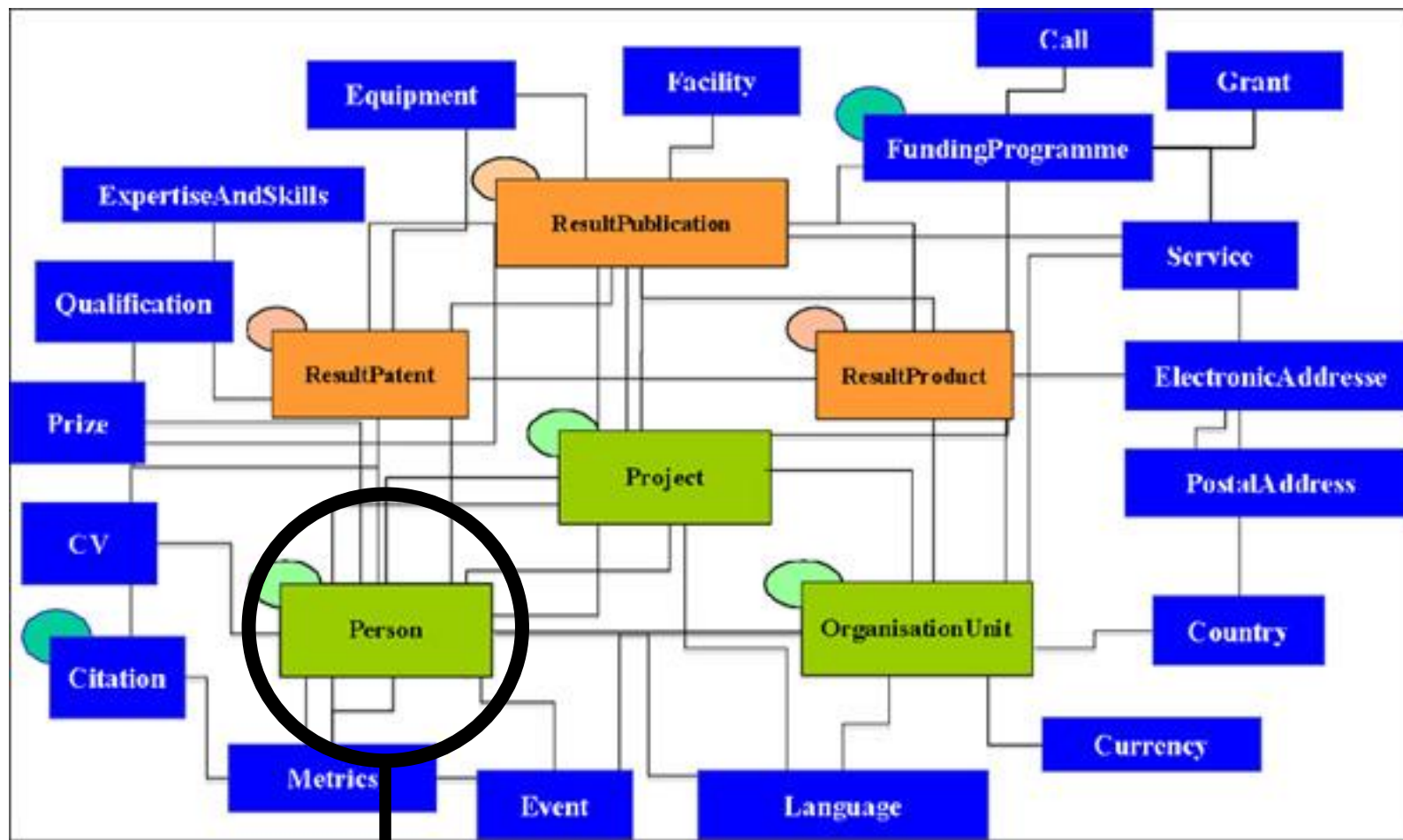


The SAFARI Syndrome

Implementing CRIS and Open Science

By Joachim Schöpfel, University of Lille 3





From object to subject

The human factor

- Objectives
 - Evaluation, reporting, funding, governance
- Values and ethics
 - Academic freedom, efficiency of management, control
- Rights
 - Privacy, intellectual property
- Public/user
 - Research manager, scientist, authority



(1) Ethics, object of CRIS

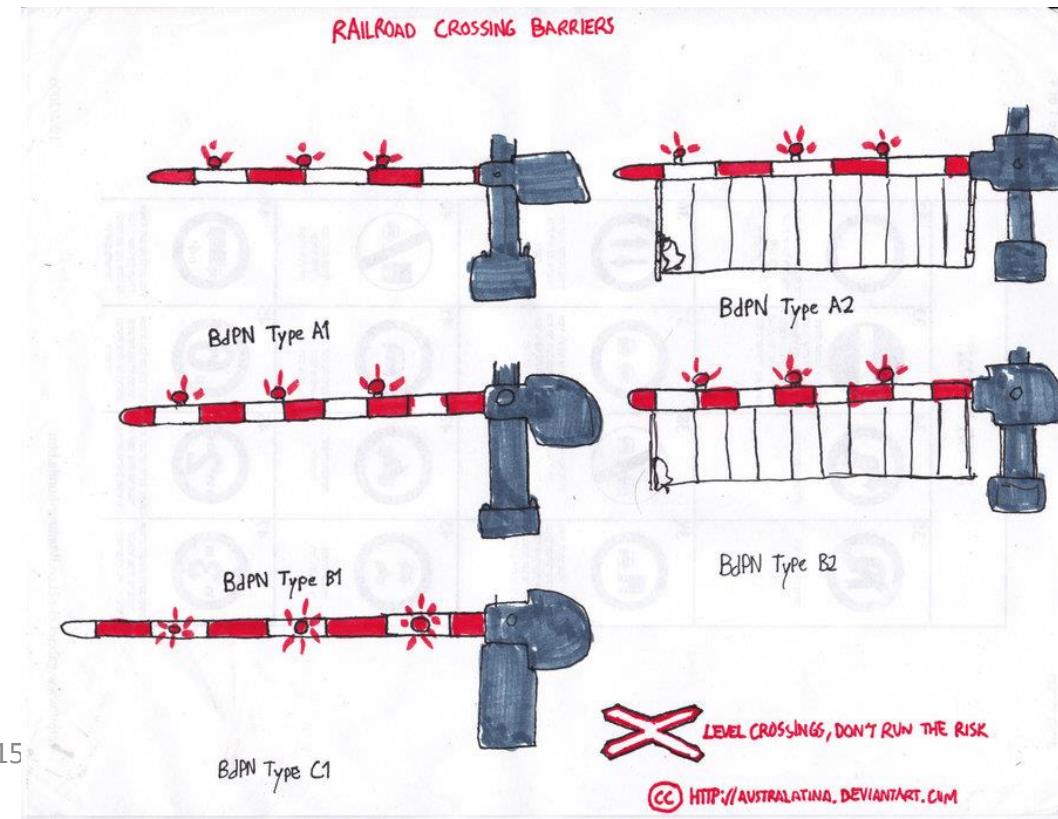
- Emerging trend
- Low priority (same level as biohazards)
- Part of legal feasibility
- Ethical handling
 - Of intellectual property
 - Of data protection laws
- Sensitive information in CRIS, reason of access restrictions
 - Related to institutions
 - Related to persons
- Part of research process
 - Ethical advisors
 - Ethical reviews



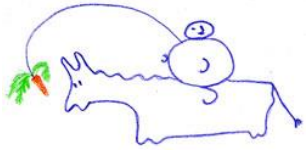
"Water-babies 2" by Linley Sambourne - The Water-babies by Charles Kingsley. Licensed under Public Domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Water-babies_2.jpg#/media/File:Water-babies_2.jpg

(2) Ethics, barrier of CRIS

- « A legalistic barrier to research data management »
- Privacy, together with national security and foreign access to sensitive data as barriers to a wider use of CRIS
- Externalized problem
 - As a societal issue
 - As a governmental issue
- « Sometimes, law and ethics get in the way »
- Fears and concerns of researchers
 - About privacy
 - About control
 - About transparency



Scepticism and Resistance – Reasons for the Fears



BIG ZUCKER



- Resistance against a **Science Management inspired by the principles of New Public Management** including Measurement & Management;
- Nostalgic **Clinging to the old German Humboldt-Model** with its Freedom of Research and Higher Education without any Measurement or Management.

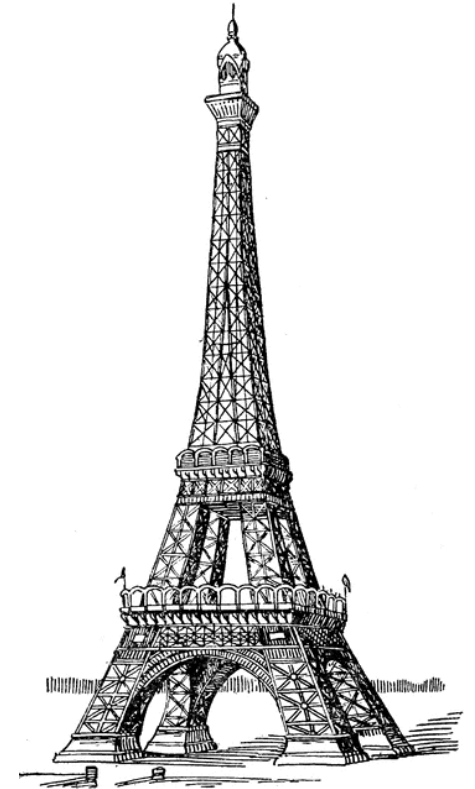
- The currently **Bad Reputation of Social Networks**, that secretly collect Informations about our Life and Disposition and sell them to who-knows-who,
- and that especially the Scientists Profiles of CRI-Systems are quite similar to.

- The currently **Bad Reputation of Companies like Google**, that secretly track us surfing in the Internet or using Google-Products like Google+ and Google-Mail and sell these Informations to who-knows-who,
- and that especially the automated Import of Data from external Sources and the intelligent Mapping to Persons reminds of.

- **Distrust in THEM in Supervisory Board, Executive Board and Central Administration**, who with a CRIS in their hands would have another instrument
 - to **end the Freedom of Academia**,
 - to **erect a System of Surveillance** made **to control and to exploit** Scientists,
 - to **justify the next Dismissals, Wage Cuts and Budget Cuts**.

In particular

- Evaluation is (often) a problem, not (always) a solution
 - Time consuming, lack of transparency, rejection rate
- Governance: ruling the campus like an automobile company?
 - Efficiency vs freedom, outcome control, assessment procedures
- Budget cuts
- Historical experience with (resistance to) centralism
 - Conflicting interests, organisational complexity (« *millefeuille* »)
 - Adopting UK, D, No, NL solutions?
- Outsourcing of data production and control
 - Where have all the start-ups gone?
- Public debate (again) on surveillance, privacy and public freedom



Tandis que le ministère de l'intérieur développe la centralisation de ses renseignements

Une division de l'informatique est créée à la chancellerie

En ordre dispersé, les départements ministériels tentent de développer à leur profit, à leur seul usage, l'informatique et son outil, l'ordinateur. Ce n'est pas tout à fait un hasard si, à l'époque où le Journal officiel va publier un arrêté créant une « division de l'informatique » au ministère de la justice, celui de l'intérieur met la dernière main à la mise en route d'un ordinateur

puissant destiné à rassembler la masse énorme des renseignements grappillés sur tout le territoire; pas un hasard non plus si le projet SAFARI (Système automatisé pour les fichiers administratifs et le répertoire des individus) destiné à définir chaque Français par un « identifiant », qui ne définit pas que lui, maintenant terminé, est l'objet de convoitises ardentes; le ministère de l'intérieur y souhaite

jouer le premier rôle. En effet, une telle banque de données, s'ouvrant opérationnel de toute autre collecte de renseignements, donnera à celui qui la possèdera, une puissance sans égale.

Ainsi se trouve d'évidence posé un problème fondamental, même s'il est rebattu : celui des rapports des libertés publiques et de l'informa-

tique. Son importance exigerait qu'il en fût, au Parlement, publiquement débattu. Tel ne paraît pas être, pourtant, la solution envisagée par le premier ministre dans les directives qu'il vient d'adresser au ministère de la justice, intéressé au premier chef si l'on s'en rapporte à la Constitution qui dans son article 66 fait de l'autorité judiciaire le gardien des libertés individuelles.

« Safari » ou la chasse aux Français

Rue Jules-Breton, à Paris-13^e, dans des locaux du ministère de l'intérieur, un ordinateur Iris-80 avec bi-processeur est en cours de mise en marche. À travers la France, les différents services de police détiennent, selon la confiance faite par un très haut magistrat, 100 millions de fiches, réparties dans 400 fichiers. Ainsi se trouve posées — et, à terme, théoriquement résolues — les données d'un problème comprenant, d'une part, l'énormité des renseignements collectés : de l'autre, la méthode à définir pour faire de cet ensemble une source unique, à tous égards, de renseignements.

L'histoire du très puissant appareil qu'est l'Iris-80 est exemplaire du secret qui entoure l'épanouissement de l'informatique dans les administrations, quelles que puissent être les informations qui filtrent ici et là.

Puissant, cet Iris-80, une comparaison le démontre sans contestation. L'appareil employé pour engranger les données de l'opération Safari, qui concerne l'identification individuelle de l'ensemble des 52 millions de Français, a une contenance de 2 milliards d'octets (1) : celle de l'ordinateur du ministère de l'intérieur est de 3,2 milliards d'octets.

C'est dire que la mise en route d'Iris-80 — dont la location coûte 1 million de francs chaque mois — a été précédée d'études, de tests pour en éprouver les possibilités. D'autant qu'à lui seul, il doit remplacer les trois GE 400 et le 10070 de la C.I.I. qu'employait jusqu'alors la Place Beauvau.

C'est sur ce dernier ordinateur qu'ont eu lieu les essais. Pour 20 % de sa capacité, il a été consacré à la gestion du personnel communal de la Ville de Paris. Mais, pour le reste (80 %), il a servi à tester les programmes devant être fournis à l'Iris-80, afin de rendre cohérentes entre elles, les données contenues dans les 400 fichiers que possèdent les services de police : renseignements généraux, direction de la surveillance du territoire, police judiciaire, etc.

A lire d'anecdote, on peut rappeler que ce 10 070 de la C.I.I., à

l'origine, budgétairement, n'était pas du tout prévu pour la tâche qu'il a finalement assurée, mais pour « traiter » les données administratives du Fichier national des constructeurs (F.N.C.). Il s'agit donc apparemment d'un détournement manifeste de crédits d'études, ce qui n'était sans doute pas le vœu du Parlement qui les vota.

De vastes ambitions

Il n'y a pas que cela. Le ministère de l'intérieur a d'encore plus vastes ambitions. Détenteurs, déjà, du fichier national du remembrement, les services de M. Jacques Chirac ont de grands efforts pour, affirmation, s'adjoint d'autres : le cadastre, le fichier de la direction nationale des impôts et, plus grave peut-être, celui du ministère du travail.

De telles visées comportent un danger qui saute aux yeux, et que M. Adolphe Touffait, procureur général de la Cour de cassation, avait parfaitement défini le 9 avril 1973 devant l'Académie des sciences morales et politiques, en disant : « La dynamique du système qui tend à la centralisation des fichiers risque de porter gravement atteinte aux libertés, et même à l'équilibre des pouvoirs politiques ».

C'est si vrai que la règle nationale des usines Renault, par exemple, dispose déjà d'une base de données établies à partir d'un fichier du personnel.

On admettra, dans ces conditions, que l'ouverture d'un débat public paraîtra particulièrement urgente pour définir les limites de l'emploi des banques de données. Or ce débat paraît, dans le principe, écarté par le premier ministre, qui, dans une lettre directive adressée voici quelques semaines à M. Jean Taittinger, avait écarté une telle procédure au profit de circulaires, voire de décrets, préservant en tout état de cause le secret de décision de l'administration. On connaît le peu d'efficacité immédiate que peut avoir devant le juge administratif le recours en annulation d'actes du gouvernement...

Ce n'est pas, pourtant, que les avertissements aient manqué. Le Conseil d'Etat en 1970, puis le ministère de la justice en 1972 (qui avait appelé le rôle dévolu à l'autorité judiciaire de « gardien des libertés individuelles » et donc réclamé voix au chapitre) ont insisté sur la nécessité d'une intervention législative qui préciserait les quelques éléments essentiels de l'emploi de l'informatique appliquée aux particuliers : réglementation de l'accès des tiers aux fichiers, de l'intercommunication de ceux-ci, droit de rectification des personnes fichées si les renseignements retenus sont inexacts, etc.

De plus, tous les exemples étrangers incitent à ce débat sur une utilisation de l'informatique à laquelle, par définition, il ne s'agit pas de renoncer, mais à qui doivent être tracées des limites, si grand est le danger qu'elle implique. La désignation par le gouvernement d'une commission de « sachsants » dans les semaines à venir ne saurait suffire à remplacer le débat parlementaire dont on se méfie si visiblement.

En fait de débats parlementaires, il y a d'ailleurs des précédents qui sont le fait, précisément, du ministère de la justice et n'ont pas compromis le développement des fichiers. Avec le casier judiciaire, depuis longtemps, la chancellerie a l'expérience de semblables fichiers. Quel que soit le jugement qui peut être porté sur le principe d'un tel

outil, il n'apparaît pas — sauf erreurs négligeables, relativement — que l'accès des tiers ou le droit à contrôle des personnes visées — par demande d'un extrait — ait jamais provoqué des bavures préjudiciables à la légalité.

De même, le fichier national des conducteurs, dans sa partie judiciaire, est prévu par une loi, et il faut regretter que les textes d'application ait permis des illégalités injustifiables — mais connues (le Monde du 8 mars).

« A la hussarde »

Fort, pourtant, de ces avantages, le ministère de la justice paraît curieusement se laisser dépasser par des querelles internes peu compréhensibles. L'arrêté signé le 18 mars par M. Jean Taittinger le montre. La création d'une « division de l'informatique », place Vendôme, serait en soi une bonne chose, du point de vue de l'efficacité, si les conditions de sa création, engagée vraiment voici trois mois, ne prenaient l'allure d'une peu élégante tentative d'élimination dirigée contre certains esprits novateurs ayant eu le mauvais goût de s'intéresser trop tôt à l'informatique.

Il serait, en effet, bien étonnant que les membres de la commission de l'informatique au ministère de la justice, que préside M. Adolphe Touffait, ne s'offusquent pas d'une décision qui, en soi, ne peut avoir pour but que de « vider de sa substance » la

dite commission. D'autant qu'il est d'ores et déjà connu que M. Touffait a été rayé de la liste des « sachsants ». Il semble d'ailleurs que les réactions vives qui sont enregistrées portent moins sur le renouvellement des structures, jugées inévitables, que sur la méthode « à la hussarde » employée par tel membre de l'entourage de M. Taittinger pour mener à bien ses projets de rénovation de la gestion dans le domaine judiciaire.

Est-ce à dire de plus que les choix que l'on entend promouvoir soient nécessairement les plus opportuns ? Tout indique, pour l'instant, que, si le ministère de l'intérieur a définitivement choisi le « matériel lourd » pour s'équiper, la chancellerie, au contraire, s'oriente vers un réseau de mini-ordinateurs placés auprès de chaque tribunal de grande instance important.

Dans cet ordre d'idée, le choix déjà décidé de M. Jean Malbec, vice-président à Bobigny (Seine-Saint-Denis), comme futur chef de la division de l'informatique (au point qu'il a, dès à présent, effectué des missions d'information à Lille, Nice, Lyon et Marseille dans les semaines passées), est significatif. Il est, en effet, à Bobigny l'apôtre d'un système « mini » qu'il souhaite étendre à l'ensemble de l'institution judiciaire. Ce n'est sans doute pas non plus par hasard si la télévision, lundi, après avoir donné des extraits du discours de M. Taittinger à Gap sur la justice civile et les nécessités

d'un *aggiornamento* technique, a illustré son discours par un large reportage sur les équipements du tribunal de Bobigny — plus réduits, donc plus rapides à réaliser, ainsi plus vite source d'orgueil pour leurs créateurs.

C'est donc un doute global qui pèse sur les intentions du gouvernement, en général, et du ministère de la justice, en particulier : ce dernier département, qui rappelle à tous sa mission de protection des libertés individuelles, a apparemment accepté sans broncher la suppression d'un éventuel débat public, ce qui jette sur les déclarations « libérales » de M. Taittinger en d'autres domaines une suspicion qui n'est pas de bon aloi.

Mais, dans cette entreprise, le ministère de la justice, même s'il fait preuve d'une grande mollesse pour la défense de ses idéaux, car il ne s'agit pas seulement à présent de « protéger des délinquants », n'est pas essentiellement en cause. Ce qui l'est, c'est une entreprise dont on a tout lieu de suspecter la pureté tant on prend soin de cacher sa réalisation.

PHILIPPE BOUCHER.

(1) L'octet, ensemble de huit « bits », est l'unité de mesure de la plupart des ordinateurs. Quand on enregistre un texte dans la mémoire, chaque caractère du texte occupe un octet. Un milliard d'octets représente, en gros, la capacité de mémoire de cinquante bandes magnétiques.

Why SAFARI?

1974 project of French government to interconnect personal data files and create a centralized database through the use of a unique identifier (INSEE code, social security number)

One day after leaking by project team members to the journal *Le Monde*, the project was suspended by French president



euroC

Océ satisfait vos besoins en copies d'une autre façon.



The SAFARI syndrome



A good (?) idea

- Increasing efficiency
- Centralized database
- Interconnection of administrative data files
- Use of personal data
- Use of a unique identifier (INSEE code, social security number)

An unexpected outcome

- Leaking by project members
- Debate on surveillance and public liberty in media and parliament
- Massive rejection of project
- « Hunting French citizens »
- At the end, the first law on privacy to protect citizens against government

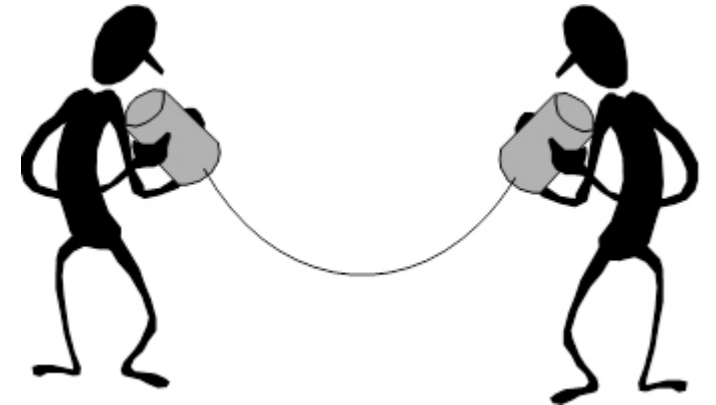
Back to the campus

- High valuation of (academic) freedom
 - #JeSuisCharlie
- Awareness of privacy issues
 - Rejection of surveillance
 - « Privacy concerns can prevent stakeholders – individuals, governments, academia, and companies - from sharing their data » (Haak et al 2012)
- Scientists (mostly) like their ivory tower
- Rejection of administrative extra work
- Capacity of creating complex situations



How to prevent the SAFARI syndrom

- CRIS is not (only) a problem of IT management
 - Political governance
 - Top-down or bottom-up?
- Development of a culture of information
- Open and proactive communication
- Integration of legal expert in project team
- Impact evaluation
 - CRIS and funding?
 - CRIS and resource attribution?
- Provision of safeguards for anonymity and privacy of research participants
 - Who controls the data?



A changing paradigm

When you're working

only 1 person
WORLD

openness is a
strange attractor
... it's inevitable
and irresistible

open
science

Baby

OPEN NOTE-



BOOK science

USEFUL CHEM.
BLOGSPOT.COM
Jean-Claude
Bradley
@jcbradley

research.
heartanthy.com

@science
of ants

CAPTURE detail

- see how it was done
- builds quality, authority

- Transparent processing
- Community science
- Open data
- etc.

organize

- blog might offer more organization than a wiki... for some
- lots of people use a wiki... but needs

STRUCTURE
& discipline

to really get the
most out of it...

Google - searchable if fine
BUT

Best to build a dB
& make everything
MACHINE READABLE

contributions from unexpected places

Just do it!

CRIS w/ or v Open Science?

- Adopting the new paradigm (and taking it seriously)
 - No two-tiered strategy
- (Re)placing the scientist at the heart of the process
- Focusing on community concerns, needs and values
- Addressing privacy issues not as barriers
 - But as objective
- Learning from historical experience
 - “Whoever wishes to foresee the future must consult the past; for human events ever resemble those of preceding times.” (Machiavelli)



©2010-2015 EddyCamara

Thank you!

Bibliography <http://www.citeulike.org/user/Schopfel/tag/safari>

Social network <https://fr.linkedin.com/in/schopfel>

Contact joachim.schopfel@univ-lille3.fr