



UNIVERSITY OF
CANBERRA

NEWS & MEDIA
RESEARCH CENTRE

10 ans
OPEN CIO
SUMMIT

Inria



PARIS
DIDEROT
université

OPEN SOURCE DIGITAL INFRASTRUCTURE

MAPPING THE FIRM-PROJECT NETWORK

MATHIEU O'NEIL - UNIVERSITY OF CANBERRA

XIAOLAN CAI - UNIVERSITY OF CANBERRA

LAURE MUSELLI - TELECOM PARIS – IP PARIS

MAHIN RAISSI - UNIVERSITY OF CANBERRA

STEFANO ZACCHIROLI - UNIVERSITE PARIS DIDEROT / INRIA

LES FIRMES ET LA PRODUCTION COLLABORATIVE OPEN SOURCE

Depuis les années 90, l'infrastructure digitale de l'internet repose principalement sur des logiciels Libres et Open Source (F/OSS)

- Linux, Apache, Mozilla, etc...

Les entreprises IT ont accepté les licences et l'éthique open source

- En 2002, IBM commence à investir dans Linux
- Adoption de licences ouvertes
- L'utilisation de F/OSS est maintenant un standard dans les entreprises
- Rôle de GitHub et Stack Overflow
- En 2018 :
 - Microsoft rachète GitHub;
 - 85% du code Linux code est produit par des employés d'entreprises

LES FIRMES ET LA PRODUCTION COLLABORATIVE OPEN SOURCE

Des entreprises paient les salaires de certains développeurs...

**... pendant que d'autres ont un comportement de 'free riding' sur le travail
bénévole des développeurs**

Ce qui nous mène à des questionnements :

- Justice de distribution des bénéfices ?
- Emergence de formes hybrides de travail et d'organisation ?
- Quelles sont les entreprises qui rémunèrent les développeurs ?
- Impact du travail rémunéré sur les projets... et l'industrie de l'IT ?

PROJET DE RECHERCHE

Albert P. Sloan Foundation / Ford Foundation “Digital Infrastructure” grant
(2019-2020)

Mapping the co-production of digital infrastructure by peer projects and firms

1. Cartographier le réseau global de connexions entre firmes et projets
2. Comprendre l’impact du travail rémunéré sur l’autonomie et la pérennité des projets
3. Implications en termes de politiques publiques ?

METHODES DE RECHERCHE

Phase 1 : Quelles entreprises paient leurs employés pour contribuer ?

1. Collecte des adresses email des contributeurs à 135 repositories Github et extraction des 50 firmes les plus représentées
2. Analyse sémantique d'articles relatifs au F/OSS sur 3 plateformes de journaux IT en ligne
 - Co-locations firme-projet
 - Analyse sémantique et de sentiment

Phase 2 : Quelles sont les préoccupations autour du travail rémunéré dans les projets ?

1. Questionnaire aux 20 repositories/projets auxquels les entreprises contribuent le plus
2. Analyse sémantique de 5 projets : sources de préoccupation autour du travail rémunéré

Phase 3 : Quels sont les discours des entreprises et des développeurs ?

Ethnographies à l'European (Lyon) et Paris Open Source Summits et au FOSDEM (Bruxelles)

PROJETS OPEN SOURCE SUR GITHUB

- GitHub est la plus grande plateforme mondiale d'hébergement de développements logiciels
- Populaire chez les développeurs F/OSS
- Créé en 2008 et racheté par Microsoft en 2018
- 37millions d'utilisateurs et 96 millions de repositories
- Collecte de données de 135 repositories : 1,048,557 contributions entre le 01/01/2015 et le 30/042019

FIRMES ET PROJETS SUR GITHUB

Summary of top 100 contributor types in 135 GitHub repositories

Category	Number of contributors	Number of commits
Firms	41	1025681
No affiliation*	20	1060168
F/OSS project	19	149293
Individual developer	19	68133
Research	1	42451
Total	100	2345726

* Email domains belongs to email servers that couldn't identify committers' affiliations, such as: @gmail.com, @outlook.com, @users.noreply.github.com, are classified as no-affiliation.

Summary of licences of 135 GitHub repositories

Licence family	Number of projects
Apache	28
MIT	53
BSD	18
GPL	19
Others	17
Total	135

* Different versions and variation of one licence are categorized into the same licence family. Project applies licence / licences other than listed families are categorized into others.

FIRMES LES PLUS ACTIVES & PROJETS SUR GITHUB

Top-20 firms by number of commits contributed

Rank	Firm	Number of commits
1	Microsoft	438220
2	JetBrains	126485
3	Google	72387
4	Red Hat	54788
5	Apple	37520
6	Intel	27613
7	Facebook	23752
8	Pivotal	15984
9	GitHub	15324
10	IBM	12543
11	SUSE	9520
12	Huawei	7980
13	Oracle	5851
14	Samsung	5773
15	ARM	5131
16	Canonical	3364
17	Nvidia	3291
18	WIX	3168
19	ZTE	2022
20	Cisco	1460

Top-20 projects by number of commits made by firm employees

Rank	Project	Number of commits
1	torvalds/linux	340710
2	NixOS/nixpkgs	125205
3	Homebrew/homebrew-core	108709
4	apple/swift	67197
5	kubernetes/kubernetes	74201
6	Microsoft/vscode	49418
7	tensorflow/tensorflow	56656
8	dotnet/corefx	32884
9	DefinitelyTyped/DefinitelyTyped	54801
10	aspnet/AspNetCore	34946
11	spring-projects/spring-boot	17855
12	ansible/ansible	31544
13	elastic/elasticsearch	33983
14	rust-lang/rust	57790
15	facebook/react-native	16908
16	moby/moby	24472
17	home-assistant/home-assistant	18876
18	pytorch/pytorch	17717
19	apache/spark	15180
20	storybooks/storybook	18968

PROPORTION DE CONTRIBUTIONS FAITES PAR LES FIRMES

Number and proportion of commits made by firms of top-20 most active projects on GitHub

Rank	Project	Total number of commits	Number of commits made by firms	Proportion of commits made by firms
1	MicrosoftDocs/azure-docs	405029	344716	0.85
2	torvalds/linux	340710	247864	0.73
3	jlord/patchwork	139868	7263	0.05
4	NixOS/nixpkgs	125205	63042	0.5
5	Homebrew/homebrew-core	108709	54352	0.5
6	kubernetes/kubernetes	74201	40041	0.54
7	apple/swift	67197	40351	0.6
8	rust-lang/rust	57790	15897	0.28
9	tensorflow/tensorflow	56656	29515	0.52
10	DefinitelyTyped/DefinitelyTyped	54801	17920	0.33
11	Microsoft/vscode	49418	37366	0.76
12	aspnet/AspNetCore	34946	16486	0.47
13	elastic/elasticsearch	33983	16051	0.47
14	dotnet/corefx	32884	25660	0.78
15	ansible/ansible	31544	16252	0.52
16	servo/servo	29155	6395	0.22
17	moby/moby	24472	11743	0.48
18	rails/rails	23022	6132	0.27
19	symfony/symfony	22477	3845	0.17
20	FreeCodeCamp/FreeCodeCamp	21468	3660	0.17

ETHNOGRAPHIES

- Observations et analyse de discours à des forums open source
- European Open Source Summit (Lyon)
- Paris Open Source Summit
- A venir : FOSDEM (Bruxelles)

PARTICIPATION DES FIRMES IT AUX PROJETS OPEN SOURCE

“We KNOW developers because we ARE developers”

(INTEL)

IMPACT SUR LES PROJETS OPEN SOURCE

1- Des fondations aux valeurs différentes en fonction des projets, instruments de gouvernance et/ou de lobbying entre firmes et projets

2- Recours croissant à des “communautés” par les entreprises :

- Contribution à des projets indispensables à leur activité (via les contributions d’employés et la participation aux fondations)
- Création de communautés autour de leurs propres projets

3- Injonctions de professionnalisation auprès des projets, pour pouvoir offrir un service de qualité aux entreprises utilisatrices (harmonisation des normes entre entreprises et projets OSS)

RECOURS CROISSANT AUX COMMUNAUTÉS

- **Best practices** pour créer des communautés et augmenter les contributions externes
- **Management des Ressources Humaines externes**
- Création d'**Open Source Program Offices (OSPOs)**
- **Open Source “beyond the code”** : adapter le modèle collaboratif open source à d'autres activités comme le **marketing**, la rédaction ou la traduction de **documentation**

PROFESSIONALISATION DES PROJETS

Projets soumis à de nouvelles normes, dans un objectif de business :

- **Performance :**

- Métriques d'évaluation des différents projets
- Marketing des projets
- Sécurité (Security) et Fiabilité (Safety)
- Compliance
- Documentation

- **Légitimation :**

- Ethique, liberté, collaboration
- Diversité
- Objectif de confiance

QUELLES IMPLICATIONS POUR LES END-USERS?

Développement d'une infrastructure digitale open source orientée business dont les firmes IT sont les grands contributeurs

- Dans certains secteurs, des end-users commencent à s'impliquer dans les projets de développement de ces plateformes
- Quel futur pour les conditions d'utilisation par les end-users en termes de prix, mais surtout d'indépendance ?

MERCI !

Pour nous contacter :

Mathieu.O'neil@canberra.edu.au

laure.muselli@telecom-paris.fr