

Collaborative Translation Efforts (CTE) Trial Report

Feb, 2019

Vision of CTE

(Tentative, TBD)

More productive translation in “Open Source” way;
with more **Openness**, more **Collaboration**,
and much more **Fun**!

Contents

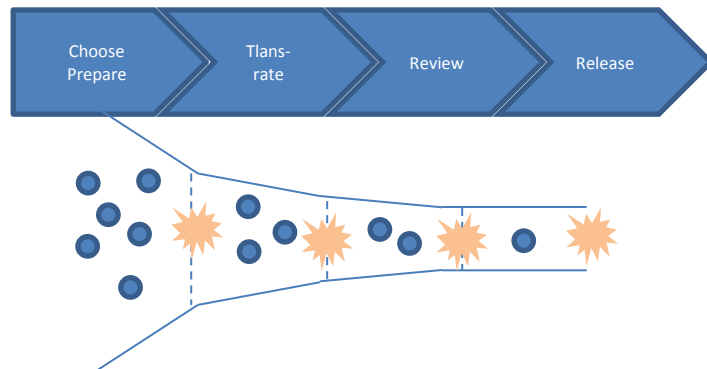
1. Challenges in translation (especially in OSS resources)
2. Aim and approach
3. Trial
4. 1st Trial : By single person.
5. 2nd Trial : with 2 persons collaboration
6. Thought so far
7. Knowledge and feeling
8. Next steps

Challenges in translation (especially of OSS resources)

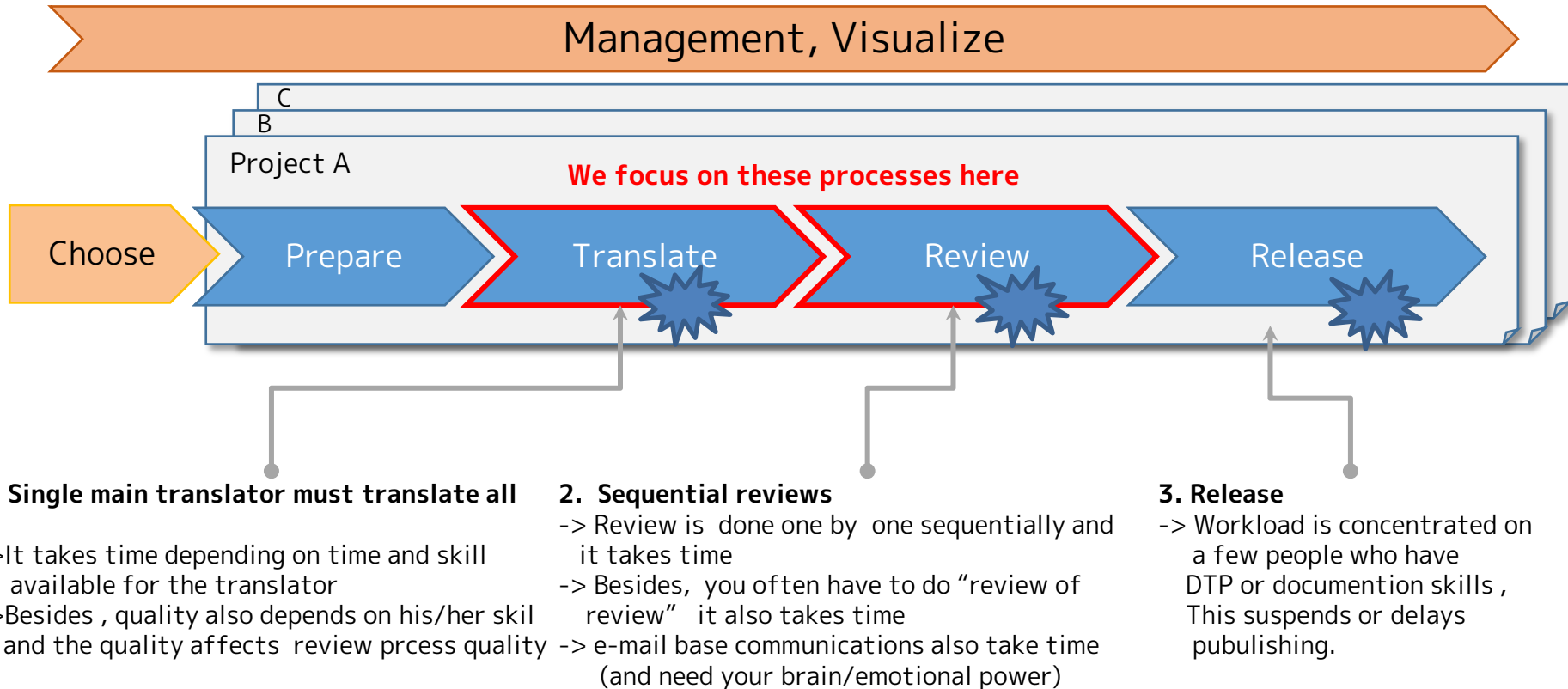
- ◆ **Question of “Quality”**: Dilemma b/w OSS expertise and translation skills
 - The OSS related translation strongly requires OSS expertise such as technology and its culture. It tends to be hard even for professional translators
 - Meanwhile, if you have good enough understanding about OSS, translation quality may not be good enough

- ◆ **Question of “Quantity”**: Unscalable process

- There exist bottlenecks for each process as in; Choose, Prepare, Translate, Review, and Release



Major bottlenecks though traslation procecss



Aim and approach

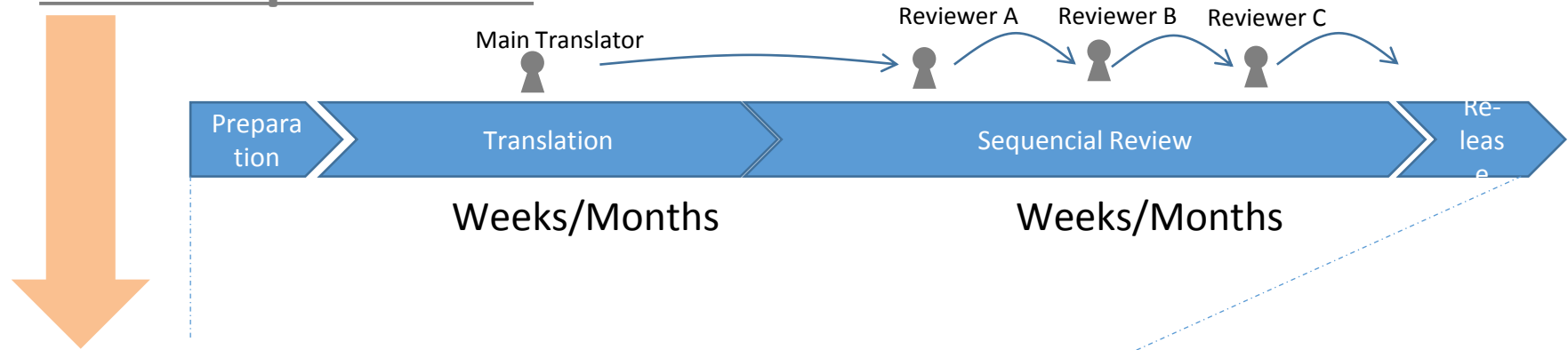
- **Aim:** Focusing on “Question of Quantity” (1 and 2 in the previous page), create virtuous cycle to “Question of Quality”
 - To be “prolific”, above all
 - Invigorate peer review, make it more fun
 - The more collaborators gather, the more prolific (and scalable)
- **Approach:** Automation of Translation process and Review in realtime using tools
 - Tool 1: **Google Machine Translation**
 - > Shorten the time in Translation process, and Focus on Review process
 - Tool 2: **Hackmd**
 - > Simultaneous, realtime review (Edit markdown contents)
 - Tool 3: **Slack**
 - > Simultaneous, realtime review (Communication)

- **Evaluation indicators :**

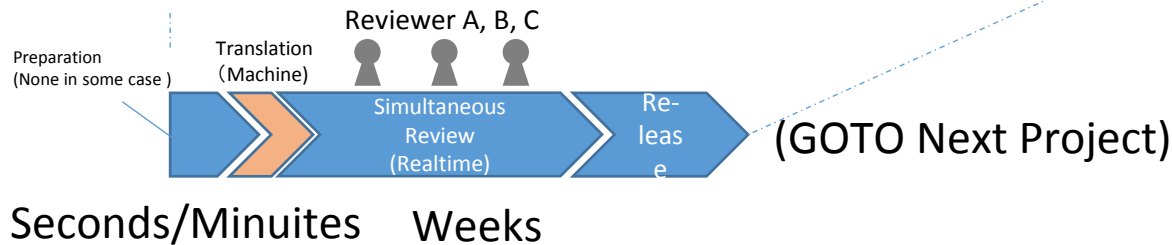
1. Progress rate	2. Efficiency
Words /day	Minuites/(person*word)
Bigger, better	Smaller,better

Aim and approach (Cont'd)

Current processes



CTE approach

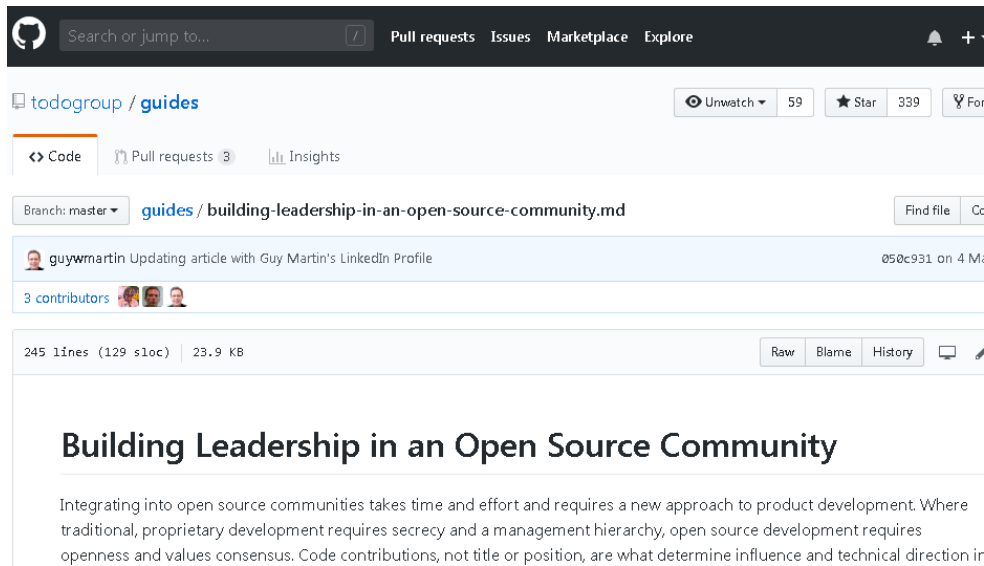


Trials: Basic policy

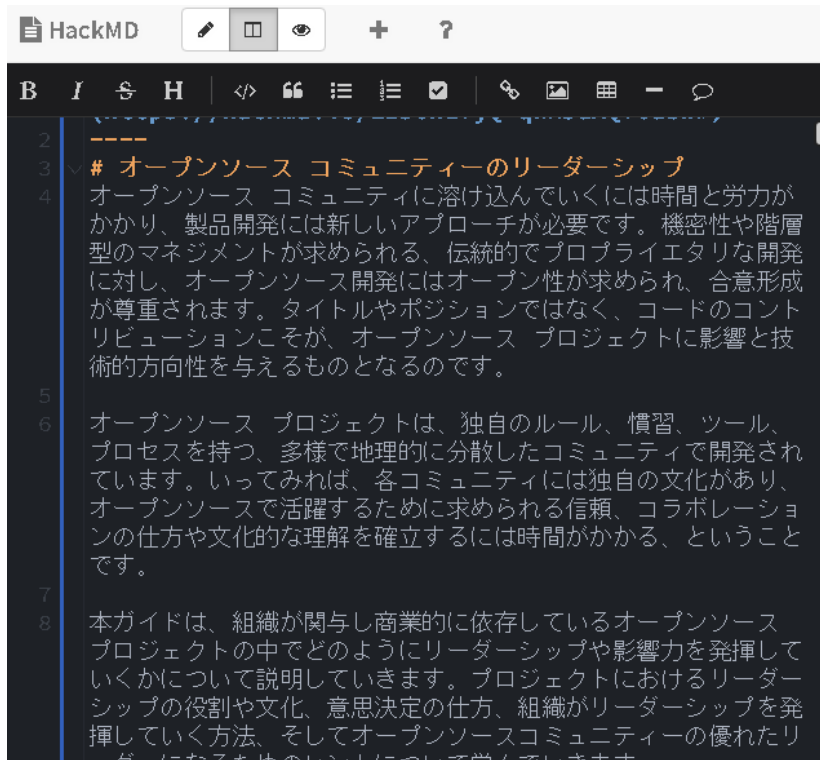
- **Measure and record time in your activity**
 - > This is an important factor for evaluation indicators
 - > NOT necessary to be so accurate
- **Online cross reviews should be done periodically , in short time**
 - > Eliminate the time you think hard on your own.
 - > Do not discuss for long time. Let's make reviews more casual
 - > Abstract and manage ToDo list
- **Online review should be done via Chat tool**
 - > Chat is more casual for many. We need more collaborators!
- **Above all, make it fun!**
 - > Enjoy original contents (Author's idea, thought)
 - > Enjoy interaction (Off topic is also important sometime)
 - > Enjoy progress (We are coming to the goal!)

1st Trial (by 1 person)

- Objective : Measure and evaluate how much single person takes
- Target Resource :
 - TODO Group "[Building Leadership in an Open Source Community](#)"
 - Approx.3600 words (24,000 letters)
- # of person : 1 (Taniguchi)
- Started: 28th Oct, 2018



1st Trial (by 1 person): Work in progress

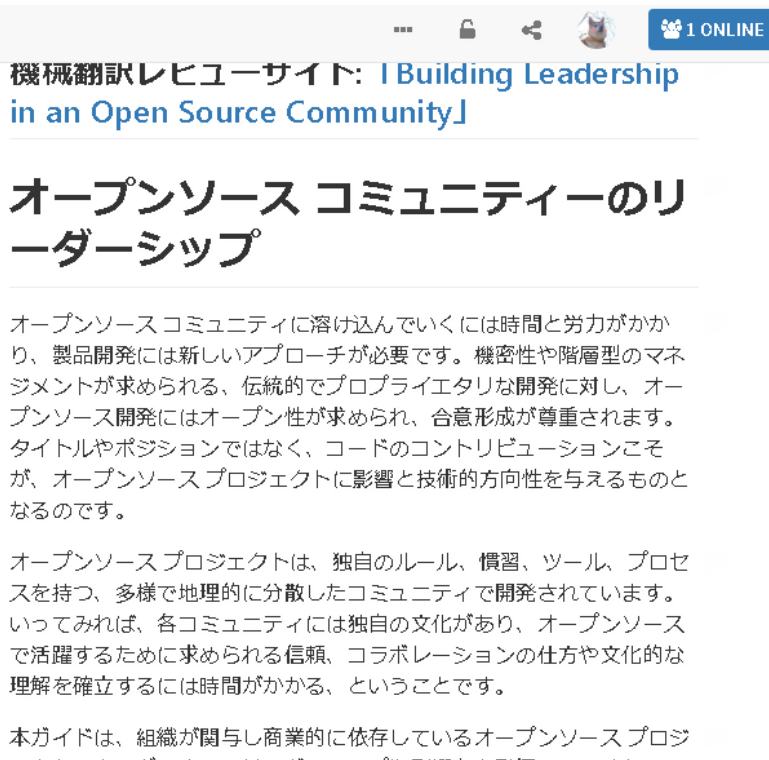


The screenshot shows the HackMD edit pane with a dark theme. The top toolbar includes icons for editing, viewing, and sharing. The document content is as follows:

```
2  ----
3  # オープンソース コミュニティのリーダーシップ
4  オープンソース コミュニティに溶け込んでいくには時間と労力がかかり、製品開発には新しいアプローチが必要です。機密性や階層型のマネジメントが求められる、伝統的でプロプライエタリな開発に対し、オープンソース開発にはオープン性が求められ、合意形成が尊重されます。タイトルやポジションではなく、コードのコントリビューションこそが、オープンソース プロジェクトに影響と技術的方向性を与えるものとなるのです。
5
6  オープンソース プロジェクトは、独自のルール、慣習、ツール、プロセスを持つ、多様で地理的に分散したコミュニティで開発されています。いってみれば、各コミュニティには独自の文化があり、オープンソースで活躍するために求められる信頼、コラボレーションの仕方や文化的な理解を確立するには時間がかかる、ということです。
7
8  本ガイドは、組織が関与し商業的に依存しているオープンソースプロジェクトの中でどのようにリーダーシップや影響力を発揮していくかについて説明していきます。プロジェクトにおけるリーダーシップの役割や文化、意思決定の仕方、組織がリーダーシップを発揮していく方法、そしてオープンソースコミュニティの優れたリ
```

Edit pane(markdown)

*Paste machine translation output here at first



The screenshot shows the HackMD view pane with a light theme. The top toolbar includes icons for editing, viewing, and sharing, along with a '1 ONLINE' indicator. The document content is as follows:

機械翻訳レビューサイト: [Building Leadership in an Open Source Community](#)

オープンソース コミュニティのリーダーシップ

オープンソース コミュニティに溶け込んでいくには時間と労力がかかり、製品開発には新しいアプローチが必要です。機密性や階層型のマネジメントが求められる、伝統的でプロプライエタリな開発に対し、オープンソース開発にはオープン性が求められ、合意形成が尊重されます。タイトルやポジションではなく、コードのコントリビューションこそが、オープンソース プロジェクトに影響と技術的方向性を与えるものとなるのです。

オープンソース プロジェクトは、独自のルール、慣習、ツール、プロセスを持つ、多様で地理的に分散したコミュニティで開発されています。いってみれば、各コミュニティには独自の文化があり、オープンソースで活躍するために求められる信頼、コラボレーションの仕方や文化的な理解を確立するには時間がかかる、ということです。

本ガイドは、組織が関与し商業的に依存しているオープンソース プロジ

View pane

1st Trial (by 1 person): Result

- Outout:
 “オープンソース コミュニティのリーダーシップ”
https://hackmd.io/aibsz3_JTqStRbyTdVO7rA
- Duration: 32days (28th Oct, 2018- 26th Nov, 2018)
 - Translation : Google Machine Translation -> **30 min.** (manual work a little)
 - Review -> **870 min.** (14.5h)
 - Review (correction by tool) -> **30 min.**
 - Release (Image linking) -> **30 min.**
 - **Sum** **960 min.** (16h)
- Evaluation:
 - 1. Progress rate : **112.5 words/day**
 (3600 words/ 32days)
 - 2. Efficiency (Translation + Review) : **0.26 min/(person*word)**
 (930min/1 person/3600 words = 0.26)

Note: Evalutation scores become worse because review must be done more than 1 person essentially

1st Trial (by 1 person): Result(Outcome)

<https://hackmd.io/zgthoZZcTl-s3JXAg1pgkw?view>

The screenshot shows a HackMD document interface. At the top, there's a header bar with the HackMD logo, editing tools (pencil, eraser, eye), a plus sign, a question mark, a menu icon, a lock icon, a share icon, a user icon, and a badge indicating '1 ONLINE'. Below the header, the document title is '機械翻訳レビューサイト: 「Building Leadership in an Open Source Community」'. A 'Watch' button is visible. The main content area has a large heading 'オープンソース コミュニティーのリーダーシップ' followed by a paragraph of Japanese text. On the right side, there's a sidebar with the text 'オープンソースコ...' and links for 'Expand all', 'Back to top', and 'Go to bottom'.

HackMD

CHANGED 2 MONTHS AGO

Watch

オープンソースコ...

Expand all

Back to top

Go to bottom

機械翻訳レビューサイト: 「Building Leadership in an Open Source Community」

オープンソース コミュニティーのリーダーシップ

オープンソース コミュニティに溶け込んでいくには時間と労力がかかり、製品開発には新しいアプローチが必要です。機密性や階層型のマネジメントが求められる、伝統的でプロプライエタリな開発に対し、オープンソース開発にはオープン性が求められ、合意形成が尊重されます。タイトルやポジションではなく、コードのコントリビューションこそが、オープンソースプロジェクトに影響と技術的方向性を与えるものとなるのです。

オープンソース プロジェクトは、独自のルール、慣習、ツール、プロセスを持つ、多様で地理的に分散したコミュニティで開発されています。いってみれば、各コミュニティには独自の文化があり、オープンソースで活躍するために求められる信頼、コラボレーションの仕方や文化的な理解を確立するには時間がかかる、ということです。

本ガイドは、組織が関与し商業的に依存しているオープンソース プロジェクトの中でどのようにリーダーシップや影響力を発揮していくかについて説明していきます。プロジェクトにおけるリーダーシップの役割や文化、意思決定の仕方、組織がリーダーシップを発揮していく方法、そしてオープンソースコミュニティの優れたリーダーになるためのヒントについて学んでいきます。

2nd Trial (2 persons)

- Objective: Evaluate effectiveness by 2 collaborators
- Target resources:
 - LF 「[Certification Preparation Guide](#)」
 - (1) DL site with introduction (104 words, 591 letters)
 - (2) PDF Slide (w/ 21 slides, 4,212 words, 27,799 letters)
- # of person: 2 (Mieko-san and Taniguchi)
 - ※ Besides that, Inou-san from NEC Solution Innovator joined and did technical check as an engineer
- Started : 27th Dec 2018



Certification Preparation Guide

Becoming a Linux Foundation Certified System Administrator or Engineer is no small feat, so to help you with your preparation, we've created this guide to give you some extra assistance. In this guide, you'll find:

- Some critical things to keep in mind on exam day
- An array of both free and paid study resources to help you be as prepared as possible
- A few tips and tricks that could make the difference at exam time
- A checklist of all the domains and competencies covered in the exam

This guide is full of helpful information, and best of all, it's completely free.

DL site



PDF slide

2nd Trial (2 persons) Work in progres

The image shows a HackMD editor window on the left and a Slack chat window on the right. The HackMD editor displays a document titled "# Linux Foundation 認定試験準備ガイド" (Linux Foundation Certification Preparation Guide) dated February 2019. The document content includes a table of contents with sections like "試験の準備5" (Preparation for the exam 5), "試験を受ける14" (Taking the exam 14), and "試験当日のチェックリスト20" (Checklist for the exam day 20). The Slack chat window shows a conversation in the "#translation" channel, with messages discussing the guide and its updates. The chat interface includes a sidebar with channel lists and a main message area with user avatars and timestamps.

HackMD Editor Content:

```
# Linux Foundation 認定試験準備ガイド
2019年2月
The Linux Foundation Training
training.linuxfoundation.org

> [name=Masa Taniguchi]
> 原文は2018年10月のものだが、これからLFCE/LFCS-JPができることを記載するため
> 日本語版はupdateしておく。
---
```

Slack Chat Content:

#translation
2月2日(土)

シングルウォークスルー「1」のこのところの網掛けはやめました^^
くだいなと
ほぼこれで完成版という感じがはいかたしょう

Mieko 12:19
谷口さん、ありがとうございます！！

Mieko 12:29
えーと、
・ p16の右側のリンクが壊れていますね。気づきませんでした。いつからだったんでしょう。http:// だけになっていた。わたしかな。(編集済み)
・ それと、p15の右のトップの1行が左に行くと、もっといいな、と。(編集済み)
・ ちょっと気になっていたことなんですが、裏表紙の文言の1行目が、「ブ」で切れないほうが読みやすいと思う。
どうもうるさくですみません。

Tani 15:17
ファイルは直しますね。

Mieko 15:30
ありがとうございます！！

Tani 15:36
私も挑戦してみようかな

1件の返信 15時間前

Tani 16:59
p15とp16と裏表紙を修正したものととなります。

PDF

LF_Training_WP_CertificationPrepGuide(JP)_pre-

Realtime multi-user editing on Hackmd

Chat on Slack

2nd Trial (2 person): Result

- Output:

“Linux Foundation 認定試験ガイド”

(1) DL site

<https://hackmd.io/LCiwQGyfQKysQBVQDj61kw>

(2) PDF Slide

<https://github.com/maabou512/CertPrepGuide>

- Duration :

(1) **Half a day** (27th Dec 2018 - 27th Dec 2018)

- Translation: Google Machine Translation -> **5 min.**
- Review -> **45 min.**

Sum 50 min.

(2) **32 days** (28th Dec, 2018 – 28th Jan, 2019) (Incl. Release process by InDesign DTP)

- Translation: Google Machine Translation -> **30 min.** (mainly Copy and Paste)
- Review (Self review) -> **1140 min.** (19h : 609min for Taniguchi), 540 min for Mieko-san)
- Review (Online cross-review) -> **550 min.** (5 times)
- Release (DTP by InDesign) -> **420 min.**

Sum 2140 min. (35h)

- # of ToDo in online review: approx. 20 (closed all)

2nd Trial (2 persons) :Result(cont'd)

- Evaluation

(1) DL site

- Progress rate : 240 words / day
(104 words/0.5 day)
- Efficiency (Translation + Review) : 0.24 min / (person*word)
(50 min/2 persons /104 words=0.24)

(2) PDF Slide

- Progress rate : 131 words / day
(4,212 words/32days*)
- Efficiency (Translation + Review) : 0.20 min. / (person*word)
(1,720 min/2 persons/4,212 words=0.20*)

* Note: Actually, above scores include "Release" activity, so scores must be (much) better

2nd Trial: Result (Outcome)

DL site introduction



The screenshot shows a HackMD document interface. The title is 'Linux Foundation 認定資格取得のための準備ガイド'. The content includes an introduction to the Linux Foundation certification and a list of preparation items.


Linux Foundation 認定資格取得のための準備ガイド

Linux Foundationの認定を受けたシステム管理者(Linux Foundation Certified System Administrator) またはエンジニア(Linux Foundation Certified Engineer)になることは簡単なことではありません。そのため、皆さんの準備のため、このガイドでさらなるサポートを提供しています。このガイドでは、次のことをまとめています。

- 試験日の重要な留意事項
- 事前準備となる無料および有償の学習リソース
- 試験で差が出るヒントと秘訣
- 出題範囲とコンピテンシーのチェックリスト

このガイドには非常に有益な情報が豊富にあります。何よりも、フリーです。

PDF Slide (Prep Guide)



The PDF slide features a blue background with a network diagram of interconnected nodes. The title is 'Linux Foundation 認定試験準備ガイド' and it is dated '2019年2月'. It is a publication of The Linux Foundation Training.

The Linux Foundation 認定試験の準備ガイドへようこそ。

本ガイドは、(認定取得に必要なすべてについて) 網羅的な情報提供をするものではなく、また、ここで書かれている内容をすべて活用することで Linux Foundation 認定試験の合格を保証するものでもありません。

Linux Foundation 認定試験準備ガイド
2019年2月

A publication of The Linux Foundation Training
training.linuxfoundation.org

www.linuxfoundation.org

Trials Round-up

	1 st Trial	2 nd Trial (1)	2 nd Trial (2)
# of Person	1人	2人	2人
Words	3,600	104	4,212
Translation time [min.] (Machine Translation)	30	5	30
Review time [min.]	960	45	1690
Result 1. Progress rate [words/day](bigger better)	113* ¹	240* ²	131* ²
Result 2. Efficiency [min./((person*word))] (smaller better)	0.26* ¹	0.24* ²	0.20* ²
Special notes	-	-	Abstracted approx.20 ToDos (All closed)

* 1 : Evalutation scores become worse because review must be done more than 1 person essentially

* 2 : Acutually , above scores includes "Release" activity , so scores must be (much) better



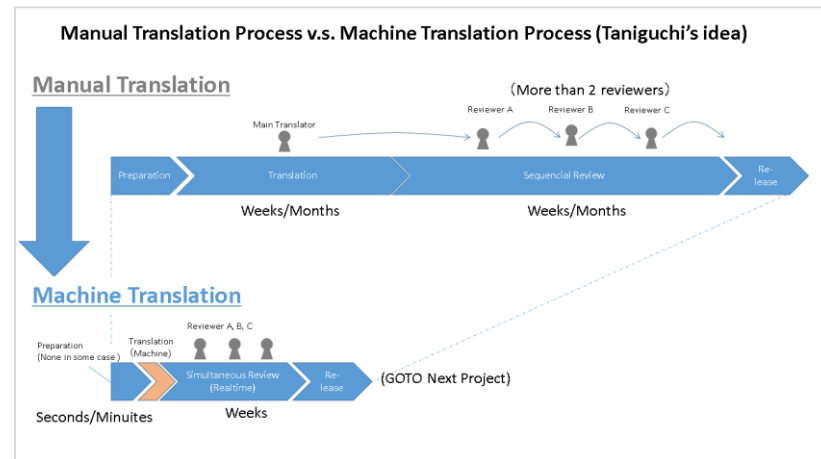
Thoughts (so far)

- This approach (Machine Translation + Realtime cross review) has possibility about;
 - **Shorter leadtime** (Observed 1 -> 2 person)
 - **More efficiency** (ditto)
 - And **Quality improvement** (More issues and ToDos can be abstracted and tackled on accurately)



So, it can be expected that the more we collaborate, the better translation in terms of quickness, efficiency, and quality.

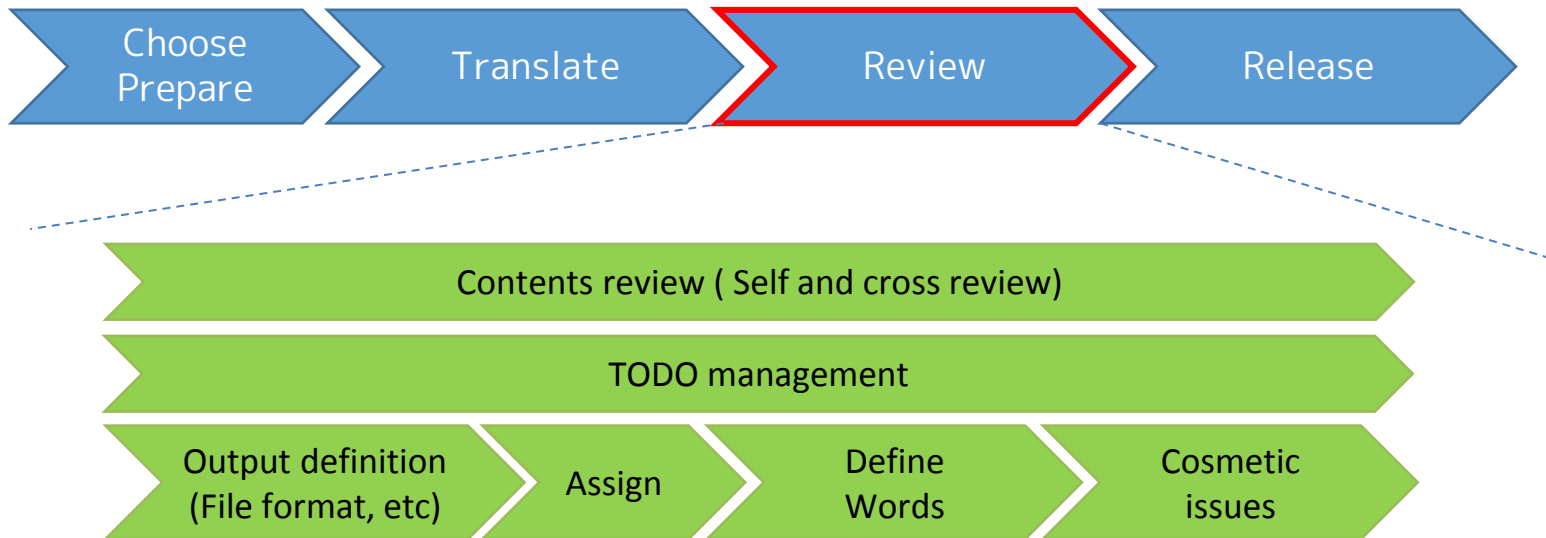
The diagram we intended in P7 (see right) is not comparable directly, but it can be effective indirectly.



Knowledge and Comments

[About processes]

- **Ease of review.** We don't need to take "who translate this?" in consideration. Machines did it, you don't care of who did.
- Review process can be broken down like below (just idea basis)



Knowledge and Comments

[Google machine translation]

- Very interesting translation experience. (e.g . Same original sentences not always to be translated into same translation/It does mistake, but not like human)
- Basically , Quality is very low, **99%** of its output must be modified ultimately.
- **Characteristics of Machine Translation**
 1. Context is not understood semantically, so sometimes it can be even unreadable
 2. Subjects are mistaken or missing sometimes
 3. "De-Aru" and " Desu-Masu" style often mixd
 4. Very weak about "metaphor"
 5. Expressions are not coherent though a material
 6. Errors human beings never do

Next steps (Incl. Issues)

- **Improve processes** -> 3rd Trial ? What if 3 – 4 persons?
 - Shorten leadtime (Plan in advance and book members' schedules)
 - More efficient (Automate progress measurement, etc.)
 - Quality improvement (Create Check list, and so on)
 - And above all , much more fun!
- **Improve tools**
 - **Hackmd** -> Enhance evaluation indicator (Work time measurement) , Chat addon, Response improvement (Cloning Tokyo region), etc.
 - **Slack** etc. -> Function for ease of ToDo mgmt.
 - **Indesign** -> Realtime design check (not supported so far)
- **How to improve Machine Translation's quality?**
 - We "teach" good translation in OSS domain.
 - Can we have so-called "Reinforcement Learning" process?

Other Resources

- 「 Collaborative Translation Effort (CTE) 2nd Trial Spec (Japanese)」
<https://hackmd.io/xhHAQBZcTmmgEGRNgrzNhQ>

→ Motivation and background about this trial and how to do it