Crowd Sourcing Translations

A Look at How Apache Cordova's Documentation is Available in Multiple Languages

Lisa Seacat DeLuca

ldeluca@us.ibm.com @LisaSeacat

www.lisaseacat.com

ApacheCon

April 7th 2014

© 2014 IBM Corporation

Goal

- Make Apache Cordova ubiquitously accessible to people of all language backgrounds

Machine Translation

- “On a basic level, MT performs simple substitution of words in one natural language for words in another, but that alone usually cannot produce a good translation of a text because recognition of whole phrases and their closest counterparts in the target language is needed” - http://en.wikipedia.org/wiki/Machine_translation

Crowd Translation

- “Crowd-assisted translation refers to employing large numbers of bilingual human translators”

Transifex

getLocalization

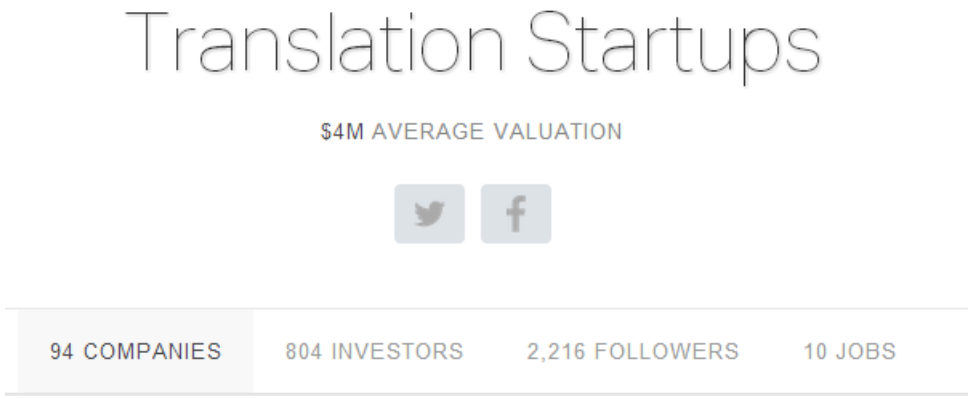
crowdin

STAR TS Translation Services

Confidence in a Translated World.

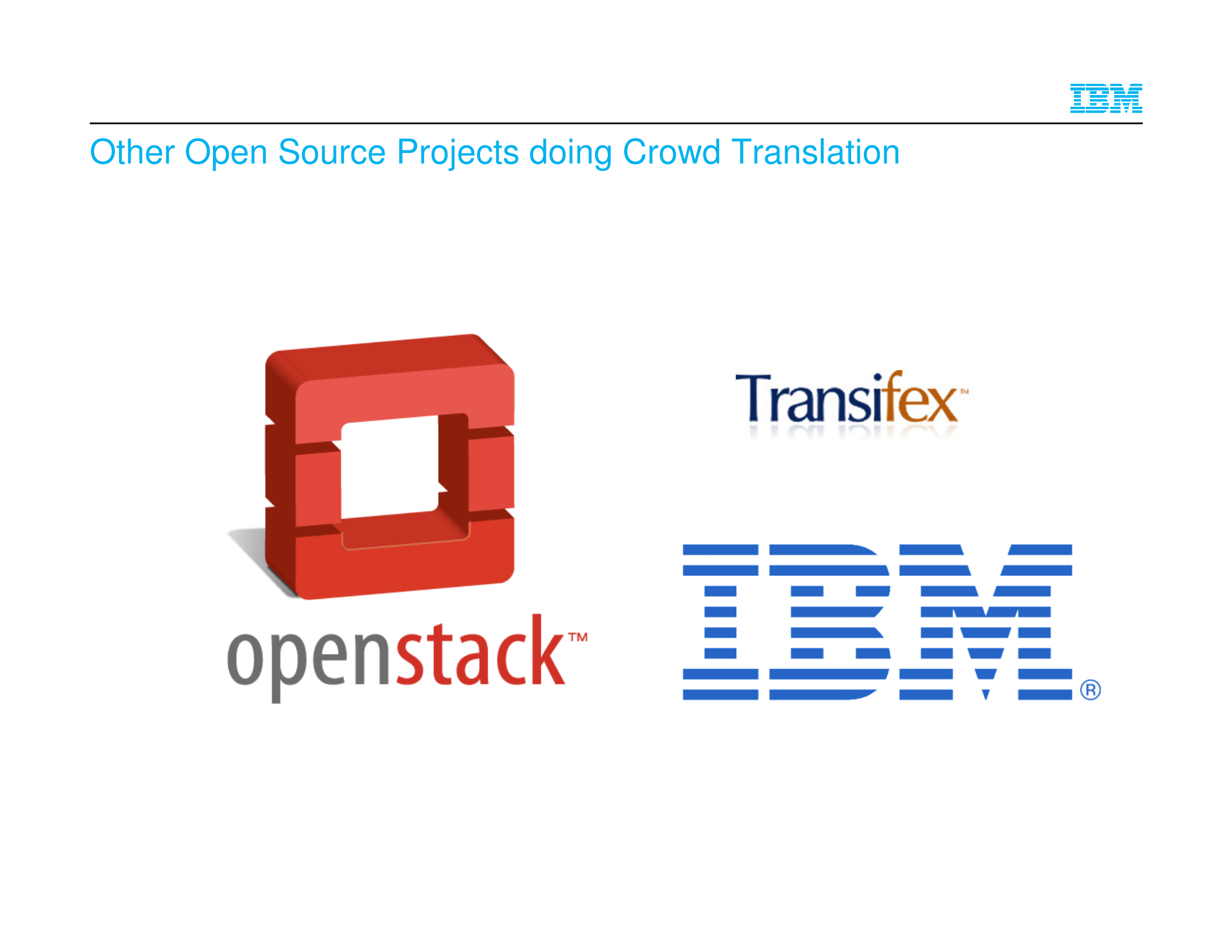
Translation Startups

- https://angel.co/translation
  - $4.2M Average Valuation
  - 94 companies
  - 804 investors



Free Open Source and Academic Licenses

Unlimited projects, unlimited strings and collaborators


IBM
Other Open Source Projects doing Crowd Translation
openstack™
Transifex™
IBM®

1. Markdown (.md) file support
2. Free for open source projects

crowdin.net/project/cordova

Projects / cordova

cordova

Apache Cordova is a set of device APIs that allow a mobile app developer to access native device function such as the camera or accelerometer from JavaScript. Combined with a UI framework such as jQuery Mobile or Dojo Mobile or Sencha Touch, this allows a smartphone app to be developed with just HTML, CSS, and JavaScript.

Owner:

Lisa Seacat DeLuca
ldeluca
Contact

Source language:

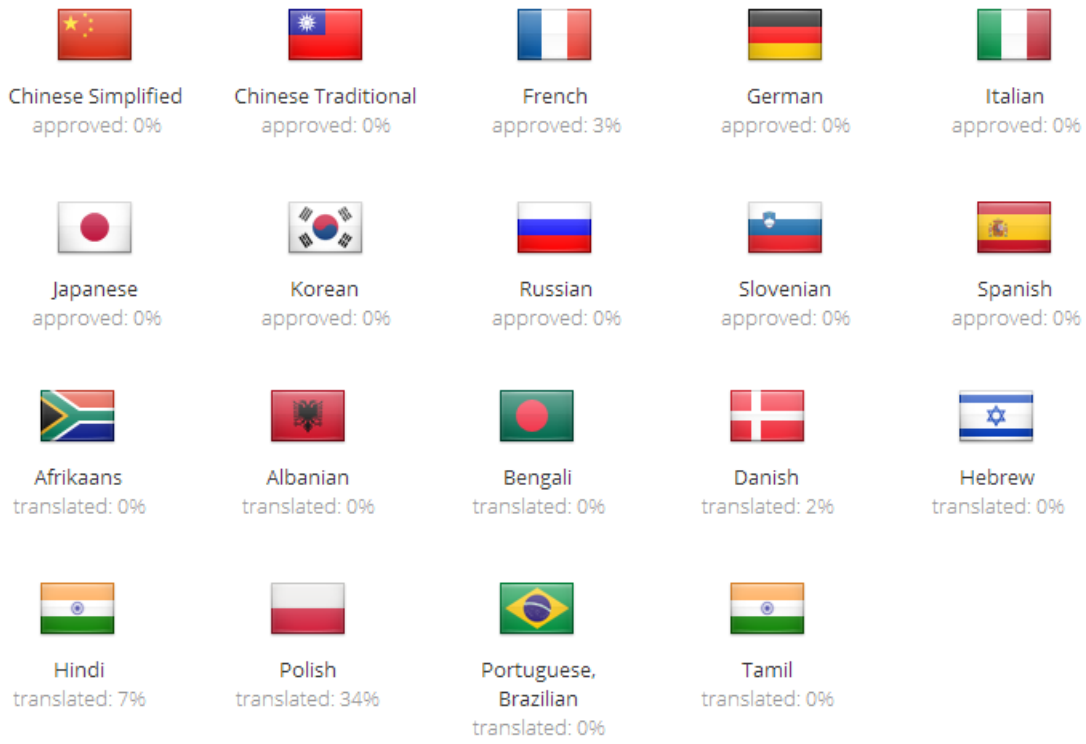
English

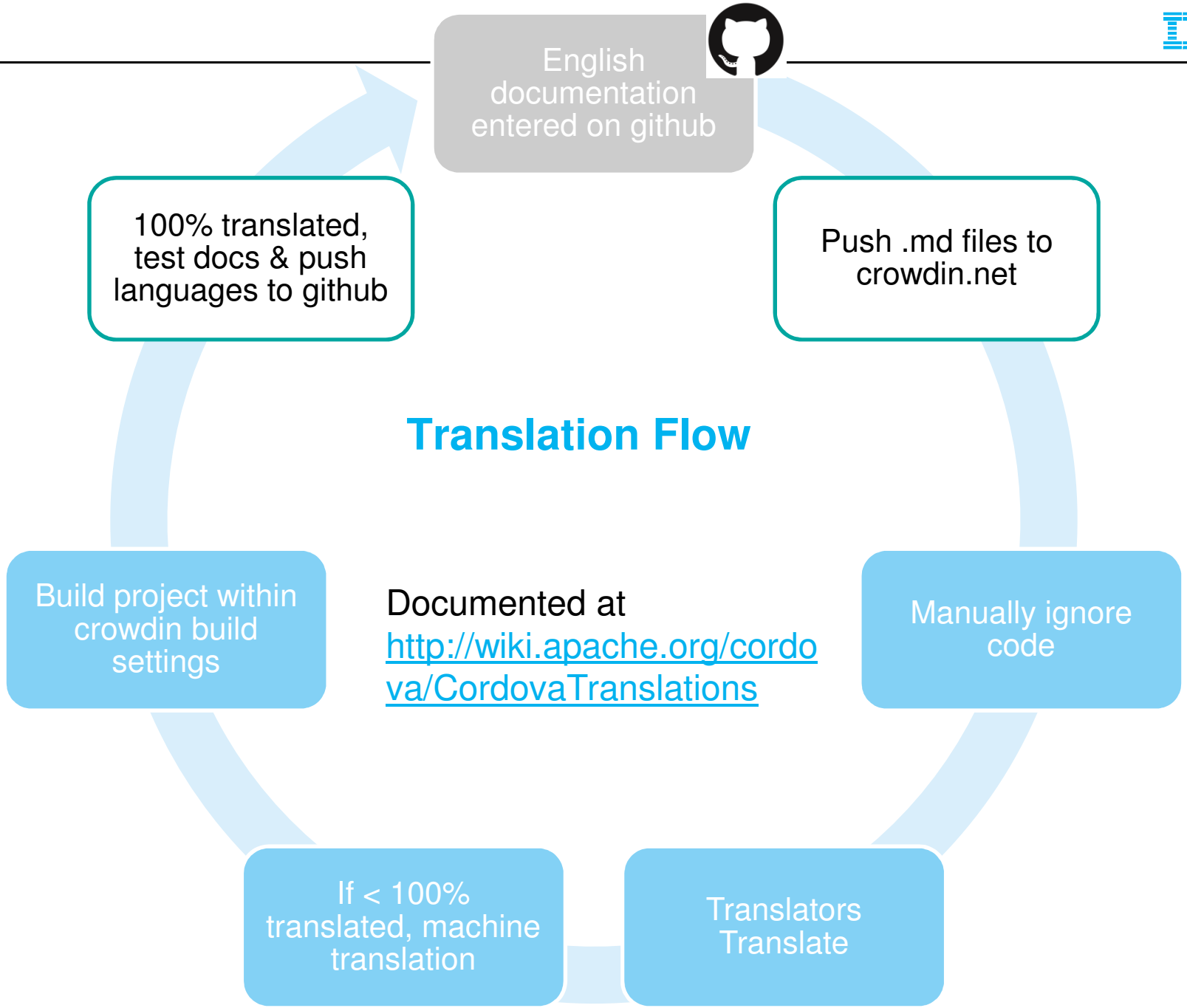
Created: 7 months ago
Last Activity: 46 minutes ago

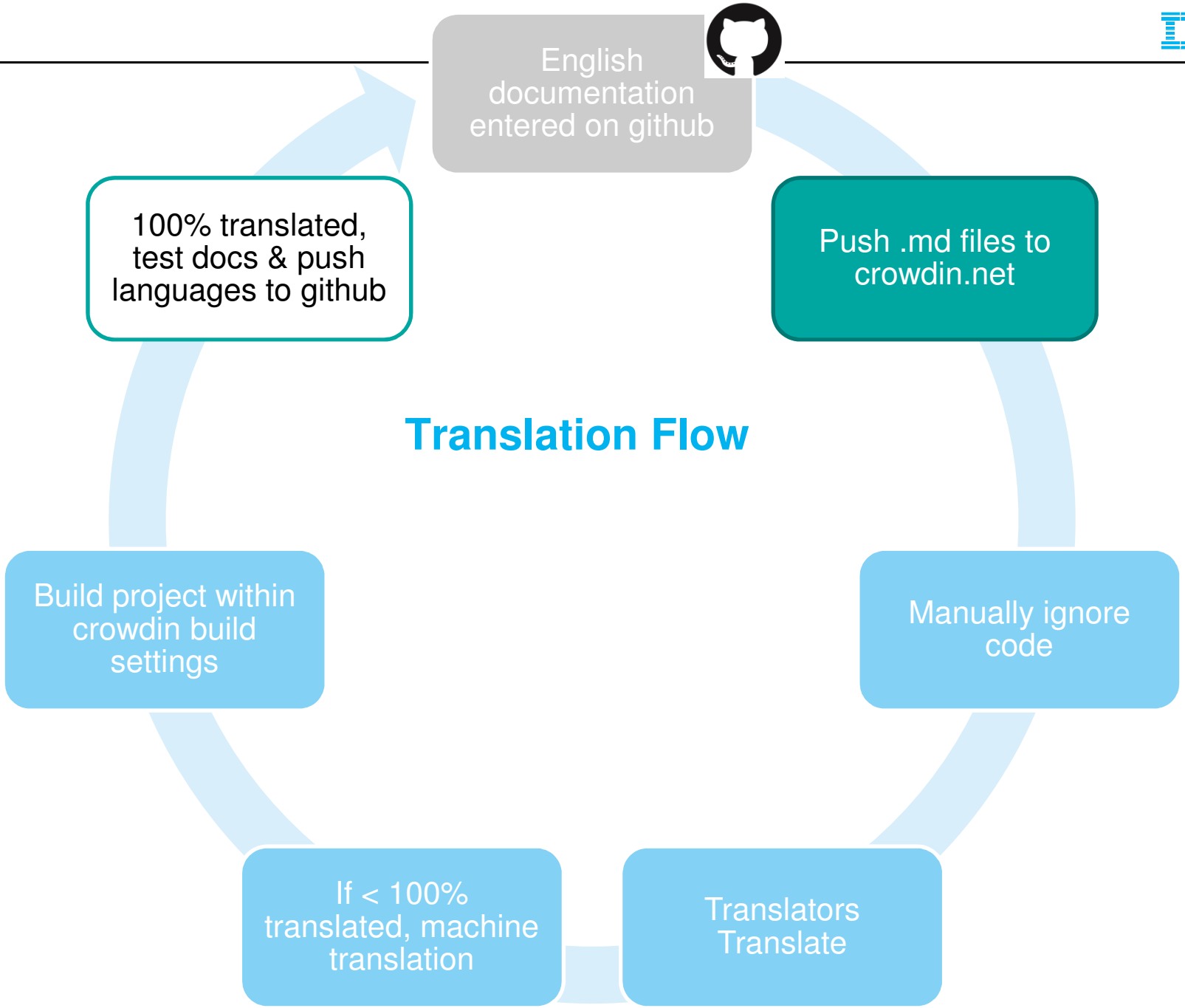
75 users participate in this project

Crowdin Stats

- 75 users
- 72 markdown files
- 10 Primary Languages
- 9 Additional Languages




English documentation entered on github
Push .md files to crowdin.net
Manually ignore code
Translators Translate
If < 100% translated, machine translation
Build project within crowdin build settings
100% translated, test docs & push languages to github
Translation Flow
Documented at
http://wiki.apache.org/cordova/CordovaTranslations


English documentation entered on github
Push .md files to crowdin.net
Manually ignore code
Translators Translate
If < 100% translated, machine translation
Build project within crowdin build settings
100% translated, test docs & push languages to github
Translation Flow

Automation Script

- Pulls mardown (.md) files from github fork
- Pushes to crowdin

```
#FIRST make sure fork is up to date

$ git checkout master
$ git pull apache master
$ git push origin master

#next, move into a branch that is tied to a JIRA issue for pulling the latest

$ git checkout -b CB-1234
$ git branch ----------this lists which branch you are on to double check
```

Administrative Scripts

- Crowdin CLI
  - Jar file must be installed
  - crowdin.yaml

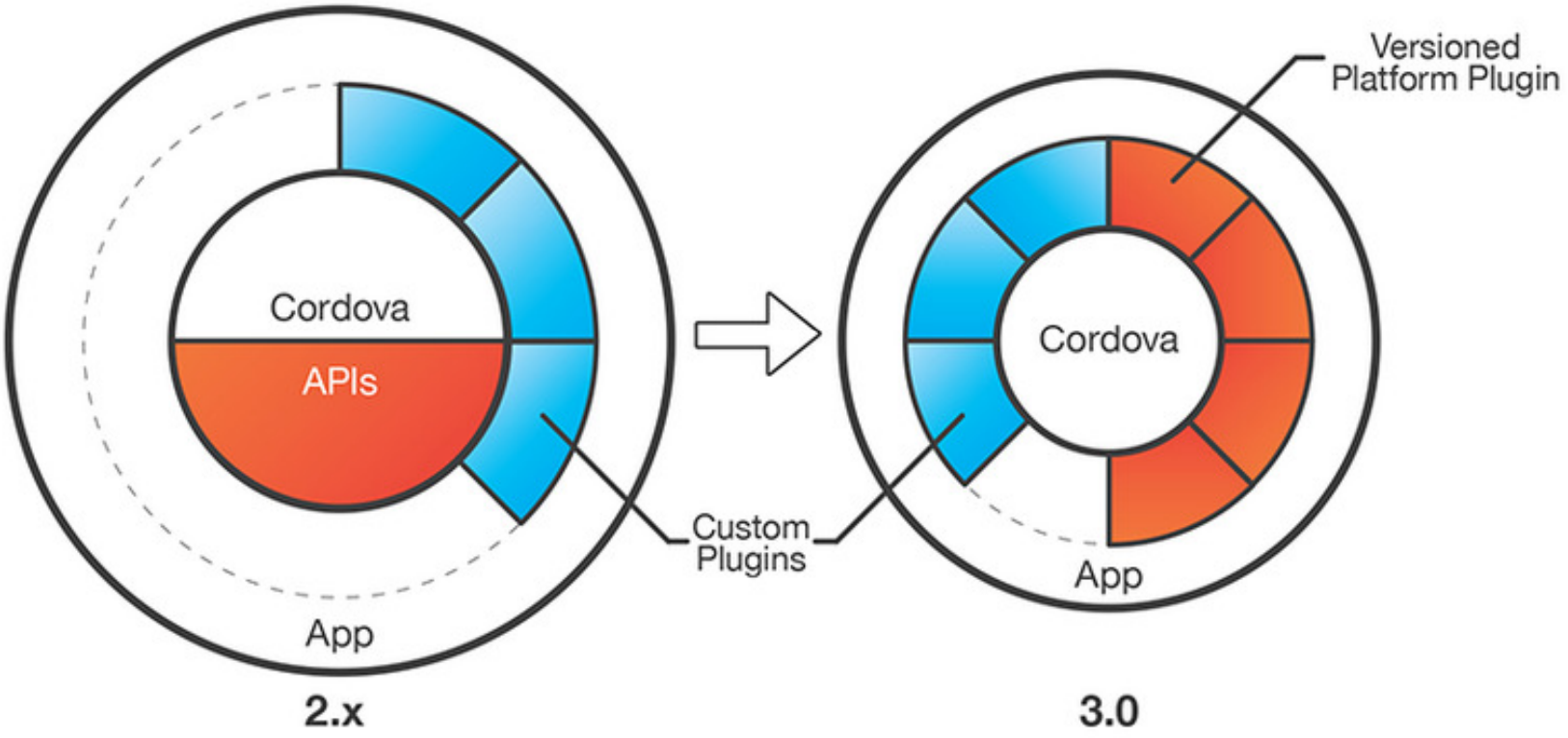
```
project_identifier: cordova
api_key: _____ourAPIkey_________
base_path: /home/<username>/git/cordova-docs #working copy path

files:
   -
         source: "/docs/en/edge/**/*.md"
         translation:
"/docs/%two_letters_code%/edge/**/%original_file_name%"
         ignore:
                  - /.git
```

- github-crowdin.sh

```
DOMAIN_NAME='http://api.crowdin.net'

#---CHANGE THE VARIABLES BELOW---
GIT_REPO_PATH=/home/<username>/git/cordova-docs
CROWDIN_CLI_PATH=/path/to/crowdincli
PROJECT_IDENTIFIER='cordova'
PROJECT_KEY='___________same API key as above________'

#---

cd $GIT_REPO_PATH
git pull origin master
git push origin

cd $CROWDIN_CLI_PATH

java -jar crowdin-cli.jar upload sources

curl $DOMAIN_NAME/api/project/$PROJECT_IDENTIFIER/status?key=$PROJECT_KEY > result.xml

read_dom () {
    local IFS=\>
    read -d \< ENTITY CONTENT
}

while read_dom; do
      if [[ $ENTITY = "code" ]] ; then
        code=( "${code[@]}" "$CONTENT" )
      fi
      if [[ $ENTITY = "translated_progress" ]] ; then
        progress=( "${progress[@]}" "$CONTENT" )
      fi
done < result.xml

for (( i = 0; i < ${#progress[@]}; i++ )); do
   if [ "${progress[$i]}" = "100" ]; then
      index=( "${index[@]}" "$i" )
   fi
done

for element in "${index[@]}"; do
    java -jar crowdin-cli.jar download -l ${code[$element]}
done

# fix crowdin issues:
echo "About to fix crowdin errors with resulting files"
find /home/<username>/git/cordova-docs/docs -name \*.md -exec sed -i "s/\* \* \*/---/1" {} \;
find /home/<username>/git/cordova-docs/docs -name \*.md -exec sed -i "s/## under the License./   under the License.\n---/g" {} \;
echo "Done with crowdin fix"

cd $GIT_REPO_PATH
  git add .
  git commit -m 'Synchronization with Crowdin'
  git push origin
```

French fr 6076 6076 208 100 3 Spanish es-ES 6076 6076 23 100 0 Afrikaans af 6076 0 0 0 0 Danish da 6076 152 0 2 0 German de 60766076 0 100 0 Hebrew he 6076 46 0 0 0 Italian it 6076 6076 0 100 0 Japanese ja 6076 6076 0 100 0 Korean ko 6076 6076 0 100 0 Polishpl 6076 2117 0 34 0 Russian ru 6076 6076 2 100 0 Slovenian sl 6076 6076 0 100 0 Albanian sq 6076 0 0 0 0 Chinese Simplified zh-CN6076 6076 0 100 0 Chinese Traditional zh-TW 6076 6076 0 100 0 Portuguese, Brazilian pt-BR 6076 42 0 0 0 Tamil ta 6076 0 0 0 0 Bengalibn 6076 34 0 0 0 Hindi hi 6076 455 0 7 0

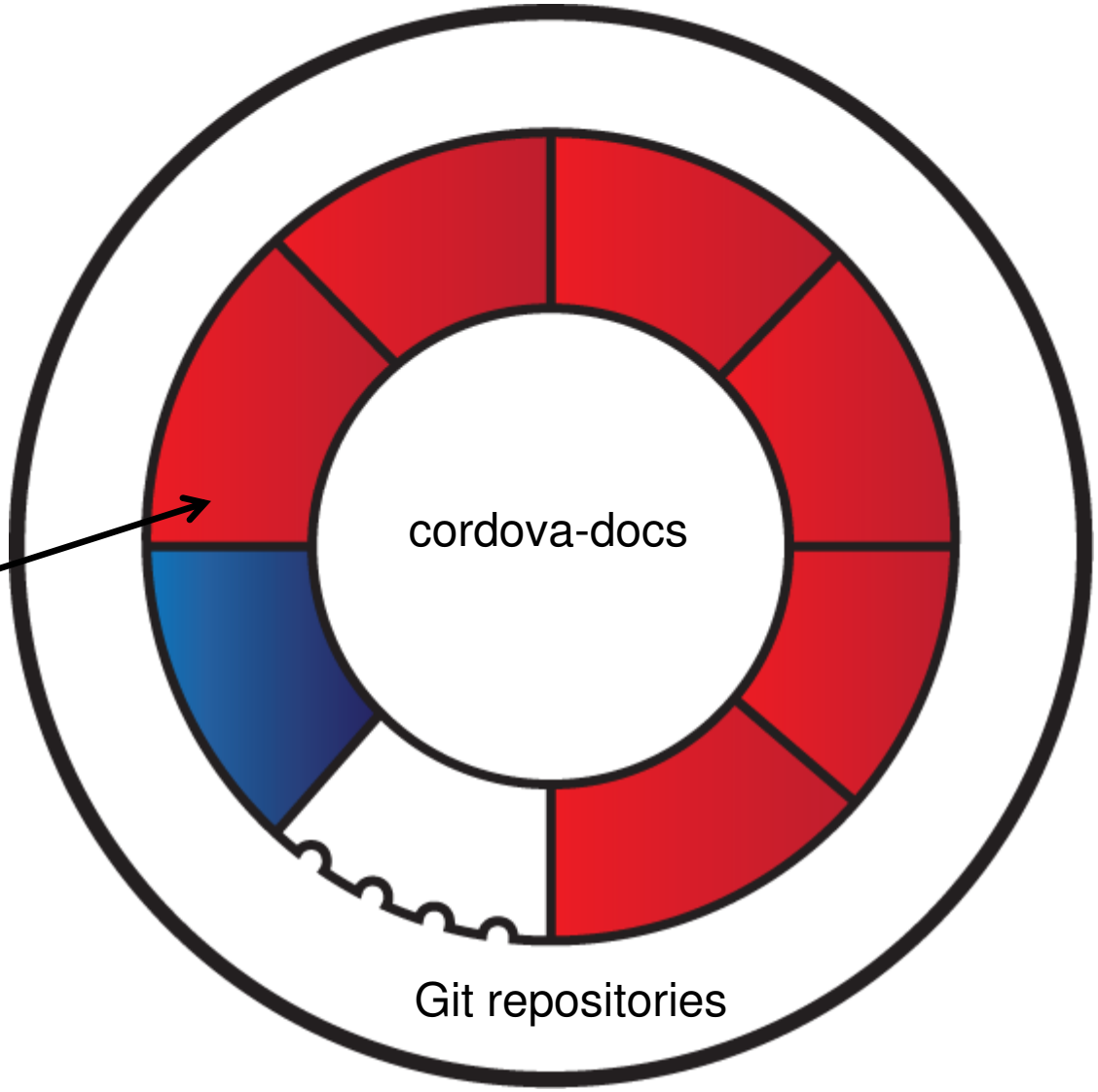
Cordova Plugin Change in Version 3.x



Cordova documentation git repositories

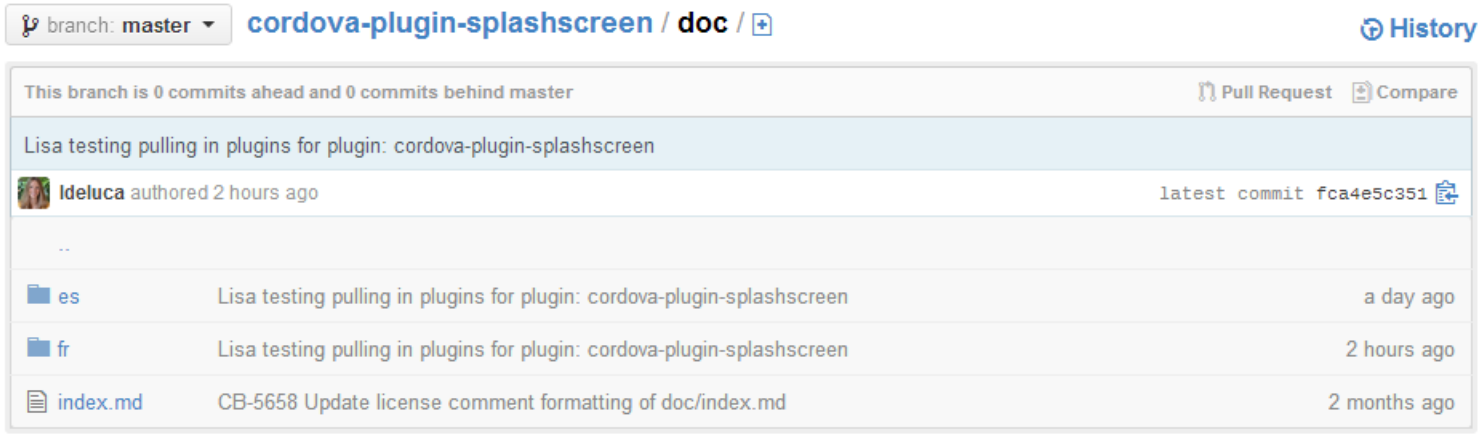
cordova-plugin-*

- Battery status
- Camera
- Contacts
- Device
- Device motion
- Device orientation
- Dialogs
- File
- File transfer
- Geolocation
- Globalization
- In app browser
- Media
- Media capture
- Network information
- Splashscreen
- Vibration



.yaml Update

```
project_identifier: cordova
api_key: ____ourAPIkey_________
base_path: /home/<username>/git #working copy path

files:
  -
    source: "/cordova-docs/docs/en/edge/**/*.md"
    translation: "/cordova-
docs/docs/%two_letters_code%/edge/**/%original_file_name%"
    ignore:
      - /.git
  -
    source: "/**/doc/*.md"
    translation: "/**/doc/%two_letters_code%/%original_file_name%"
    ignore:
      - /.git
```



.sh Update

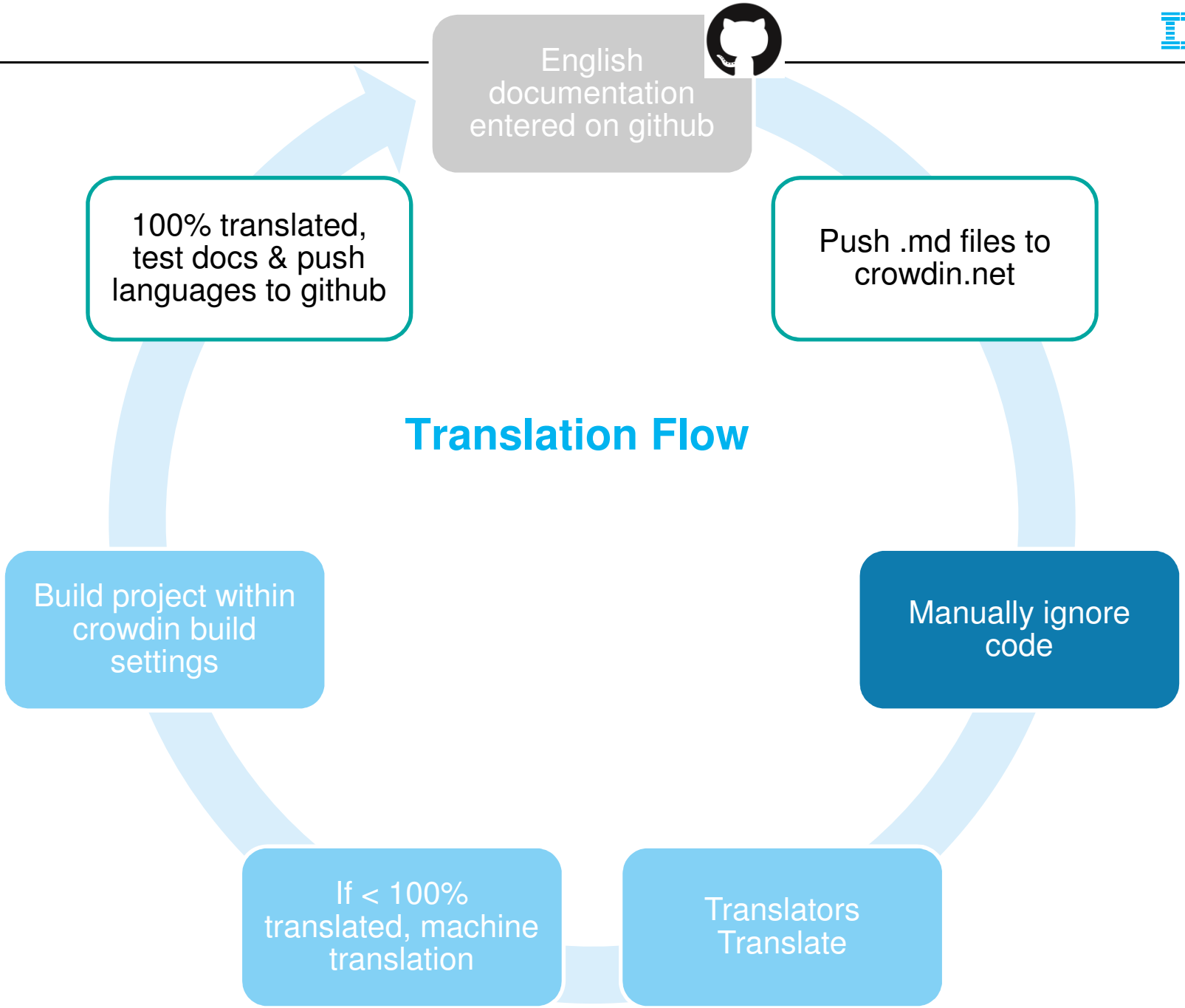
```
declare -a gitrepos=("cordova-docs" "cordova-plugin-battery-
status" "cordova-plugin-camera" "cordova-plugin-contacts"
"cordova-plugin-device" "cordova-plugin-device-motion" "cordova-
plugin-dialogs" "cordova-plugin-file" "cordova-plugin-file-
transfer" "cordova-plugin-geolocation" "cordova-plugin-
globalization" "cordova-plugin-inappbrowser" "cordova-plugin-
media" "cordova-plugin-media-capture" "cordova-plugin-network-
information" "cordova-plugin-splashscreen" "cordova-plugin-
vibration")
```

```
## now loop through the gitrepos array
for i in "${gitrepos[@]}"
do
   echo "***************** $i **************************"
   cd $BASE_GIT_REPO_PATH/$i

   # below pushes changes from local to fork
   git pull origin master
   git push origin
done
```

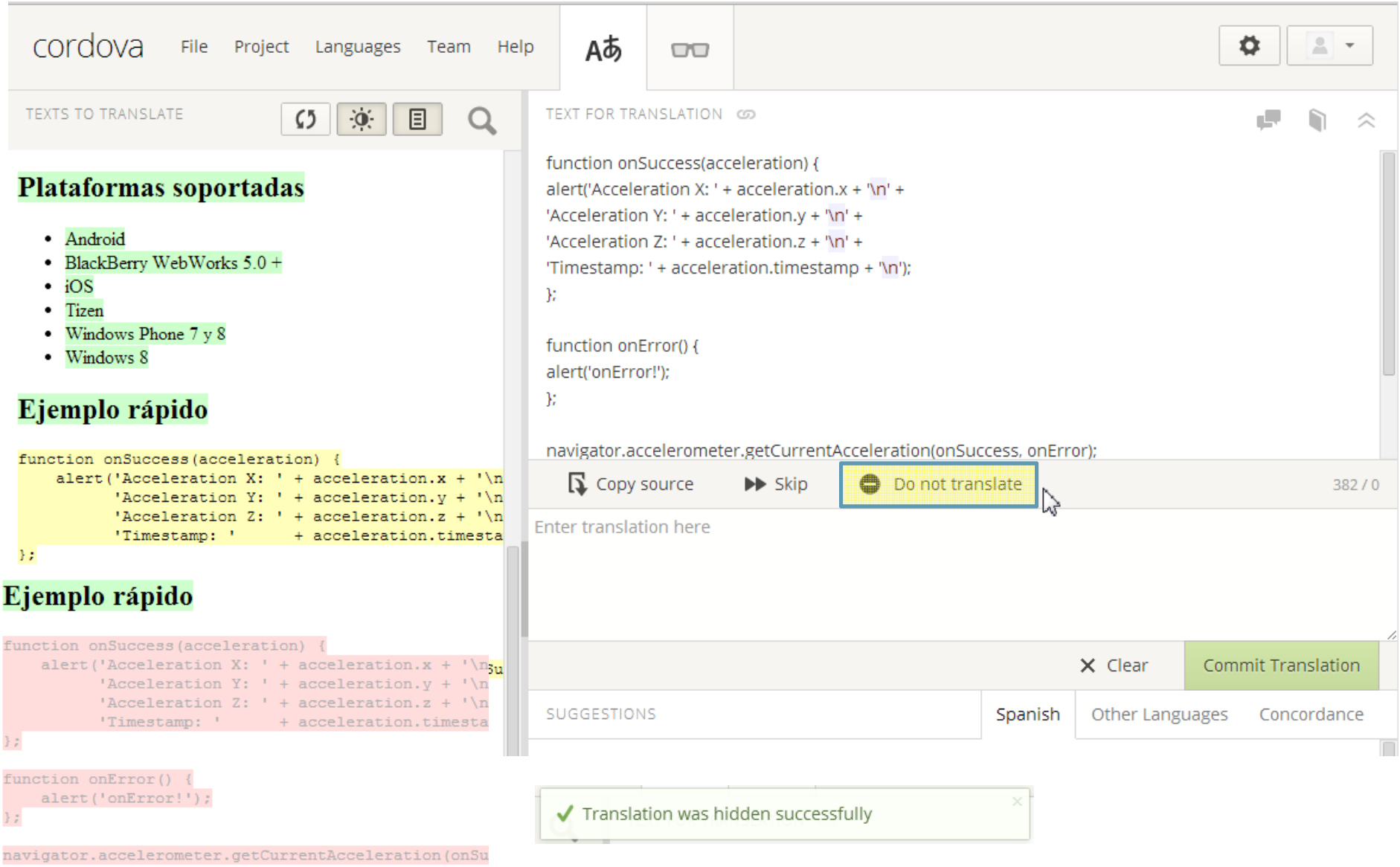
Demo

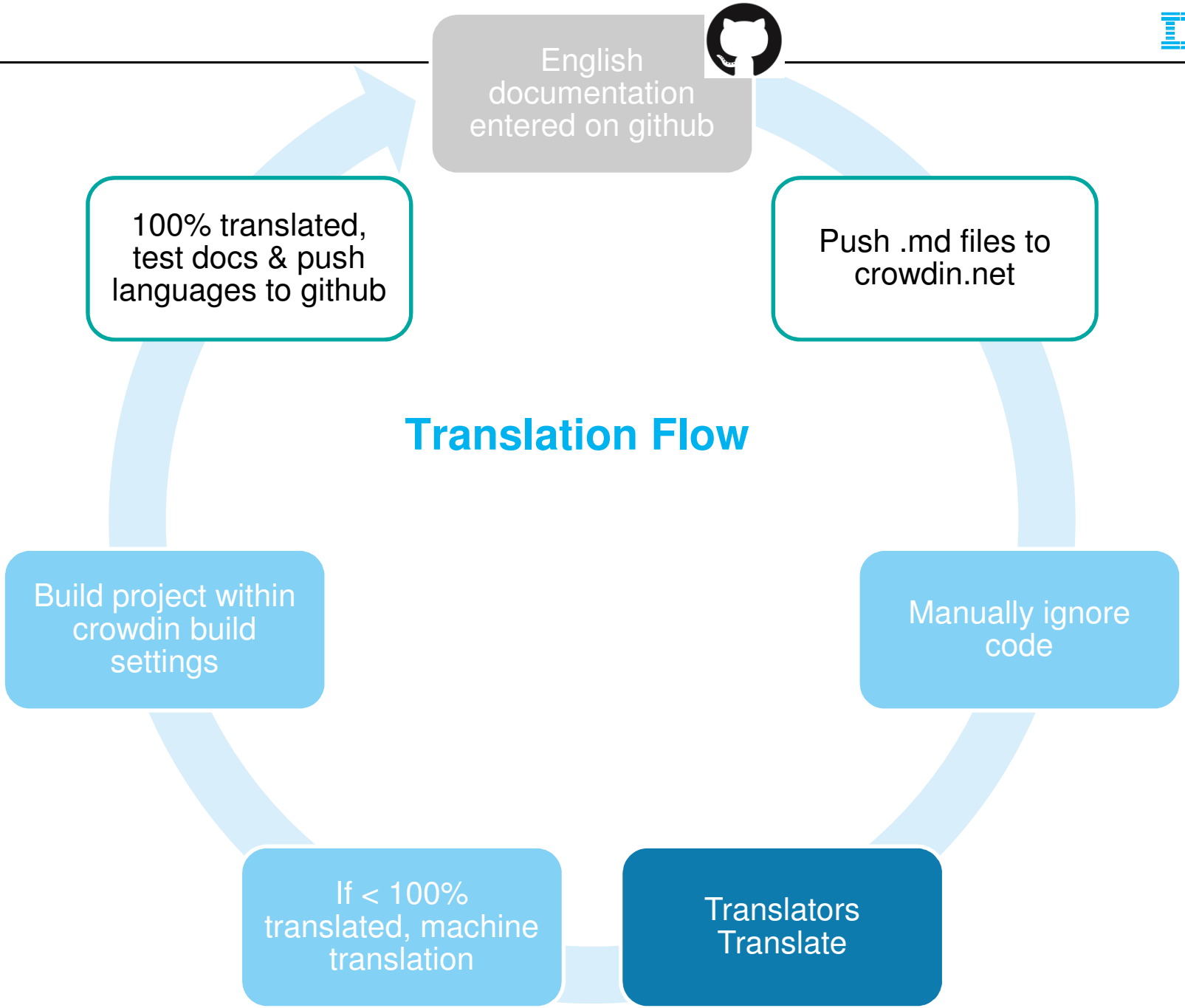
SEE THE CORDOVA CROWDIN SCRIPT IN ACTION

CROWDINCORDOVADEMO-LISA.SWF


English documentation entered on github
Push .md files to crowdin.net
Manually ignore code
Translators Translate
If < 100% translated, machine translation
Build project within crowdin build settings
100% translated, test docs & push languages to github
Translation Flow

Manual Step: Ignoring Code Snippets




English documentation entered on github
Push .md files to crowdin.net
Manually ignore code
Translators Translate
If < 100% translated, machine translation
Build project within crowdin build settings
100% translated, test docs & push languages to github
Translation Flow

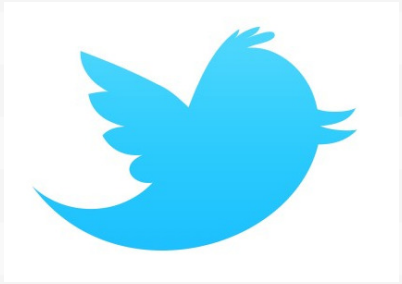
Translators

cordova settings

General | Translations | Screenshots | Glossary | TM & MT | File Manager | Translators | Strings | Reports | API | Branding | Extras

Invite | Professional Translations | Filter | Select all on this page | Contact all selected | Contact all

Name	Role	Status	Info
alexandresantos	translator	active	
badneoguy	translator	active	
bdylvr	translator	active	
bugmeniet	translator	active	
carloshbcabral	translator	active	
creakie92	translator	active	
csantana23	translator	active	
daniel009840	translator	active	
danigomez	translator	active	
DavidHormigo	translator	active	
delkant	translator	active	
dnm	translator	active	
eobriot	translator	active	
erikusaj	translator	active	
ewaldhorn	translator	active	

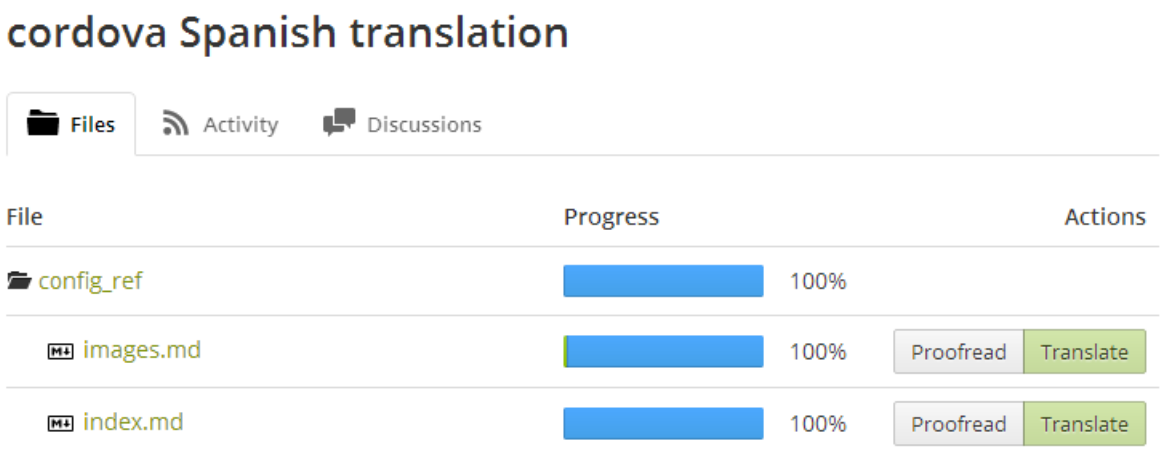


Page 1 of 5

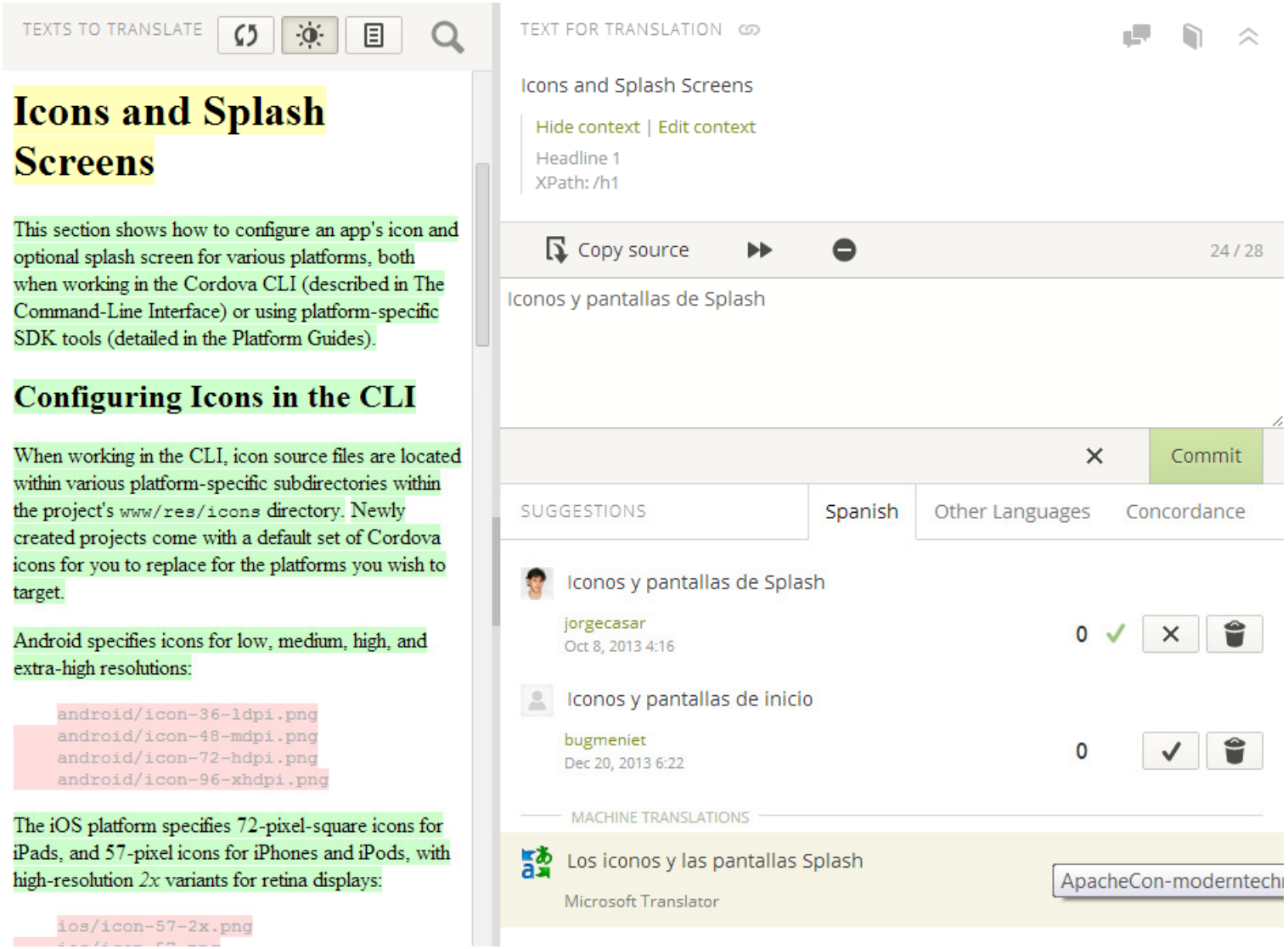
Displaying 1 to 15 of 73 items

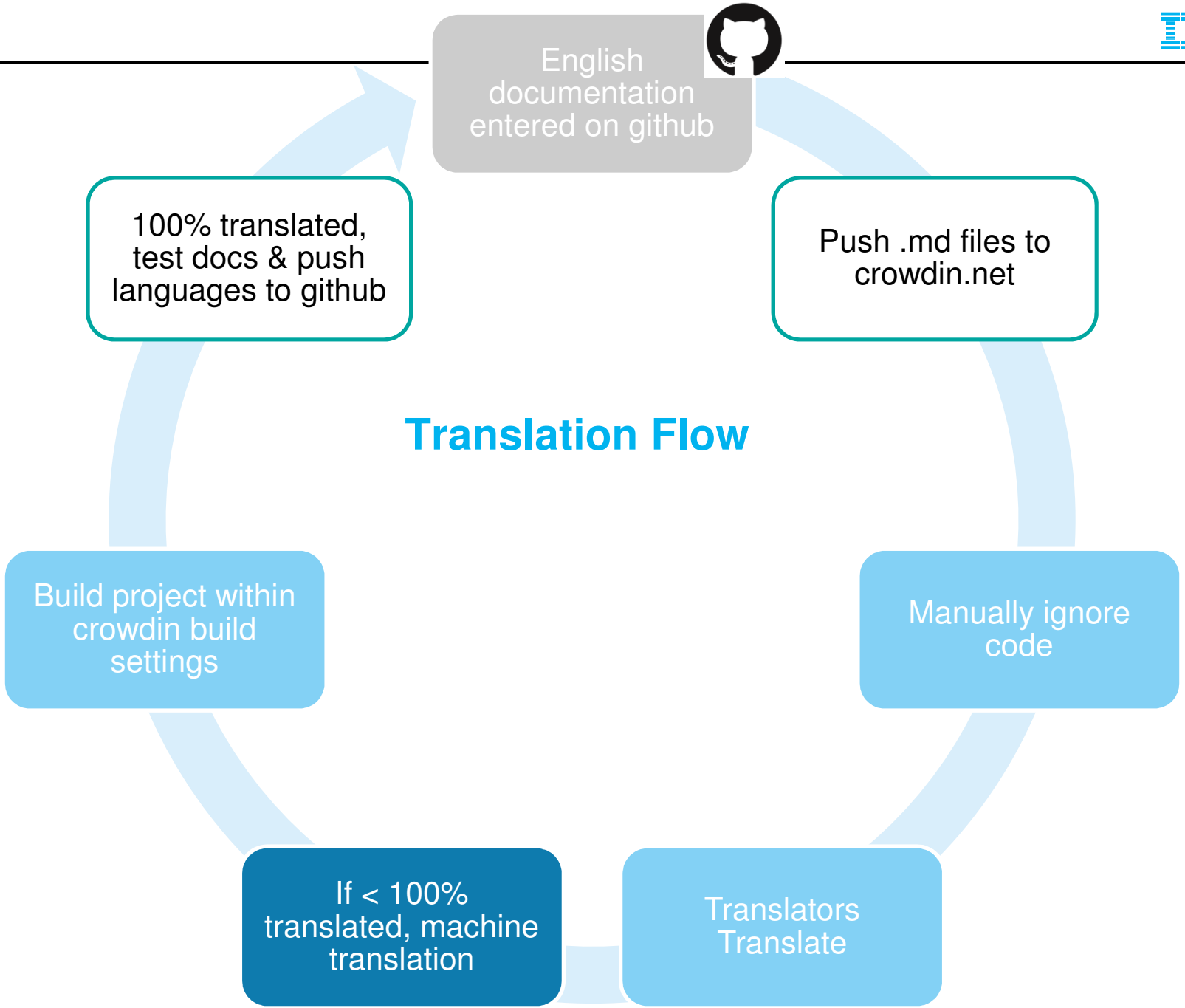
Process for Translators

1. Create a free account with Crowdin.net http://crowdin.net
2. Search for and find the Cordova project http://crowdin.net/project/cordova/
3. Scroll down to find the languages that are currently being translated.
4. Choose a language and click on the language to start contributing. e.g. Spanish http://crowdin.net/project/cordova/es-ES
5. Next to each of the markdown files you will see a button labeled "Translate". Clicking that button will open a translate page where the English version of the file is on the left.

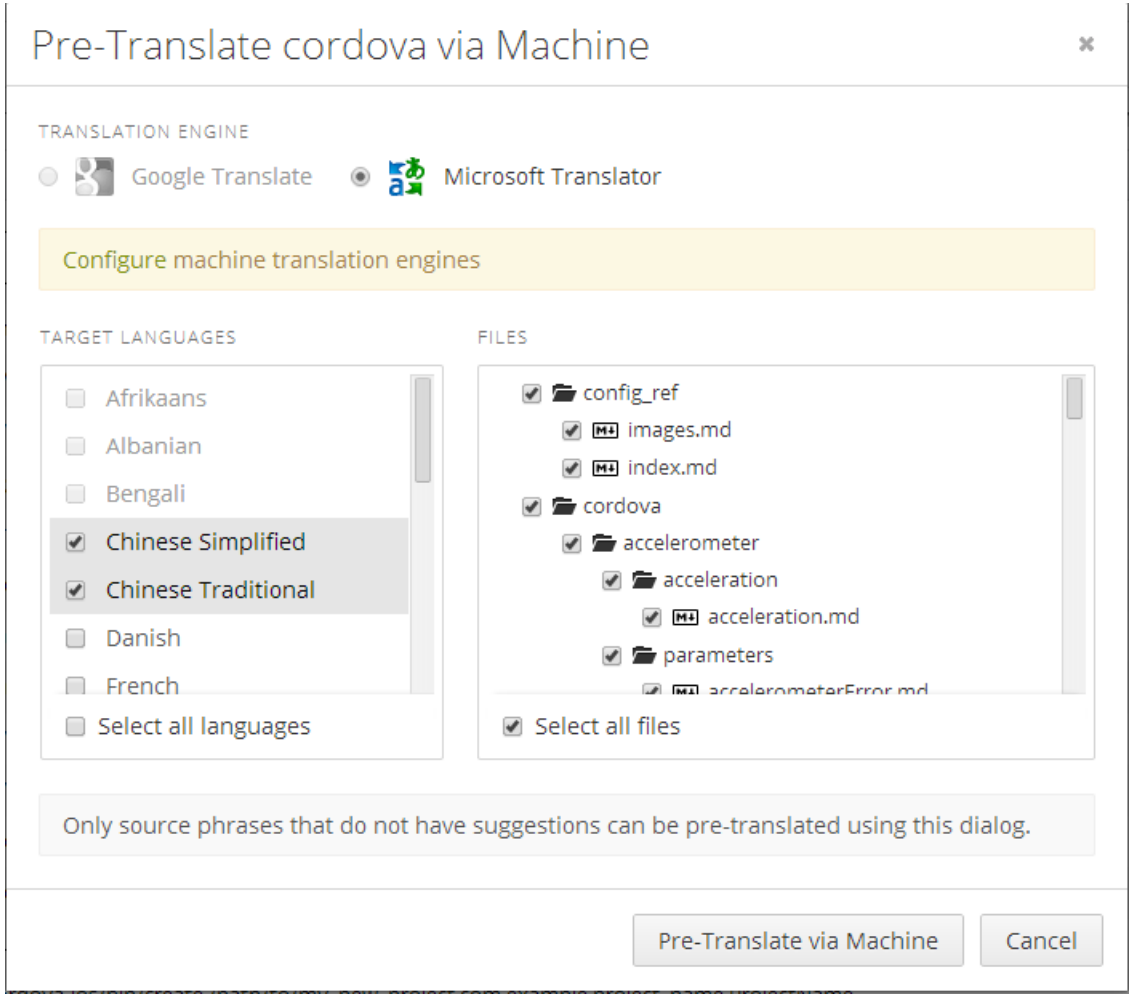


6. Click on a phrase in the left-hand panel. Then type a translation in the text area. Click the "Commit" button.

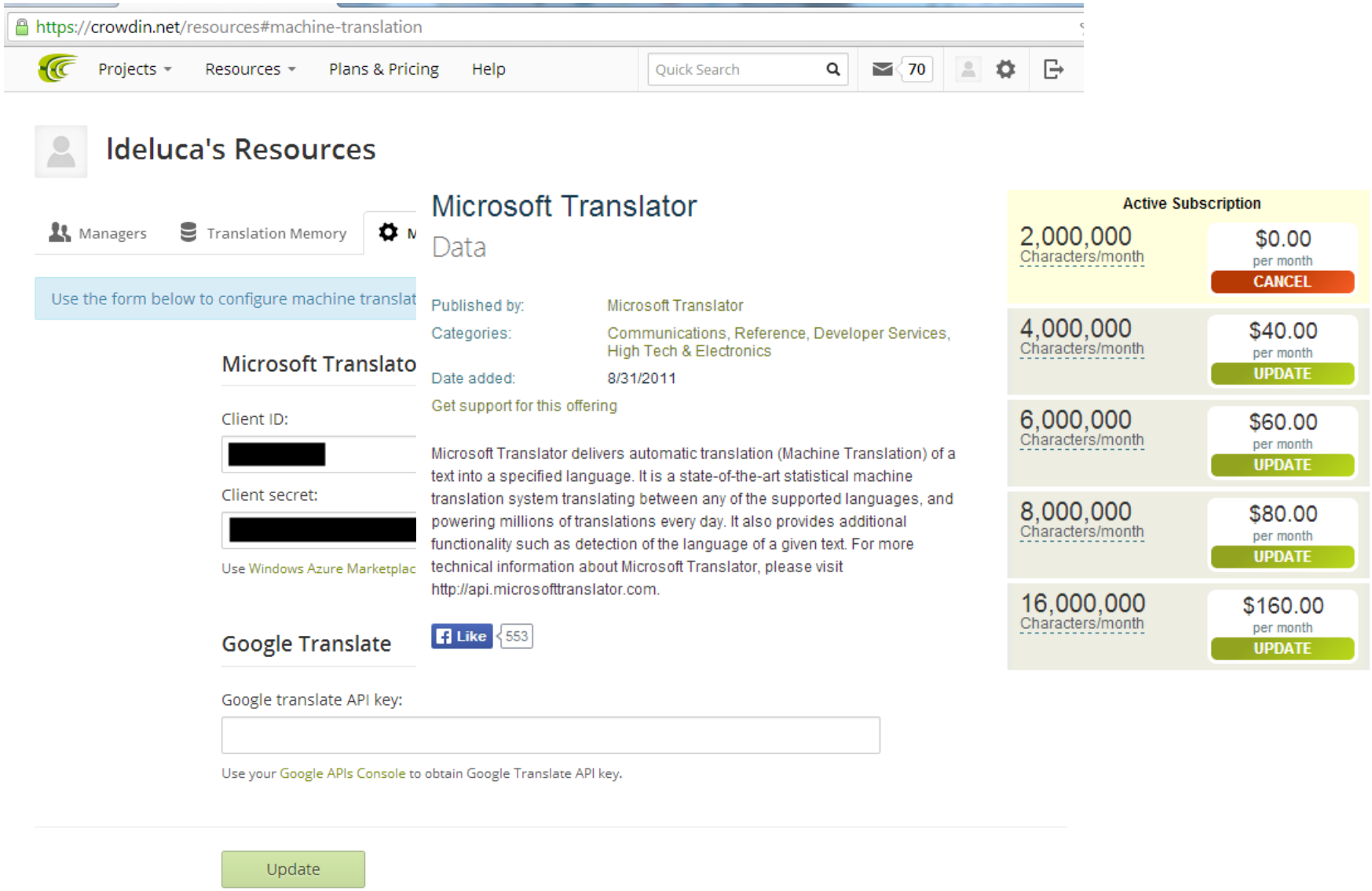


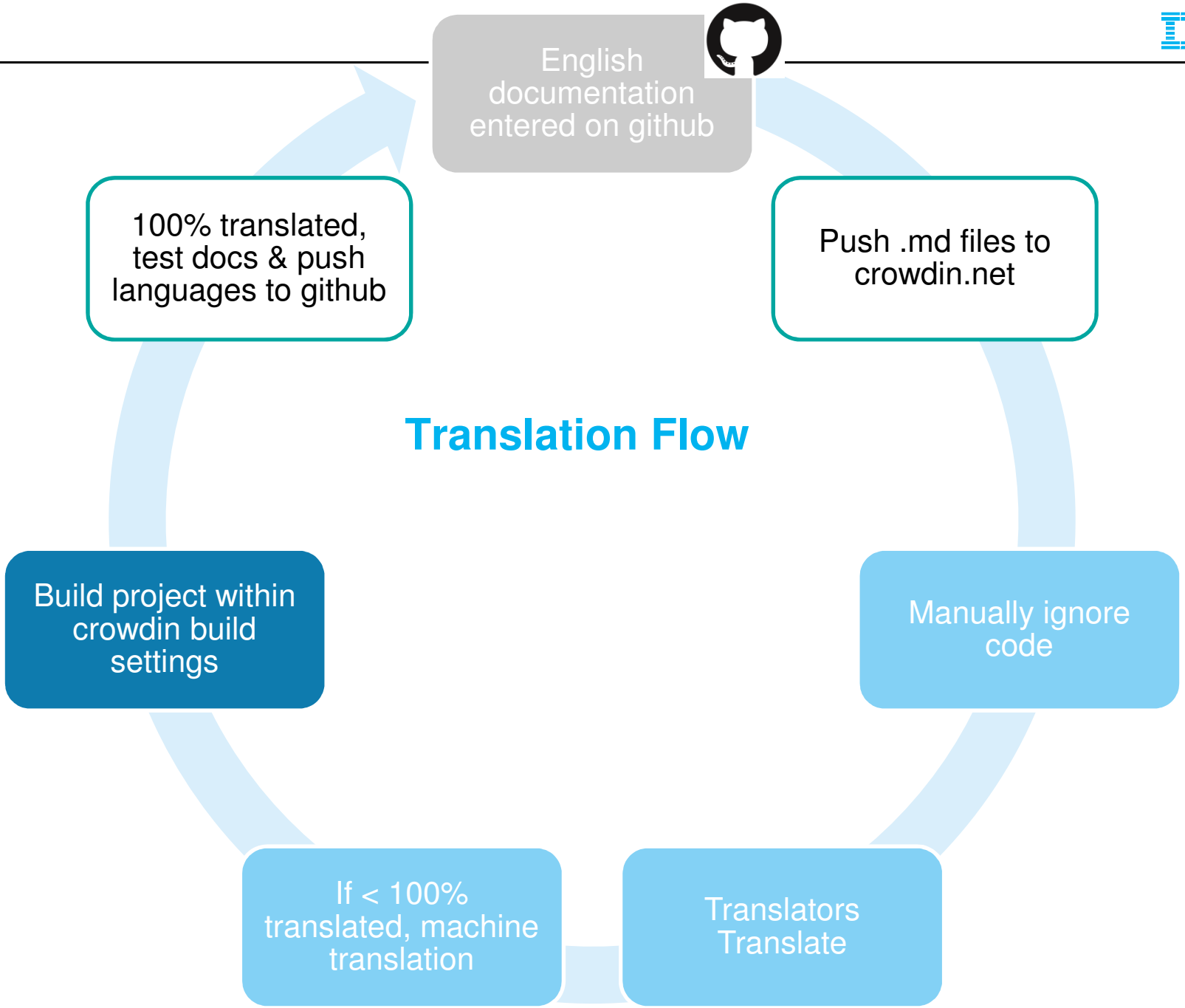

English documentation entered on github
Push .md files to crowdin.net
Manually ignore code
Translators Translate
If < 100% translated, machine translation
Build project within crowdin build settings
100% translated, test docs & push languages to github
Translation Flow

Manual Step: Machine Translation

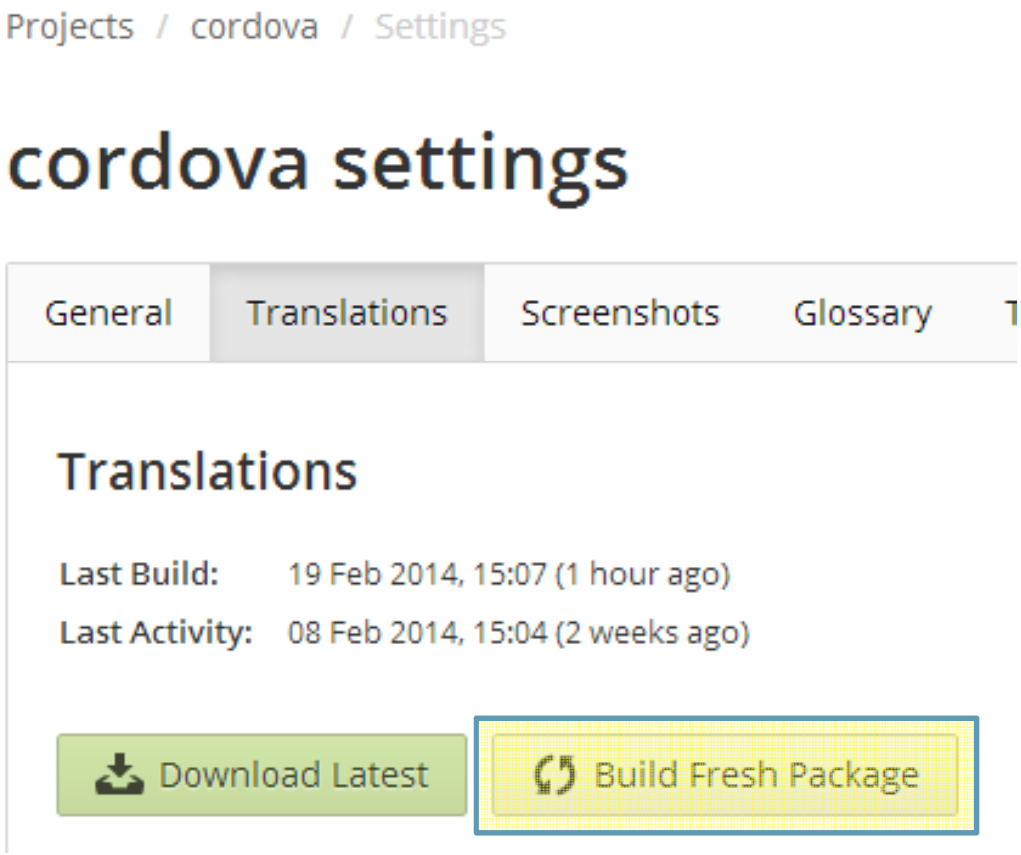


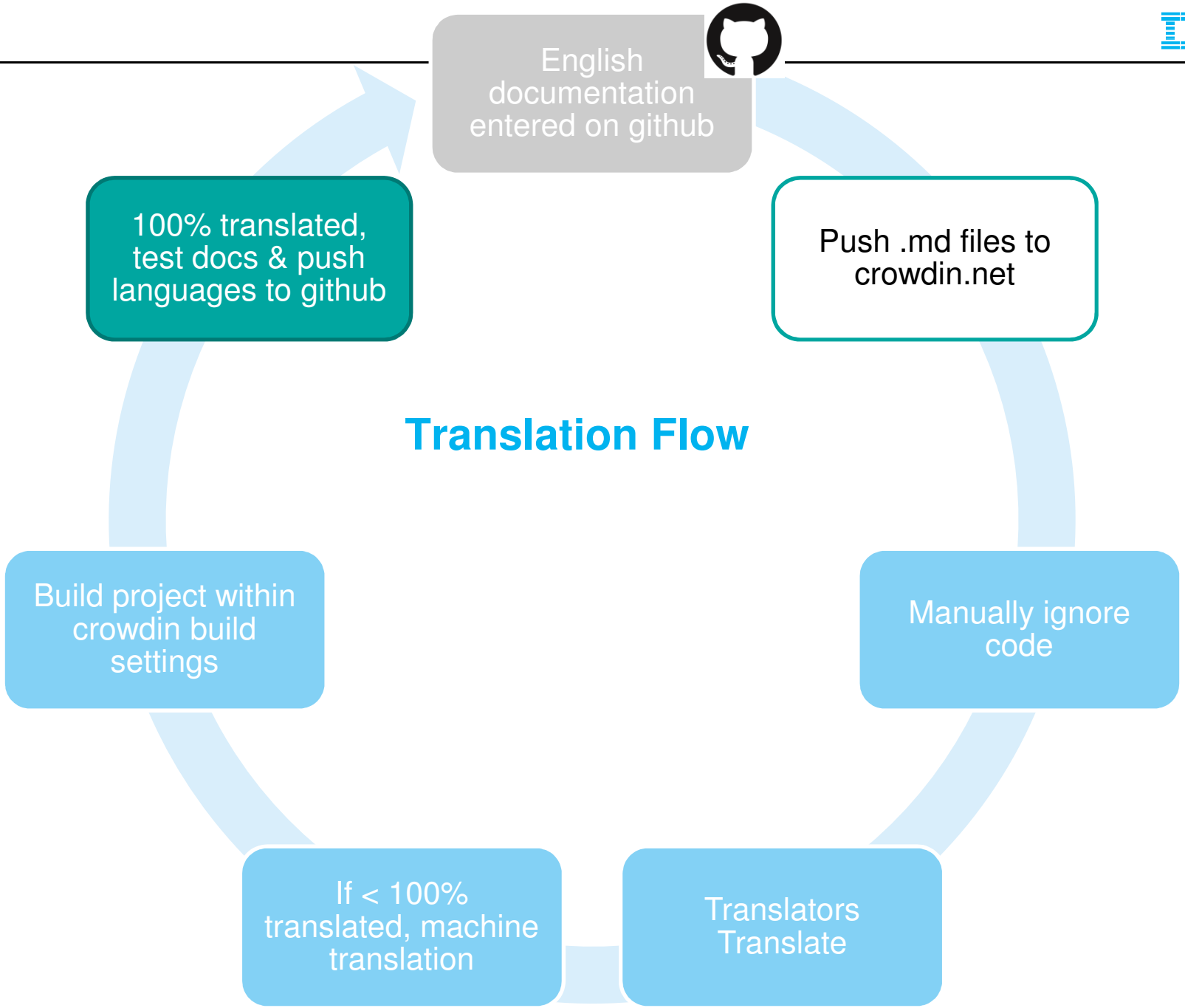
Machine Translation Engines for Crowdin




English documentation entered on github
Push .md files to crowdin.net
Manually ignore code
Translators Translate
If < 100% translated, machine translation
Build project within crowdin build settings
100% translated, test docs & push languages to github
Translation Flow

Manual Step: Build Crowdin Project




English documentation entered on github
Push .md files to crowdin.net
Manually ignore code
Translators Translate
If < 100% translated, machine translation
Build project within crowdin build settings
100% translated, test docs & push languages to github
Translation Flow

Manual: Test Building Docs

- Generate the documentation

```
# test the languages (preferrably before committing ☺)
~/git/cordova-docs$ bin/generate ru edge
```

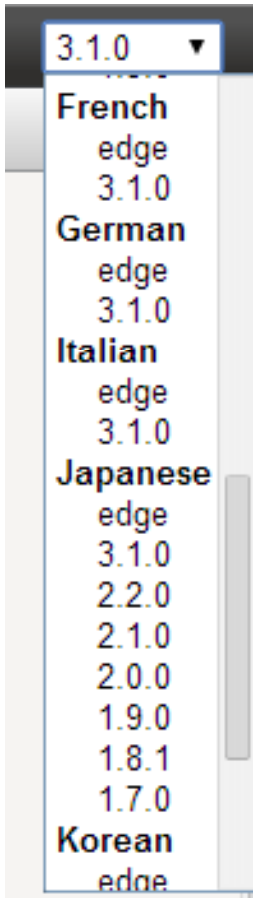
- Watch out for headings!!
  - If a word is translated in two or more ways the links will be broken

Script – Committing translations

- Edit commit message

```
git add .
git commit -m "Lisa @ApacheCon "
git push origin
```

- Re-run github-crowdin.sh
- Push from fork to Apache stream
- Build documentation (Run rake)
- Visit documentation site: http://cordova.apache.org/docs/en/edge/index.html



Other Gotchas

- Inline code is hard for translation
  - Can’t select a single word or part of a paragraph to ignore.
- Languages must be at 100%
  - Specific to our script
- Machine Translation Limitations
  - There is a limit with the free MT services
  - Advice: Ignore before using MT
- Checkout projects using ssh
  - Ex. git clone ssh://git@github.com/ldeluca/cordova-plugin-battery-status.git
  - Otherwise prompted for credentials for each repo

Call to Action

- We are still looking for translators! Join us!

http://crowdin.net/project/cordova

How to Add Crowdin To YOUR Open Source Project

- Visit our wiki to learn about what we're doing.
  http://wiki.apache.org/cordova/CordovaTranslations

- Scripts are freely available

- Email me: ldeluca@us.ibm.com

Cordova Hackathon

Sponsored by

IBM & RedHat

- **When:** Tuesday
- **Where:** McCourt room
- **What:** Race to Apache Cordova Contribution & Hybrid Mobile App Challenge
- **Who:** Everyone is invited. 8+ Apache Cordova Committers will be in the room answering questions. Come stump the experts!
- **Why:** See what it's like to participate in the Apache Cordova community.

ibm.com/mobilefirst