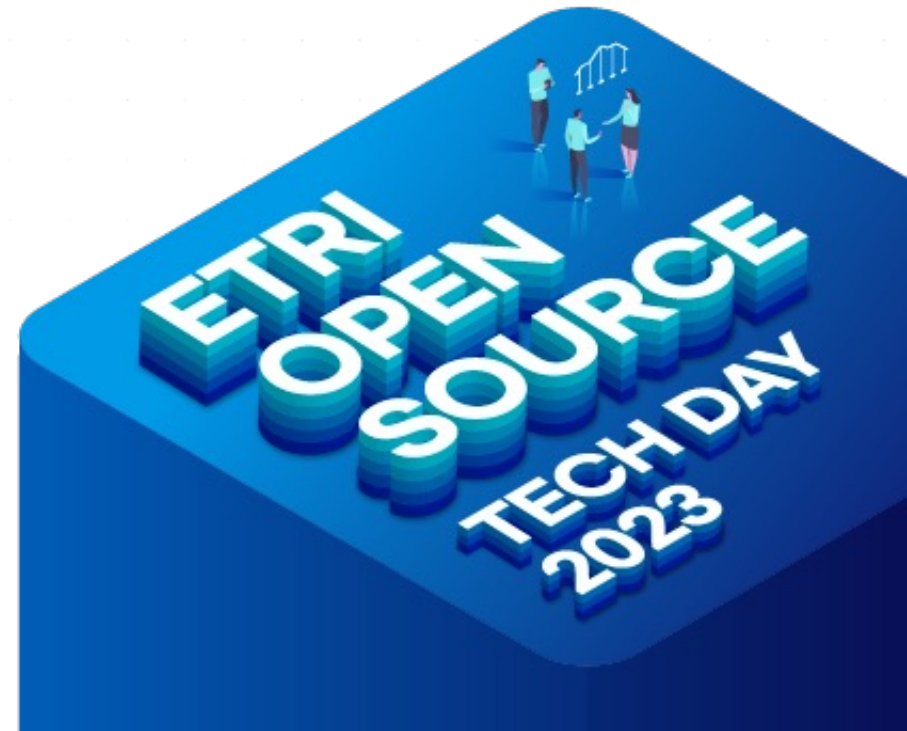



Open Source, 자연스럽게

이민석
국민대학교





CONTENTS

- 00 우리는 어디에 있는가?
 - 01 그러나
 - 02 Open Source, 자연스럽게
 - 03 요약
- 



00

우리는 어디에 있는가?



Open Source 사용

~ 100%

in all fields of

Business, R&Ds, and Hobbies

출처: 거의 틀림없는 느낌



Open Source 생산



100+ million Developers

4+ million Organizations

330+ million Repositories



아직 초기 단계의 Open LLM

22 Open LLMs

10 Open LLMs for Code

2 Open LLM datasets for pre-training

3 Open LLM datasets for instruction-tuning

1 Open LLM datasets for alignment-tuning



Open Source 소프트웨어 개발자 수입원

23% I'm a semi-professional maintainer, and earn some of my income from maintaining projects

13% I'm a professional maintainer, and earn most of my income from maintaining projects

4% Other

14% I'm an unpaid hobbyist and do not want to get paid for maintaining projects

46% I'm an unpaid hobbyist, but would appreciate getting paid for maintaining projects

60%
UNPAID

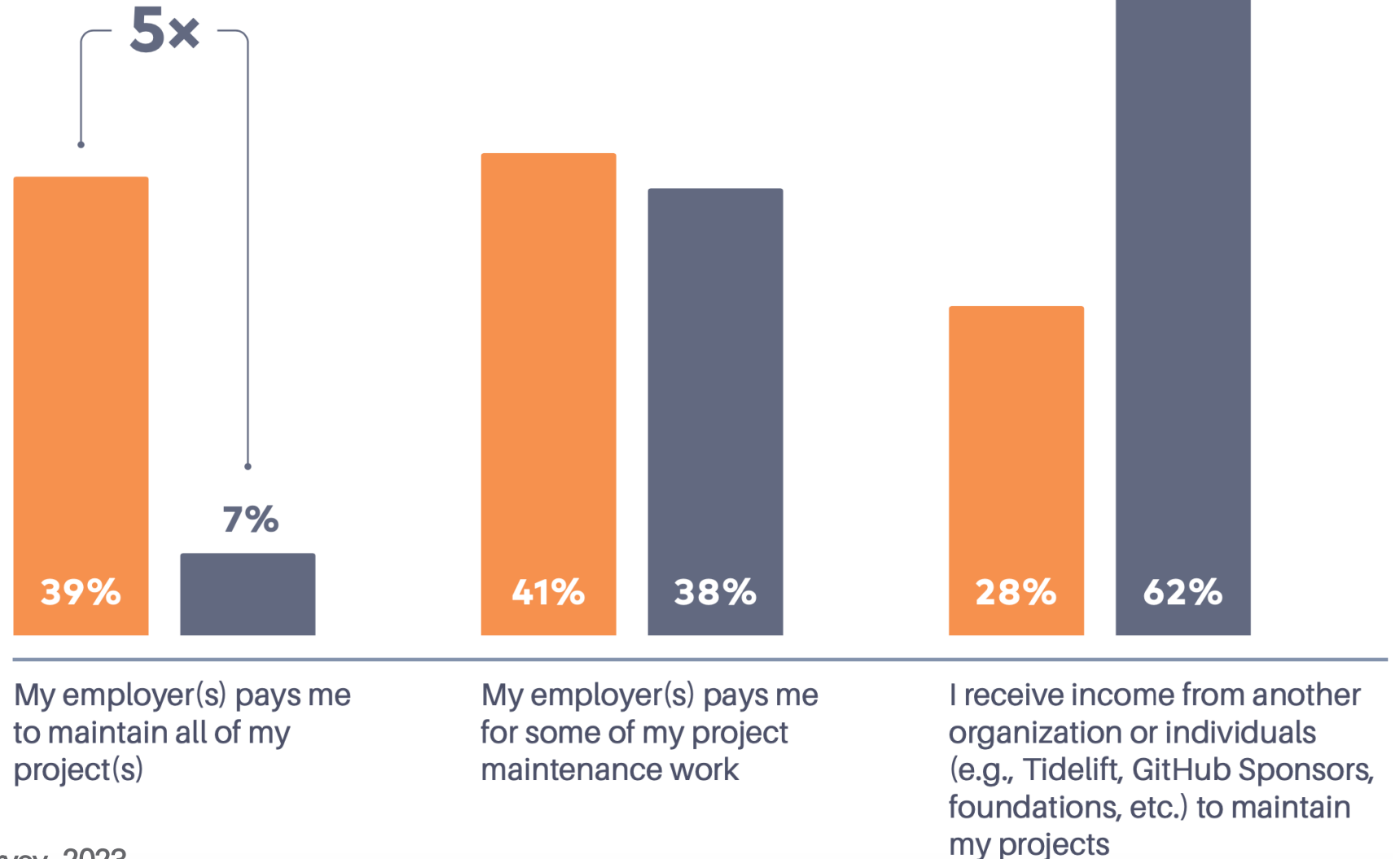
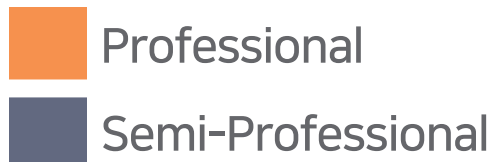
n=326

Tidelift Open Source Maintainer Survey, 2023



Open Source 소프트웨어 개발자 수입원

Maintainer





Open Source in Cloud Era

AWS, Microsoft, Google 주도의 규모 있는 생태계 형성

-- 기술 계층에 따른 서비스 시장이 잘 구축됨

수평적, 수직적 Value Chain이 잘 만들어지고, Data 쪽으로 확산 중

-- Subscription이라는 전통적 Business Model이 잘 동작



01

그러나



Open Source in Cloud Era

Open Source 철학에 대한 도전:

기술 자체 vs. Source, 수익 vs. 기여 ...

주로, 기술 집약적 Open Source가 Managed Service로 제공될 때

-- 기술적으로 꽤 안정된 상태에 이른 후, 수익 배분이 이슈

-- 그 결과,

- SSPL (ref, Mongo DB) : Managing Service의 소스 공개를 요구
- BUSL (red, Terraform) : 상업적 Service 제한 기간의 설정
- RSAL (ref, Redis) : 상업적 Managed Service 자체를 금지



우리나라의 오픈소스 기여

RANK	ORGANIZATION	ACTIVE CONTRIBUTORS ?	TOTAL COMMUNITY ?
1	Google (2 - 48K)	5633 +201	10355 +274
2	Microsoft (1 - 67K)	3999 +184	7945 +314
3	Red Hat	3128 +85	4615 +100
4	Intel	1940 +114	3951 +120
5	Amazon (5 - 47K)	1626 +87	3474 +144
6	IBM	1400 +58	2986 +133
7	Oracle (27 - 30K)	810 +34	1404 +43
8	VMware	787 +27	1321 +45
9	SAP	730 +33	1234 +47
10	GitHub	655 +18	1407 +50

13	Huawei (47)	515 +17	1272 +59
25	Alibaba (19)	309 +12	763 +40
28	Tencent	289 +14	798 +40

34	3	Samsung (21 - 48K?)
101	3	LG Electronics

Ranked by the volume of their employees' year-to-date open source contributions on GitHub

<https://opensourceindex.io/> 2023.8

Active Contributors : 10 개 이상의 commit을 한 사람 수
Total Community : 1개 라도 commit을 한 사람 수

() 안은 시총 순위, 50위까지, Forbe 2023.5
개발자 수, outtalent.com



01 우리나라, 긍정적 조사

“Open Source 활용 사례가 늘어나고 있으나 (94.2%),
자체적 라이선스 검증절차가 있는 기업은 32.1%”

“Open Source를 단순히 그냥 활용하는 사례가 많다 (47.1%)”

“Open Source 프로젝트에 대해 리뷰나 수정 등으로 적극
참여하는 비율을 27년까지 70% 수준으로 늘릴 예정”



01 중간 요약

기업들의 Open Source 의지는 높다.
AI 영역도 복잡하지만 Open Source는 큰 의미를 가진다.
저작권 중심의 거버넌스보다, 보안이 더 중요해지고 있다.
Cloud 시대, Open Source 산수가 복잡해졌다.

우리나라는 약간의 분발이 필요하다.



02

Open Source, 자연스럽게



02 Open Source, 자연스럽지 않았던 이유

Open Source를 Business 모델로 인식

-- Open Source는 혁신의 방법론

아직 100년도 안된 소프트웨어 산업

-- 아직 100년도 안된 인터넷, 아직 100년도 안된 ...

-- 잘 정의가 안되었던 소프트웨어의 사용 영역 (전쟁, 비윤리적 ...)

그리고, 가치 평가 기준의 빈곤함

-- (기술) Ownership에 대한 자신감/대표성 부족



Open Source, 자연스럽게 해야하는데

커뮤니티가 다 엮는데...

그가 나의 이름을 불러주기 전에는
나는 다만
하나의 몸짓에 지나지 않았다.

그가 나의 이름을 불러주었을 때,
나는 그에게로 가서
Open Source가 되었다.



02 Open Source, 자연스럽게

그런데, 커뮤니티가 그 커뮤니티가 아니다.

기술 Ownership을

개별 기업, 순수 개발자 커뮤니티가 만들기 어려워졌다.

-- 더 이상 Low Hanging Fruit은 없음

Foundation, Consortium, Forum, 국가적 R&D ...

-- 많은 조직이 참여하여, 산업 표준을 만들어 global lock-in 구조를 형성

-- Open Source 보안 문제에 대한 조직적 접근도 필요하다.



02 Open Source, 자연스럽게

그래서, 개발자도 그 개발자가 아니다.

전업으로 Open Source에 기여하는 개발자의 증가

-- 더 많은 조직들이 상업적 소프트웨어 개발을 Open Source로 한다.

개발자의 성장에 기여하는 Open Source 프로젝트가 되어야 한다.

-- Open Source 일수록 더 도전적인 개발 프로세스를 추구해야 하고

-- Open Source 일수록 더 협업과 의사결정을 위한 소통이 잘 되어야 하고

-- Open Source 일수록 더 결과물의 산업적 가치를 중시해야 한다.



Open Source AI, 자연스러울 수 있을까?

30여년 전 Free-OSS 운동 때와의 차이

-- Source 란 무엇인가?

model, training data, user data, weight, ... 는 Source 인가?

-- 볼 수 있다는 것은 무엇인가?

그 Source가 중요한가?

-- 수정은 무엇이고, 배포는 무엇이며, 파생물은 무엇인가?

-- 서비스는 무엇인가?

이미 뒤늦은 법, 특이점을 넘어가고 있는 기술 !!!

OSI의 Open Source 정의는 아직도 유효한가?



Open Source AI, 자연스러울 수 있을까?

현존 Open LLM과 그 언저리의 저작권

Open LLMs

-- Apache 2.0, MIT, OpenRail-M v1.0, CC-BY-SA-3.0/4.0, custom, ...

Open LLMs for Code

-- Apache 2.0, OpenRail-M v1.0, CC-BY-SA-4.0, BSD-3-clause, custom, ...

Open LLM datasets for pre-training, tuning

-- Apache 2.0, CC-BY-SA-3.0, ...



Open Source, 자연스럽게

Open Source라서 성공? 성공한 Project가 Open Source?

사용자의 Needs가 해결되지 않은 프로젝트는 실패한 프로젝트

- 사용자는 조직 밖에 있다. 만나야 한다.
- 사용자는 미래에 있다. 그 미래가 어떨지 이야기해봐야 한다.
- 사용자는 마음이 변한다. 빠른 release와 lean한 프로세스가 중요하다.

사용자가 개발자라면 지금 상태의 코드를 보고싶어 한다.

- 그래서 Open Source여야 한다.
- 이제는 보안 때문에, 모두가 코드도 보고 싶고, 관리도 되는지 알고 싶다.



03

요약



요약 : Open Source, 자연스럽게

대세인 것은 분명하다. 이미 오래전부터

- Open Source를 만드는 사람을 이롭게 한다.
- Open Source를 이용하는 사람을 이롭게 한다.

하지만 혼란의 시기이다.

- 세상도 변했고, 기술 환경도, 사업 환경도 변했다.

자본의 논리가 Open Source를 움직이지만,

- 아직 그 중심에는 개발자가 있다.
- 사용자와 개발자를 설득 못하는 Open Source는 Open Source가 아니다.

잘 하자.

고맙습니다

**ETRI
OPEN
SOURCE
TECH DAY
2023**



Open Source R&D



2장. 공개SW R&D 계획수립의 단계별 검토사항

1. 공개SW R&D 계획수립의 단계별 검토사항 개요

가. 공개SW R&D 계획수립의 검토 단계

나. 공개SW R&D 계획수립의 검토 절차

2. 1단계 : 공개SW R&D 추진전략 및 사업모델 수립단계

가. 공개SW R&D 추진전략 수립

나. 공개SW R&D 사업모델

3. 2단계 : 공개SW R&D 라이선스 정책수립 단계

가. 공개SW R&D에 적용할 공개SW 라이선스 분류기준

나. 사업화를 위한 공개SW R&D의 사업모델에 따른 라이선스 정책검토

다. 비영리 목적의 공개SW R&D의 라이선스 정책검토

라. 공개SW R&D 대표 라이선스 및 호환 라이선스 정책 검토

4. 공개SW 사용정책 및 절차 수립단계

5. 공개SW R&D 기술이전시의 검토사항

가. 공개 R&D 결과물의 기술이전 유형

나. 공개 SW R&D 결과물의 기술이전 유형별 검토사항

공개SW R&D 실무 수행 가이드라인

www.iitp.kr

2022년 7월

