

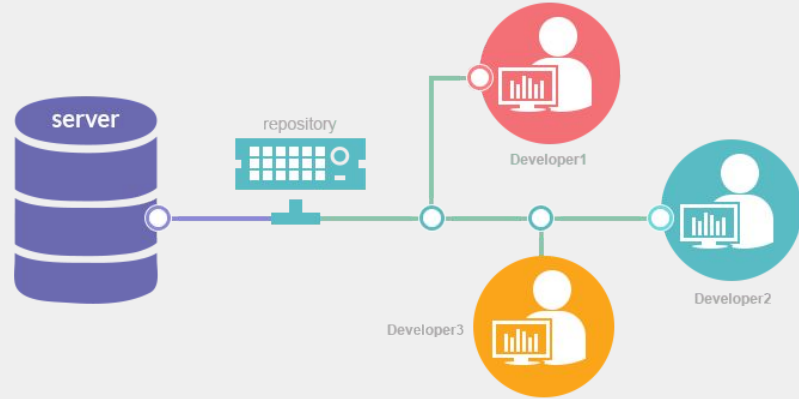
Kaynak Kodları GitHub Üzerinden
Yönetilen ve Debian Depolarında
Bulunan Yazılımların Geliştirme
İstatistikleri

Dilara Koca



Sürüm Kontrol Sistemi

Sürüm kontrol sistemi, dosyalar üzerinde yapılan tüm deęişiklikleri kaydedip sürümler oluşturan ve daha sonra istenilen sürüme geri dönebilmeye olanak sağlayan bir sistemdir.



Merkezi Olmayan Sürüm Kontrol Sistemi

- birçok “**yerel depo**” (local repository)
 - birçok kullanıcı
 - bir tane “**uzak depo**” (remote repository)
- ↗ **Commit**, yerel depoda yapılan değişikliklerin bir mesaj ile birlikte uzak depoya gönderilmek üzere sürüm olarak kaydedilmesi

Git ve GitHub

Merkezi olmayan sürüm kontrol sistemlerinin bir örneđi **Git**'tir.

Github ise, **Git** sürüm kontrol sistemini kullanan yazılım geliştirme projeleri için web tabanlı bir depolama servsidir.



Debian ve .deb Paketi

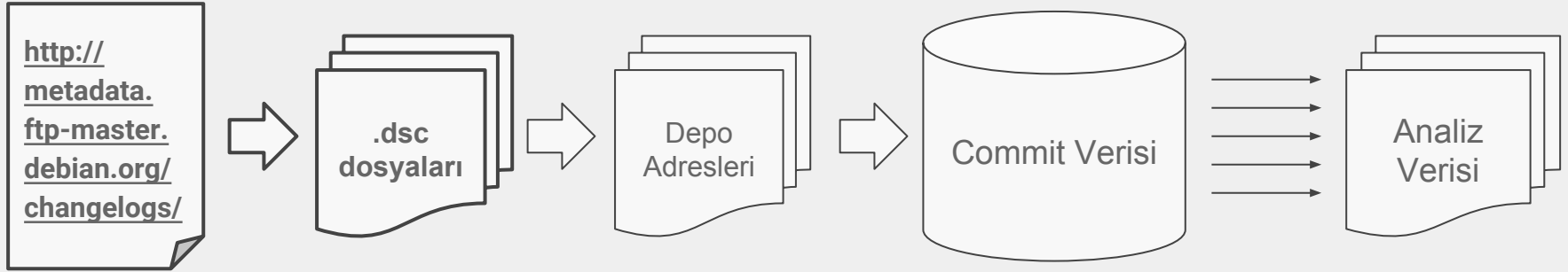
- ↗ **Debian**, Dünya'nın çeşitli bölgelerindeki gönüllüler tarafından hazırlanan; farklı çekirdek seçeneklerine dayalı tamamen özgür bir Linux dağıtımı
- ↗ En yaygın GNU/Linux dağıtımlarından biri

- ↗ Kendine özgü bir paket üslubu olan **.deb** kullanılır

“DSC” Dosyası (Debian Source Control)

- ↗ “.dsc” dosyası, Debian kaynak paketini tanımlayan ayrılmış alanlardan oluşan bir dizi içerir
- ↗ “**Homepage**”, üst geliştirici kaynak paketinin paket ile ilgili bilgi barındıran alan

Dsc Dosyalarının Elde Edilmesi



Dsc Dosyalarının Elde Edilmesi

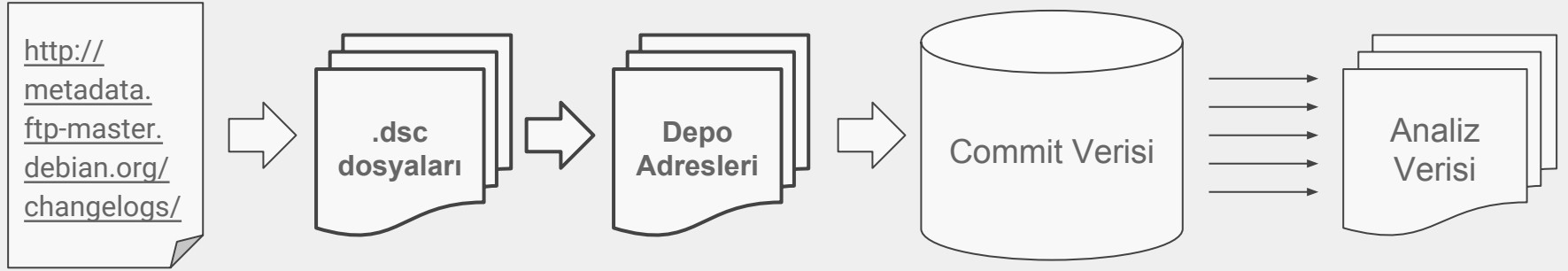
↗ Güncel Debian paket listesi

<http://metadata.ftp-master.debian.org/changelogs/> (26444 adet paket)

↗ Paketin yalnızca .dsc dosyalarının çekilmesi

```
$ apt-get source -dsc-only paket
```

Depo Adreslerinin Elde Edilmesi

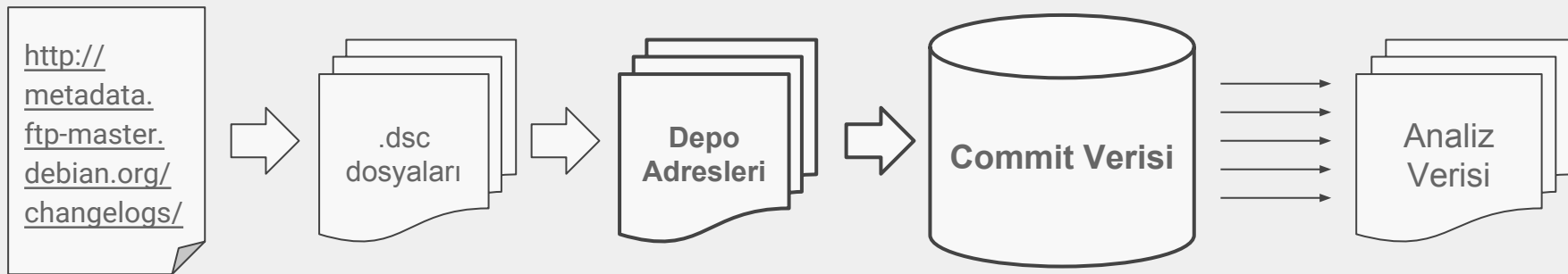


Depo Adreslerinin Elde Edilmesi

Sürüm kontrol sistemi olarak Git ve depolama servisi olarak Github kullanan paketlerin seçilmesi istenilmiştir. Dsc dosyasının “Homepage” alanında üzerinde olduğu adres bilgisini veren paketler, düzenli ifadeler kullanan otomatize bir yapı ile süzölmüştür.

26444 adet paket içinden süzme işlemi sonucu elde edilen paket sayısı 1137'dir.

Commit Verilerinin Üretilmesi



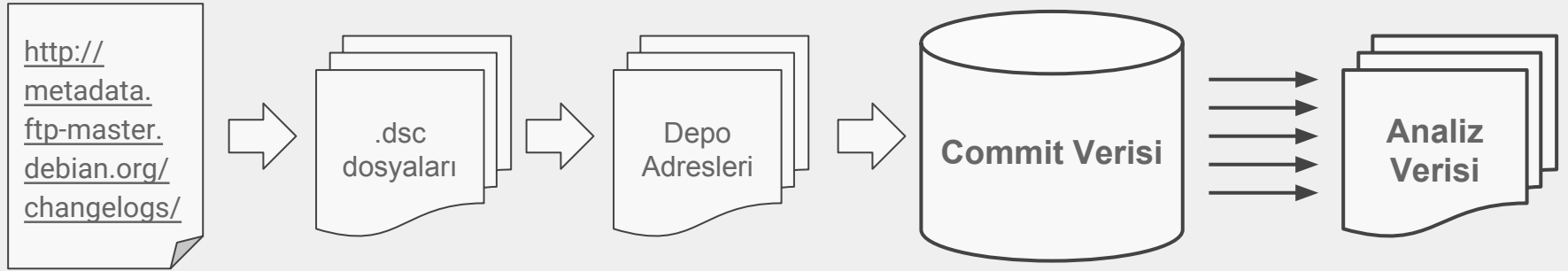
Commit Verilerinin Üretilmesi

↗ Github API

↗ Son 2 yılın commitleri

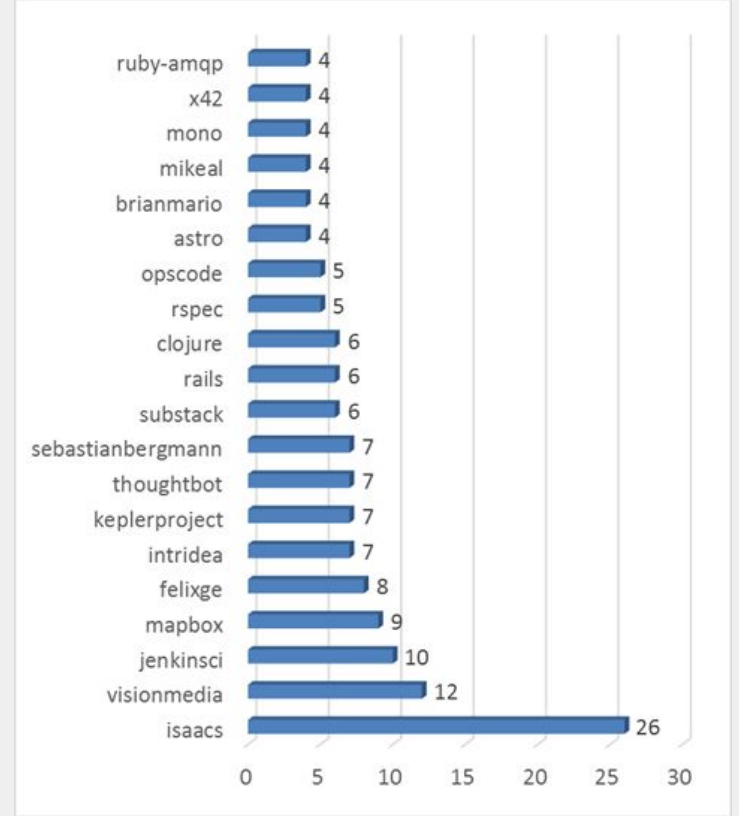
<https://api.github.com/repos/depo-sahibi/depo/commits?Since=2013-08-20>

Analizler



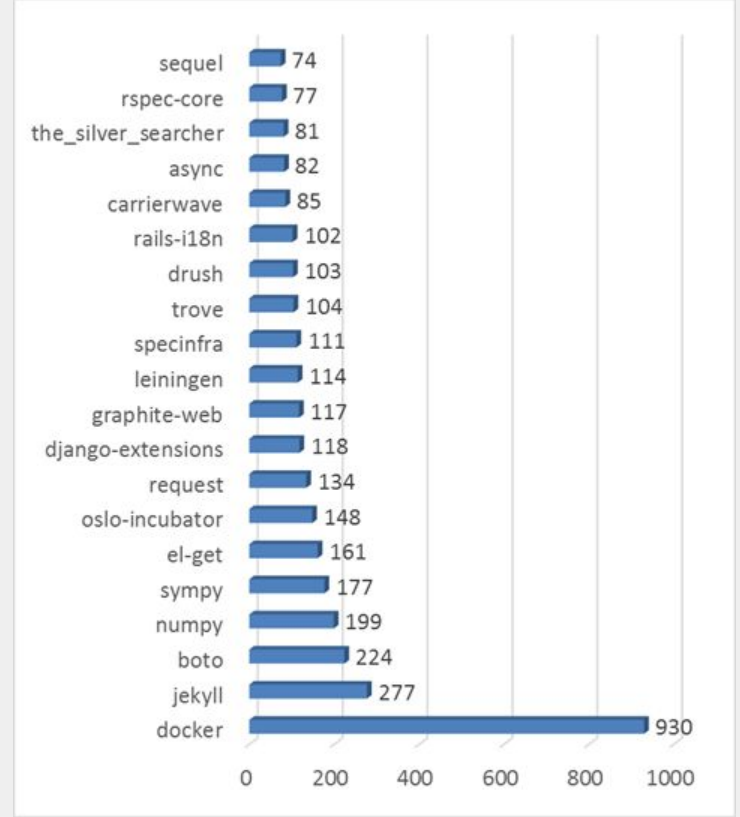
Depo Sahiplerinin Depo Sayıları

Depo sahipleri hakkında yapılan inceleme göz önüne alındığında ortalama 1 (1.38) olan kullanıcı başına düşen depo sayılarının 25'e kadar çıkabildiği gözlenmiştir.



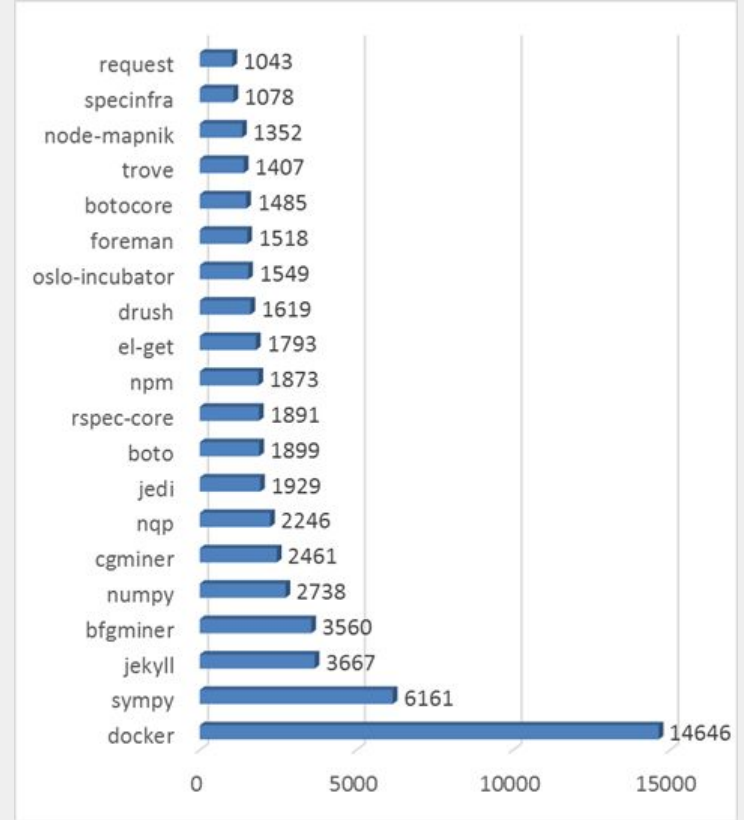
Depoların Commit Gönderen Kullanıcı Sayıları

Depolar hakkında yapılan incelemelere yakından bakıldığında depoya commit gönderen kullanıcı sayılarının ortalama 12 iken bu sayının 930 ile maksimum noktaya “docker” deposunda ulaştığı tespit edilmiştir.



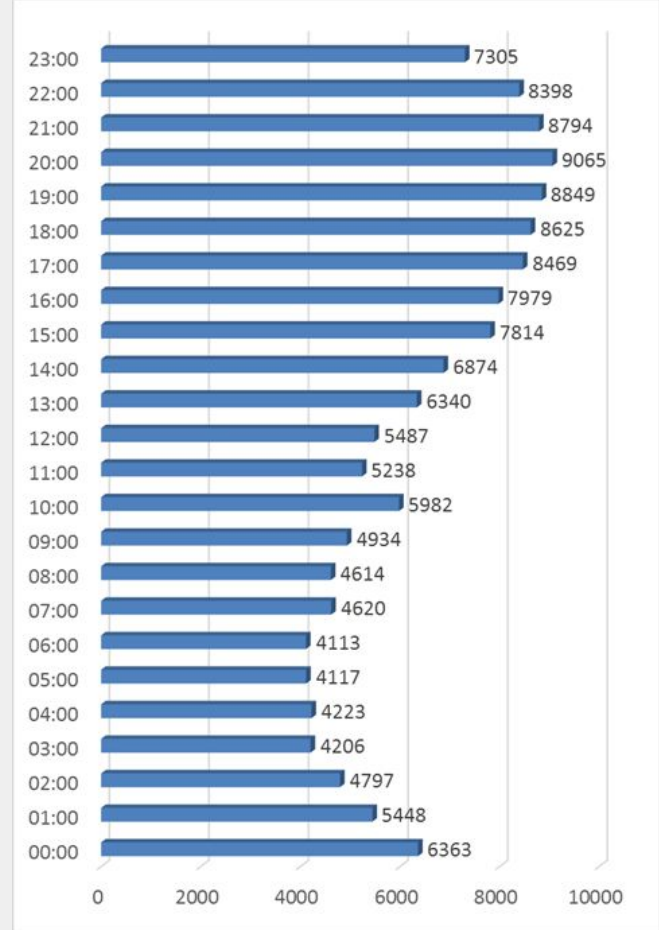
Depoların Commit Sayıları

Depolara yapılan commit sayıları ortalama olarak 159 olarak hesaplanırken 14646 commit sayısı ile “docker” deposu başı çekmektedir.



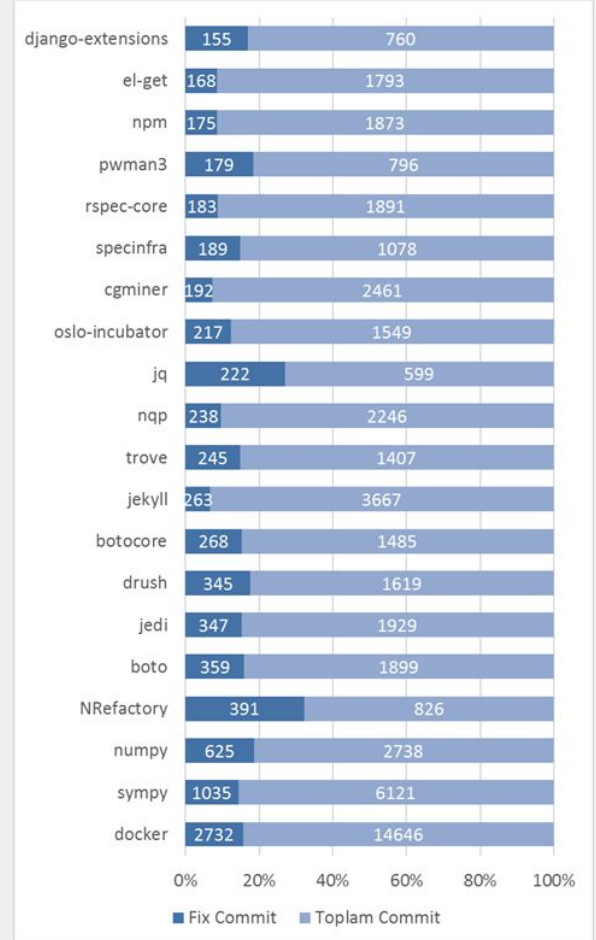
Depoların Saatlere Göre Commit Sayıları

Yapılan bu commitler ağırlıklı olarak 15:00-23:00 saatleri arasında yoğunlaşmaktadır.



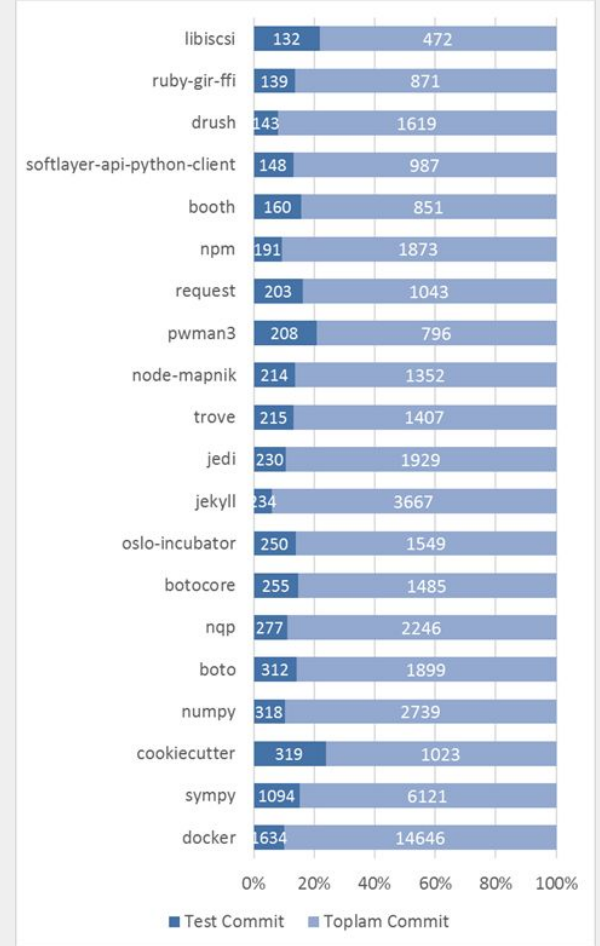
Depoların Fix Commit Sayıları

Bunun yanı sıra depolara yapılan commitlerin ortalama olarak yüzde %15'inin fix commiti olduğu görülmektedir.



Depoların Test ile İlgili Commit Sayıları

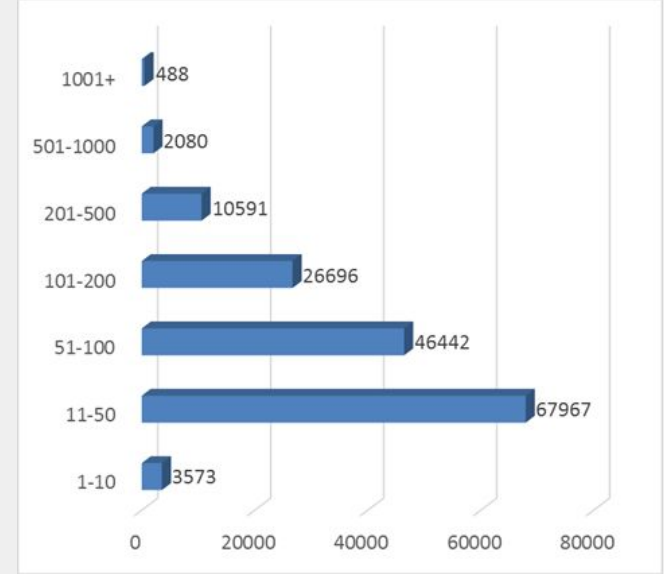
Yazılım geliştirme sürecinde büyük bir yer tutan test geliştirme süreciyle alakalı yapılan commitler analiz edildiğinde test ile alakalı commitlerin sayısının tüm commitlere oranı %9.2 olarak tespit edilmiştir.



Depoların Commit Mesajı Uzunluğu Dağılımı

Geliştirme sürecinde geçmişe dönme, gelişim sürecinin takip edilmesi gibi bir çok alanda kolaylık sağlayan commit mesajlarının anlamlı olması gerekmektedir.

- ↗ Ortalama 92 karakter uzunluğunda
- ↗ %70 oranında ortalama altında
- ↗ %30 oranında ortalama üstünde
- ↗ En uzun 65535 karakter



Kaynaklar

- ↗ Index of /changelogs, Ağustos 2015 (de indirildi) World Wide Web: <http://metadata.ftp-master.debian.org/changelogs/>
- ↗ Rodin J., Aoki O., Debian New Maintainers' Guide, Ağustos 2015 (de indirildi) World Wide Web: <https://www.debian.org/doc/manuals/maint-guide/>
- ↗ Github, Github Developer, Ağustos 2015 (de indirildi) World Wide Web: <https://developer.github.com/v3/>
- ↗ Spinellis, D. (2012). Git. Yazılım, IEEE, 29(3), 100-101.
- ↗ Knott, M. (2014). Version Control with Git. In Beginning Xcode, 371-403. Apress.
- ↗ Shaffer, K. (2013). Push, Pull, Fork: GitHub for Academics. Hybrid Pedagogy.
- ↗ Agrawal, K., Amreen, S., & Mockus, A. (2015). Commit quality in five high performance computing projects. Proceedings of the 2015 International Workshop on Software Engineering for High Performance Computing in Science, 24-29.
- ↗ Nussbaum, L., & Zacchiroli, S. (2010). The ultimate debian database: Consolidating bazaar metadata for quality assurance and data mining. Mining Software Repositories (MSR), 2010 7th IEEE Working Conference, 52-61. IEEE.