

Základní dokumentace CAT nejnovější



Obsah

- [Úvod](#)
- [Začínáme](#)
 - [Nový projekt](#)
 - [Překladová paměť](#)
 - [Správa terminologie](#)
 - [Manipulace segmentů](#)
 - [Statistika](#)
 - [Náhled](#)
 - [Generovat cílové soubory](#)
- [Pokročilé funkce](#)
 - [Strojový překlad](#)
 - [Pre-translát](#)
 - [Výběr slov pro získání významů](#)
 - [Online slovník](#)
 - [Kontrola jazyka](#)
 - [Autokompletní \(Interaktivní strojový překlad\)](#)
 - [Quickfill](#)
 - [AutoCorrect](#)
 - [Exportovat Slovo pro přezkum](#)
 - [Exportovat dvojjazyčné soubory odstavců](#)
 - [Export do markdownu s poznámkami](#)
 - [Vyhledávání a nahrazení](#)
- [Překlad všech druhů souborů](#)
 - [Soubory TXT](#)
 - [Soubory IDML](#)
 - [Soubory XLIFF](#)
 - [Soubory PO](#)
 - [Soubory PDF](#)
 - [Používání Okapi pro překládání souborů v jiných formátech](#)
 - [Překládat soubory Wordu pomocí Felixe](#)
- [Spolupráce](#)
 - [Sdílení překladové paměti a termínu](#)
 - [Používání systému Git ke spolupráci](#)
- [Ostatní](#)
 - [Nastavení vzhledu](#)
 - [Automatické zálohování](#)
 - [Ovládání verzí pomocí systému Git](#)
 - [Správa modulů](#)
- [Jak přispět](#)

Základní dokumentace CAT

- [Dokumenty »](#)
- [Dokumentace BasicCAT](#)
- [Editovat na GitHub](#)

Dokumentace BasicCATu ¶

Úvod ¶

BasicCAT je počítačový překladový nástroj s otevřeným zdrojovým kódem, jehož cílem je poskytnout jednoduchý a užitečný nástroj pro překladatele. Název je BasicCAT, kvůli jeho jednoduchosti a programovacímu jazyku - Basic. Basic se snadno naučí a každý si může na základě zdrojového kódu BasicCAT vytvořit vhodný nástroj CAT.

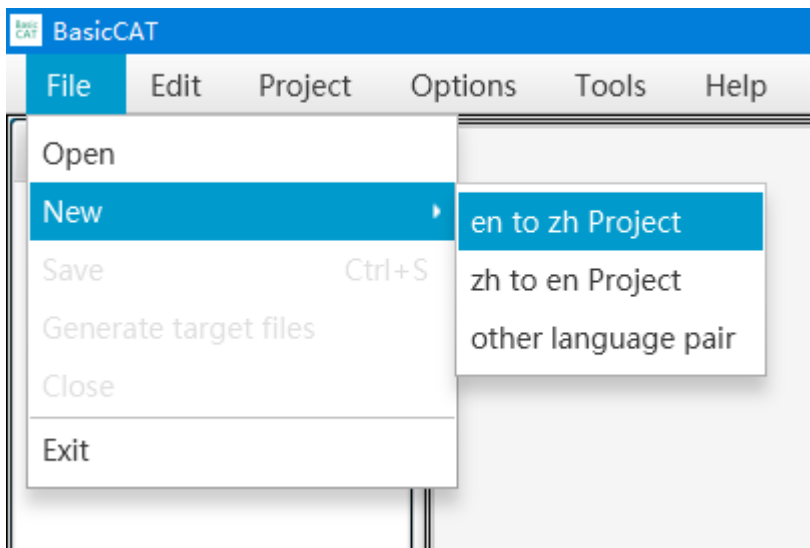
Má následující funkce:

- Překladová paměť
- Správa terminologie
- Kontrola jazyka
- Výběr slov pro získání významů
- Quickfill
- AutoCorrect
- Interaktivní překlad stroje
- Exportovat Word pro externí recenzi
- Exportovat dvojjazyčné soubory odstavců
- Export do markdownu s poznámkami
- Volně sloučit a rozdělit segmenty
- Integrace online slovníků
- Podpora Strojové překladové služby Poskytované společností Google, Microsoft, atd.
- Pretranslát založený na překladové paměti a strojovém překladu
- Podpora běžných formátů souborů: TXT, idml, xliiff, gettext Po
- Podporujte tři překladové standardy: TMX, TBX a SRX
- Sdílení překladové paměti a terminologie on-line
- Ovládání verzí a spolupráce pomocí systému Git

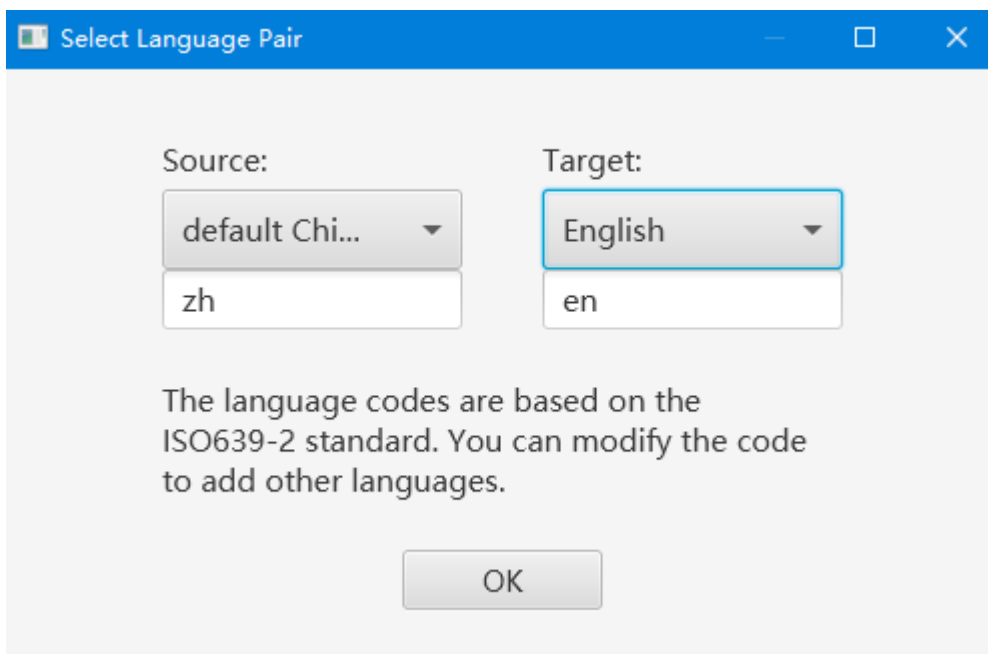
Začínáme

Nový projekt

Klikněte na „Soubor“ -> „Nový“, vyberte „en to zh“ pro vytvoření projektů v angličtině a čínštině a „zh to en“ pro vytvoření projektů v čínštině a angličtině. Vyberte jiný jazykový pár a určete zdrojový jazyk a cílový jazyk projektu. Překladová paměť a Term Base budou vytvořeny současně.



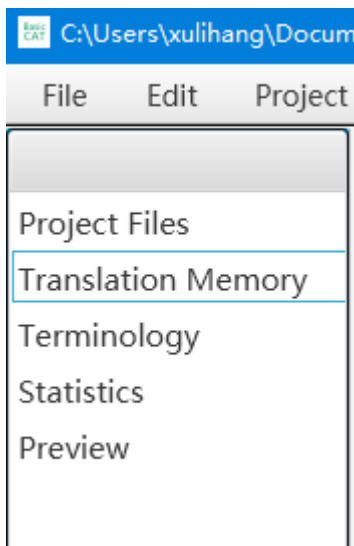
Můžete také zadat požadovaný kód jazyka, který odpovídá normě ISO 639, sami. Podrobné informace naleznete [zde](#).



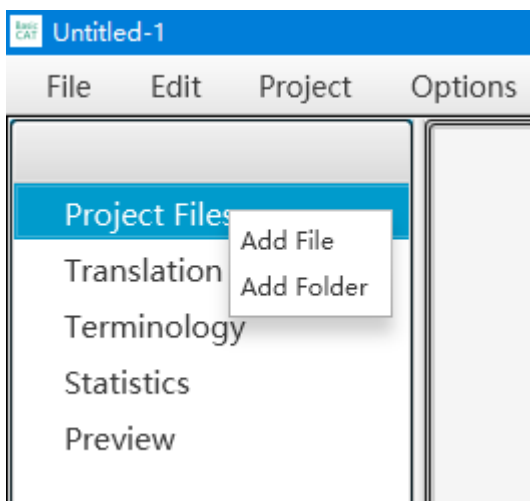
Uložte projekt před dalšími operacemi.

Přidat soubor ¶

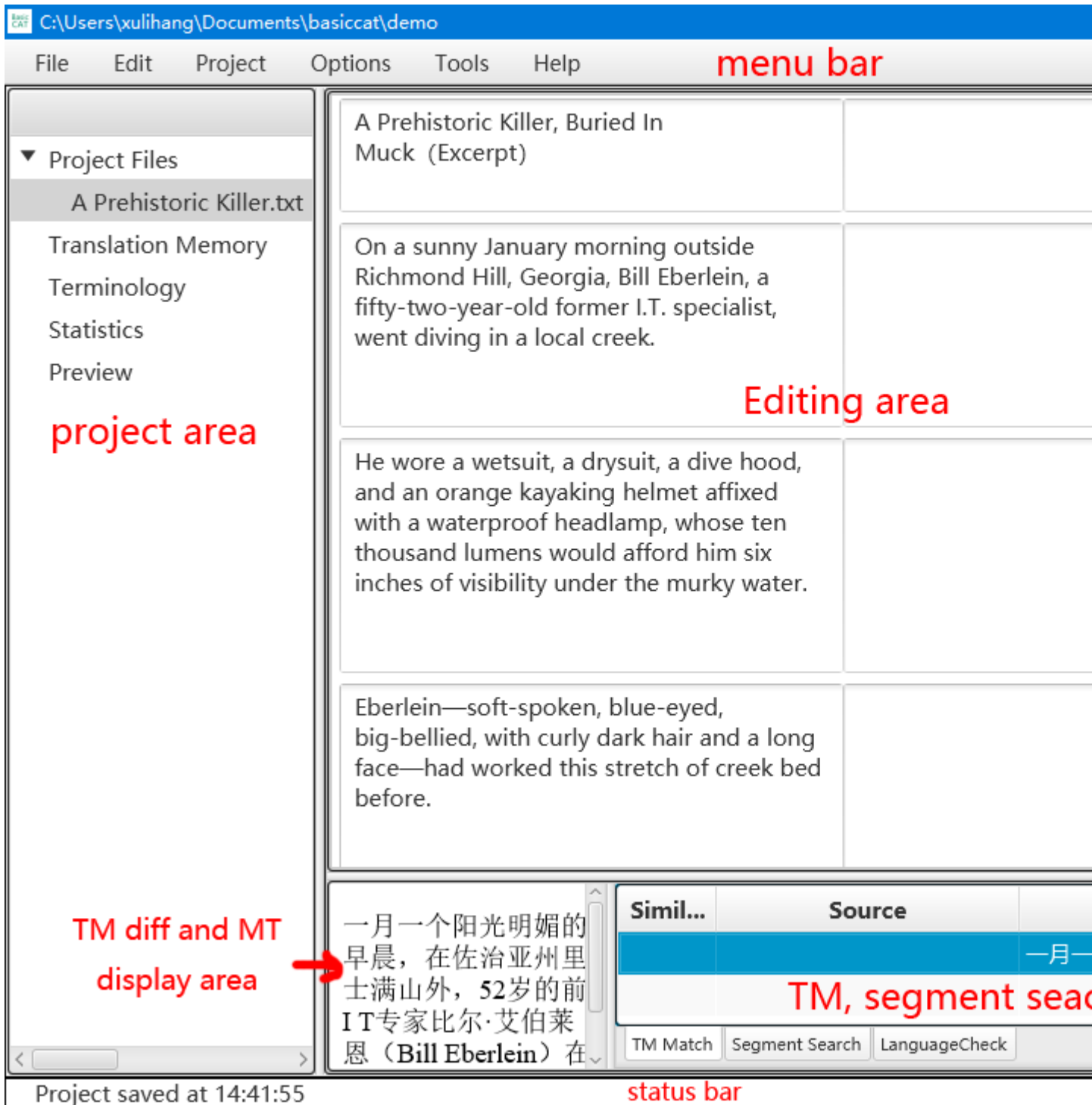
Po otevření projektu se v levé oblasti zobrazí seznam položek. Můžete spravovat soubory projektu, překladovou paměť, terminologickou základnu, zobrazit statistiky a zobrazit náhled.



Pravým tlačítkem myši klikněte na „Projektové soubory“ pro přidání souborů nebo přidání složek.

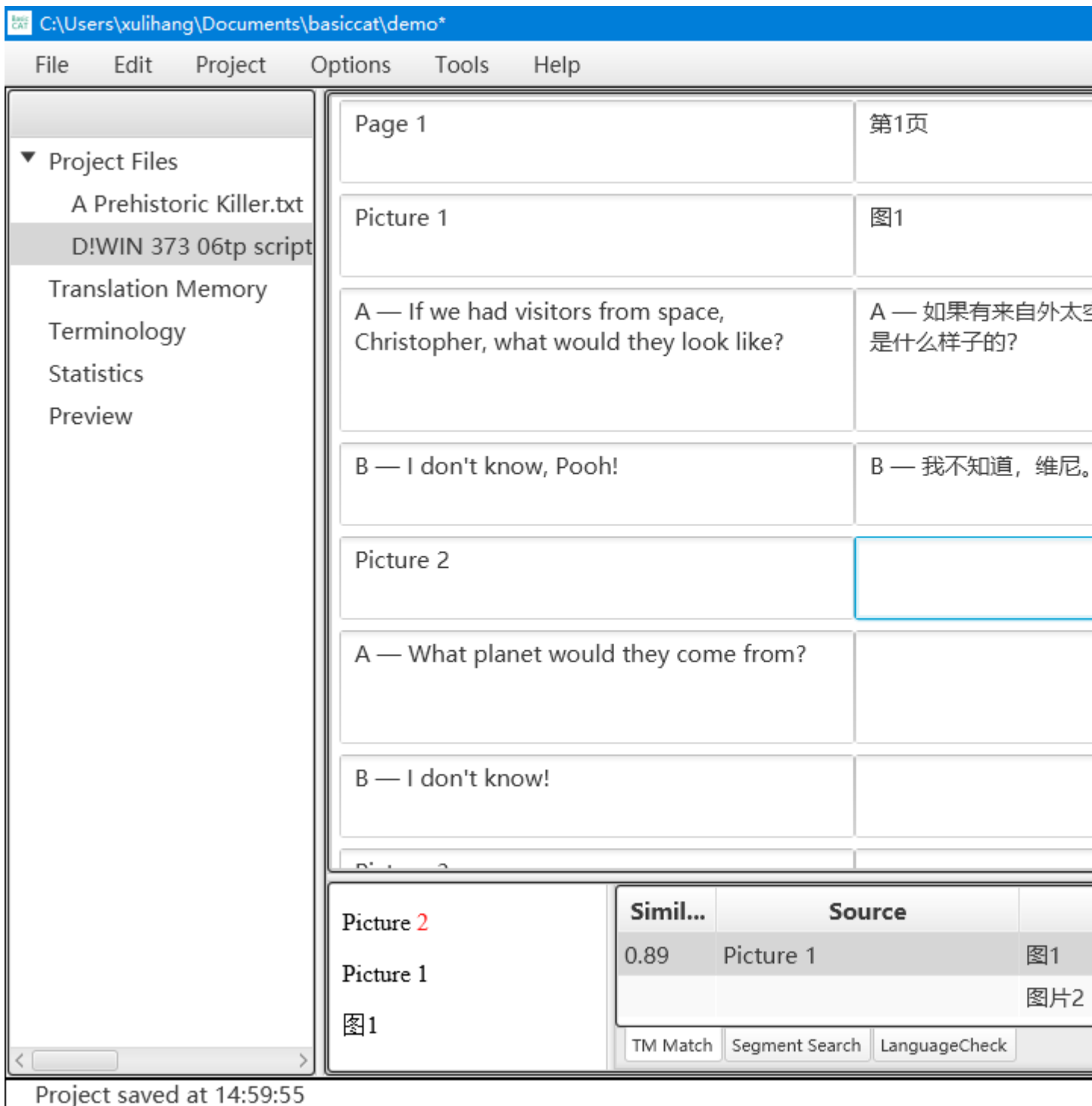


Kliknutím na názvy souborů otevřete soubory. Po otevření souboru bude rozhraní vypadat takto. Každá funkční oblast je na obrázku vyznačena. Vložte překlad do správné textové oblasti. Po dokončení překladu jednoho segmentu přejděte stisknutím klávesy „Enter“ na další.



Překlad paměti ¶

Po dokončení jednoho segmentu stisknutím klávesy „Enter“ bude překlad přidán do překladové paměti. Při překladu podobného segmentu se zobrazí v dolní oblasti. Kliknutím na tuto shodu vyplníte překlad do textové oblasti.



Hodnotu shody překladové paměti lze nastavit v Nastavení projektu. Hpdnota by měla být mezi 0,5 až 1,0.



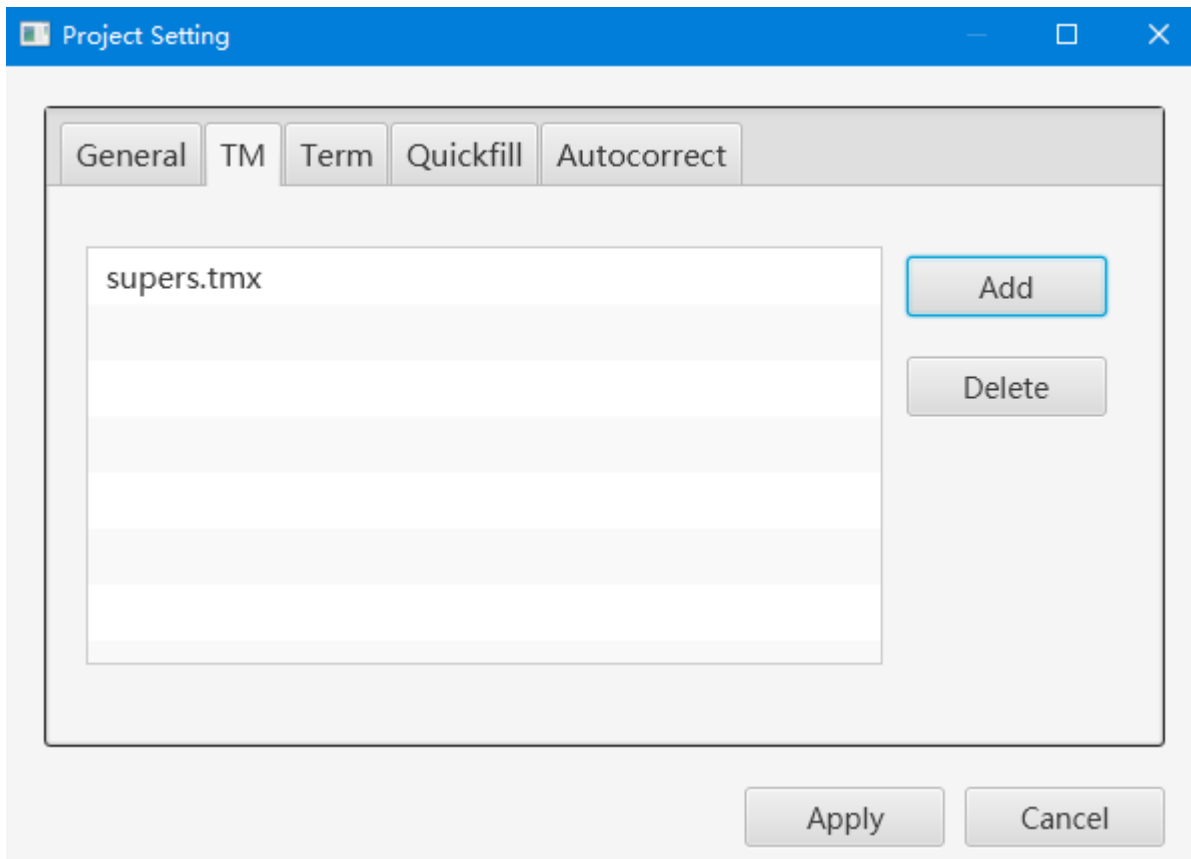
Přidat externí překlad paměti ¶

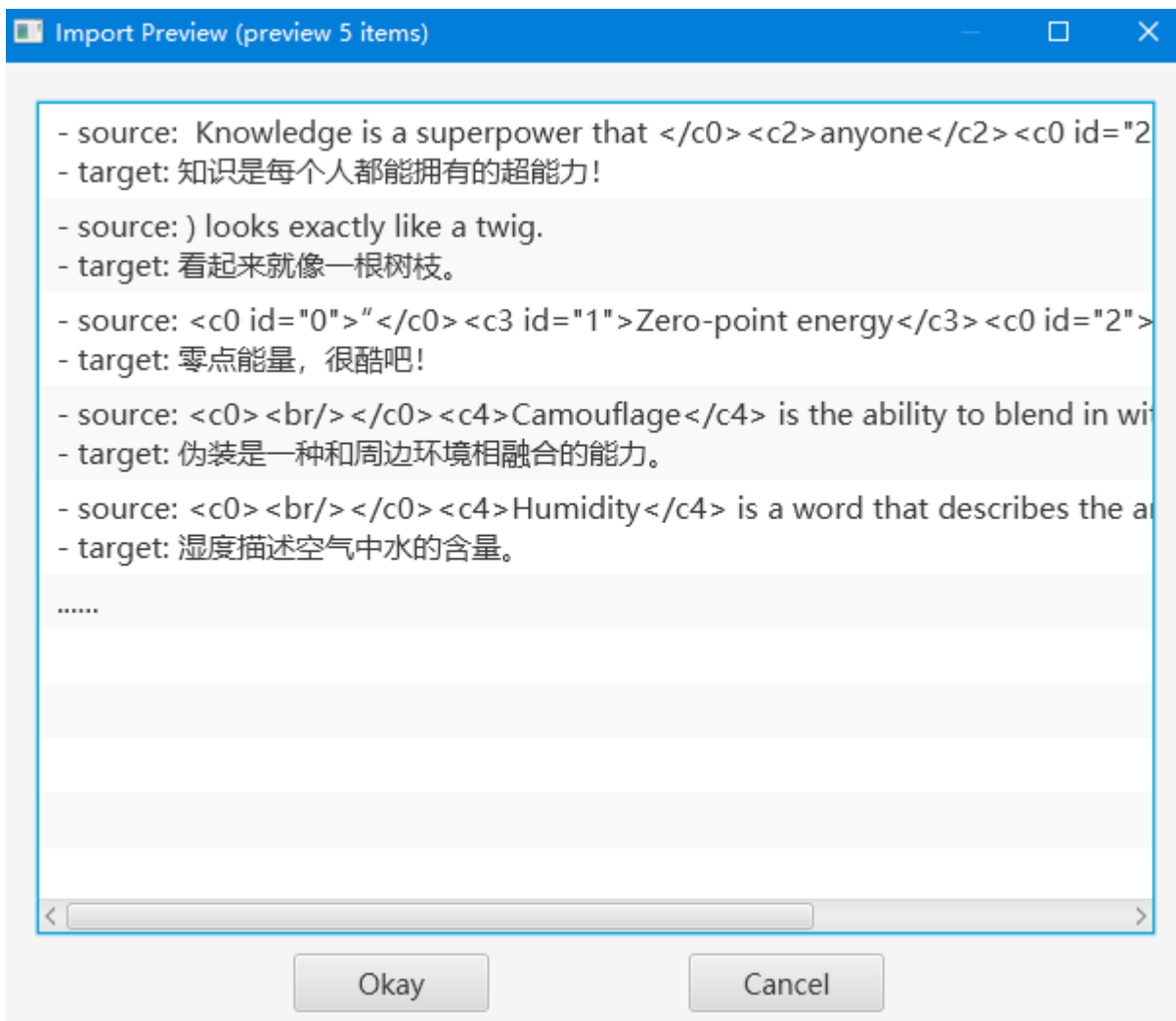
V BasicCAT existují dva typy překladové paměti. Jedním z nich je projektová paměť a druhým externí paměť. Paměť projektu ukládá paměti vytvořené při překladu souborů projektu, zatímco externí paměť zobrazuje importovanou překladovou paměť.

Klikněte na „Project->Nastavení projektu“ a v záložce „TM“ můžete spravovat

externí překladovou paměť.

Můžete importovat soubory TMX nebo soubory txt oddělené tabulátory. U souborů txt by měl být zdroj v prvním řádku a cíl v druhém. Po přidání nového souboru se zobrazí okno náhledu.





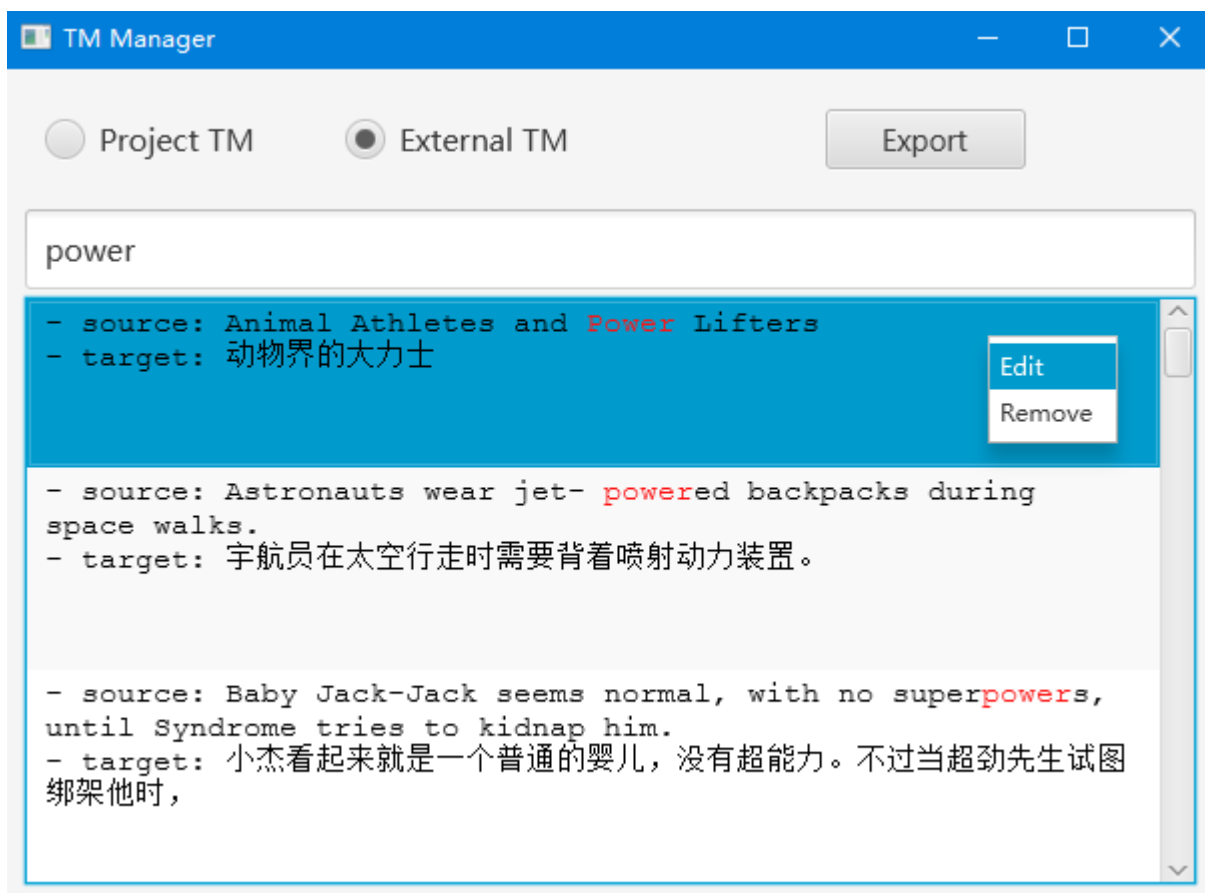
Když se shoduje segment z externí překladové paměti, zobrazí se název souboru, který označuje, odkud tato shoda pochází. Zdrojový text v překladové paměti, zdrojový text aktuálního segmentu a cílový text v překladové paměti se zobrazí v oblasti zobrazení rozdílů.

Simil...	Source	Target
0.6	Read on to discover the ...	发现超人背后的科学原理吧!
		超自然的真正科学
0.55	That's bad news for busy...	这对繁忙的超人们来说可并不

TM Match Segment Search LanguageCheck

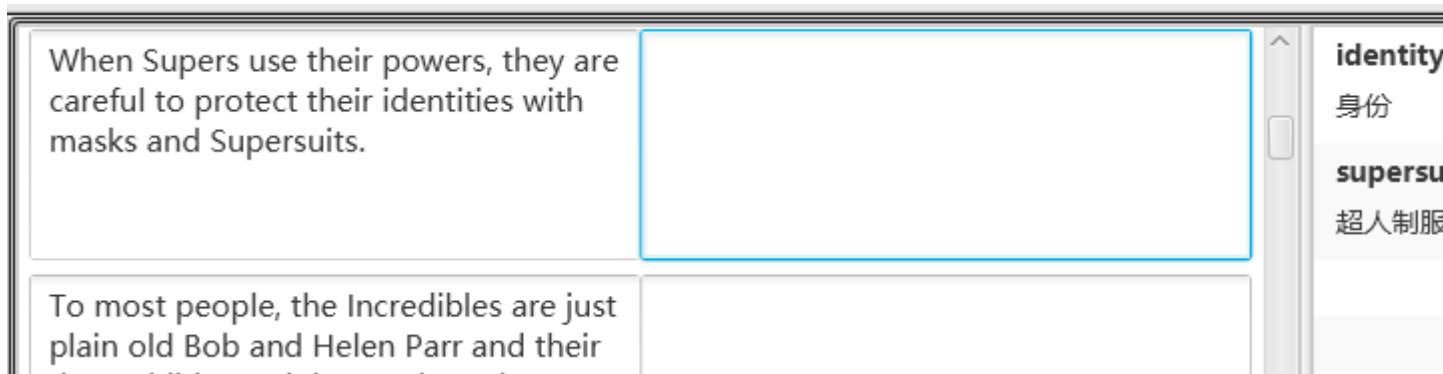
Správa paměti překladu ¶

Kliknutím na „Překladová paměť“ v oblasti Projekt otevřete Správce TM. Můžete vyhledávat, exportovat, upravovat nebo odstraňovat překladové paměti.

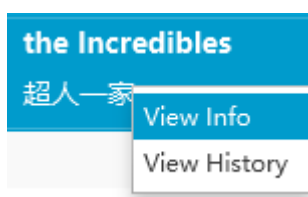


Řízení terminologie

Pokud věta obsahuje výrazy, můžete vybrat odpovídající texty ve zdroji a překlad a přidat výrazy. BasicCAT používá k lemmatizaci slov `opennlp`. Pokud tedy přidáte termín v jeho množném čísle, může BasicCAT detekovat jeho singulární formu v jiném segmentu.



Kliknutím pravým tlačítkem myši na položku výraz zobrazíte více informací a jeho historii.

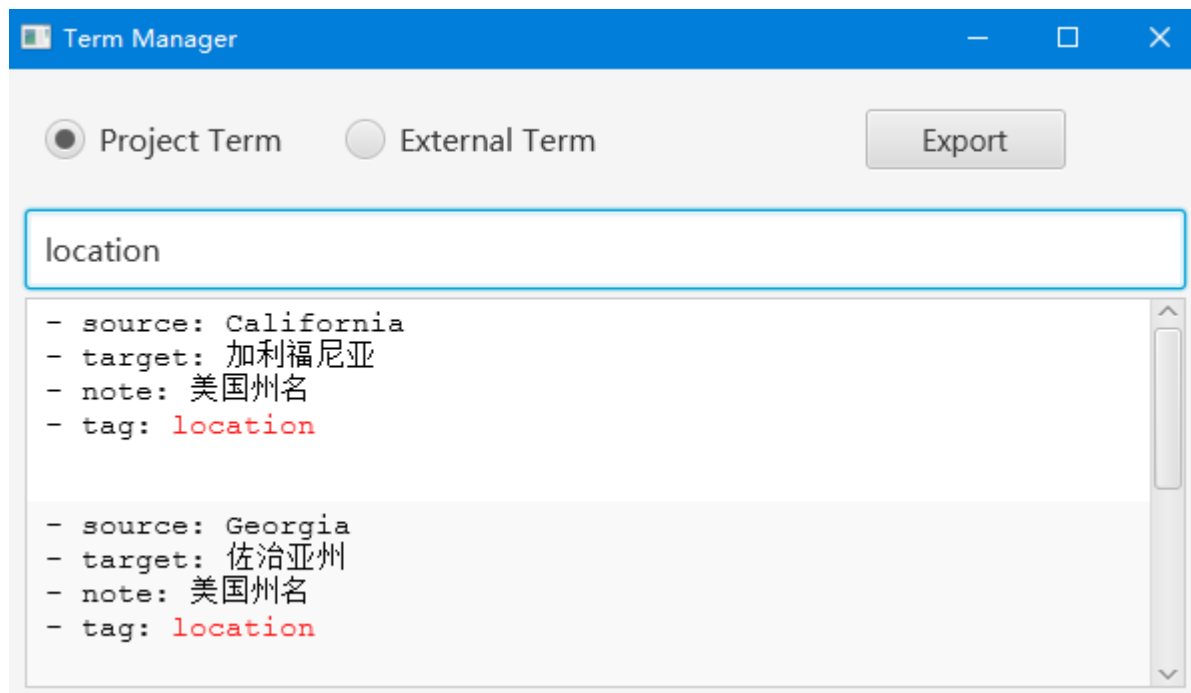


Pozor!

Jelikož externí databáze termínů může obsahovat tisíce záznamů, používá BasicCAT algoritmus HashMap k párování termínů. Lemmatizován bude pouze text ve zdroji. Termíny v externích termabázích nebudou lemmatizovány. Když tedy přidáváte termín, je lepší ho přidat v jeho původní podobě.

Import termínů je téměř stejný jako import překladové paměti. Podporovány jsou soubory TBX a txt oddělené tabulátory.

Správce terminologie je také podobný správci TM. Rozdíl spočívá v tom, že k termínům lze přidávat značky a poznámky.

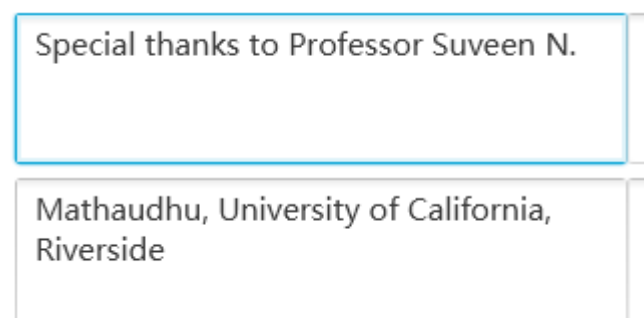


Manipulace segmentů ¶

BasicCAT uses the SRX segmentation standard to segment the text. A segment can be a sentence or a phrase.

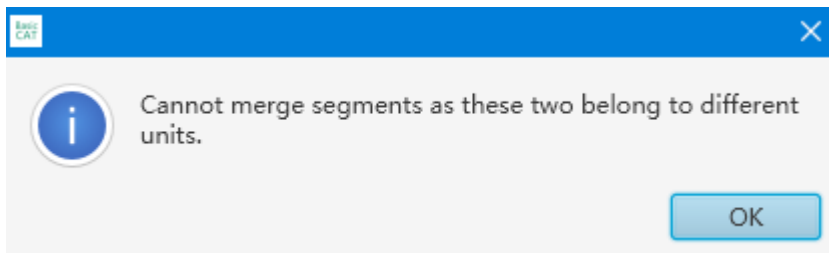
Sloučit a rozdělit segmenty ¶

Pokud narazíte na chybně segmentované jméno, jako je níže, můžete přesunout kurzor na konec a stisknete „Smazat“ pro sloučení dvou segmentů.

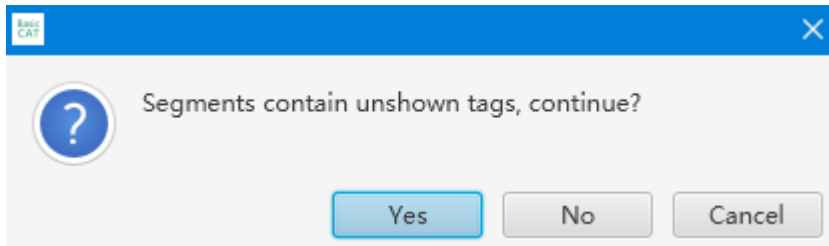


Pokud dva segmenty patří do různých souborů nebo překladových jednotek, nelze je sloučit. K takovému případu patří různé odstavce v Wordu a různé příběhy v

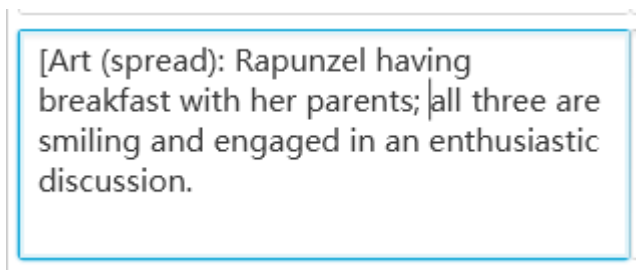
InDesignu.



BasicCAT skryje formátovací tagy, pokud je to možné. Takže pokud segmenty obsahují skryté značky, bude zde okno se zprávou, jak je uvedeno níže. Můžete se rozhodnout pokračovat a sloučený zdrojový text může obsahovat složité značky.



Když potřebujete rozdělit, jako na středníku níže, přesuňte kurzor na středník a stiskněte tlačítko „Enter“.



Zanedbávání segmentu

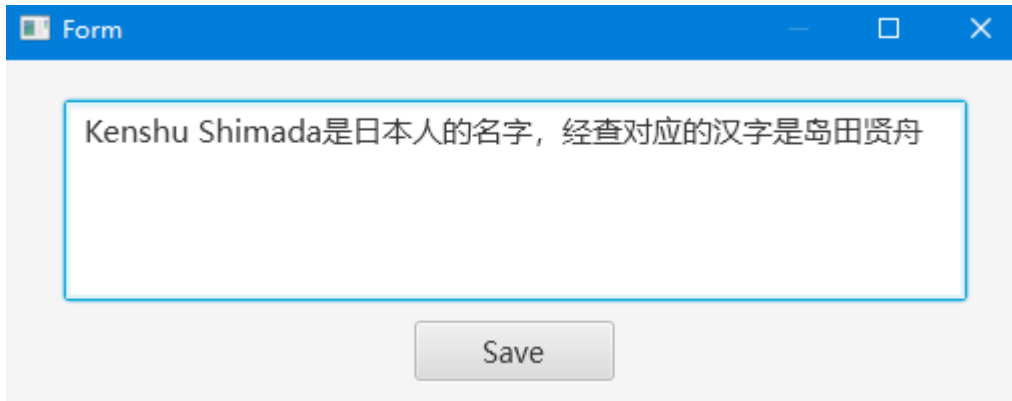
Při překladu z angličtiny do čínštiny je běžné, že první a druhý segment mají podobné významy. První můžete označit jako zanedbané a přeložit pouze druhé. Při generování cílových souborů budou tyto segmenty vynechány. K tomu použijte Nabídka Úpravy-> Označit aktuální segment jako zanedbaný.

Textová oblast zanedbaných segmentů bude šedá a neupravitelná.

The Incredibles are a family of "Supers."	
This means they are people with superhuman powers.	超人一家每个人都有超能力,

Přidat poznámky ¶

Pokud narazíte na obtížné věty, můžete si udělat poznámky, jak překlad dokončit. Chcete-li zobrazit nebo upravit poznámky, použijte Nabídka Upravy-> Zobrazit / Upravit poznámky aktuálního segmentu.



The screenshot shows a window titled "Form" with a blue header bar. Inside the window, there is a text area containing the text: "Kenshu Shimada是日本人的名字，经查对应的汉字是岛田贤舟". Below the text area is a "Save" button.

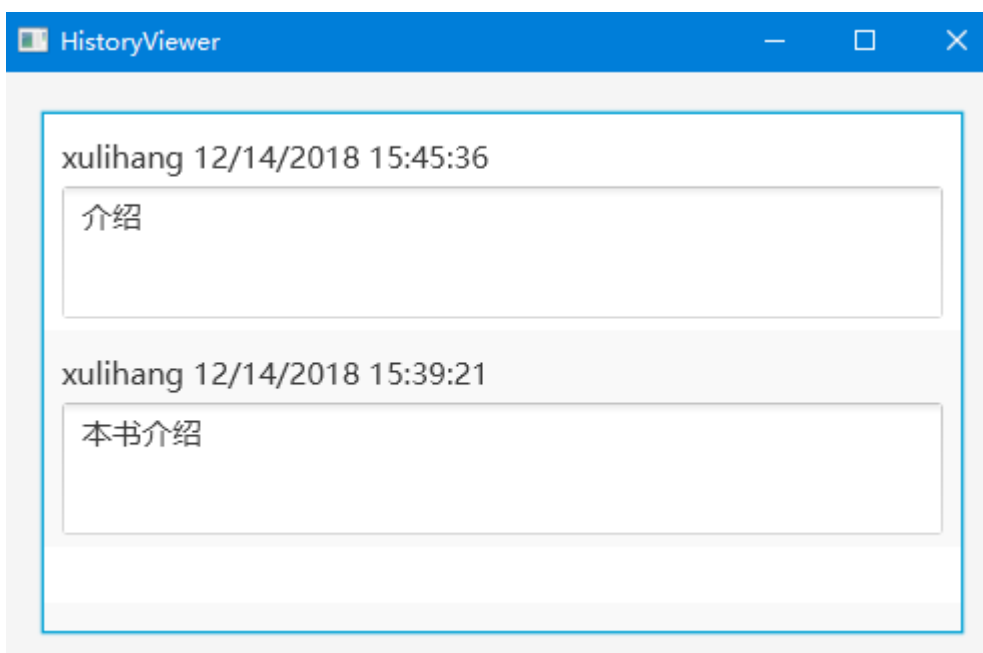
Segmenty obsahující poznámky budou mít textarea se šedým okrajem.



The screenshot shows a text editor with two segments. The left segment contains the text: "But, beyond that, according to Kenshu Shimada, a professor of paleobiology at DePaul University, 'There are a lot of things that we still don't know.'" The right segment contains the text: "但德保罗大学古生物学教授岛田贤舟（罗马文Kenshu Shimada）称目前还有很多我们不知道的。". The right segment's text area has a green border, while the left segment's text area has a grey border.

Zobrazit historii segmentů ¶

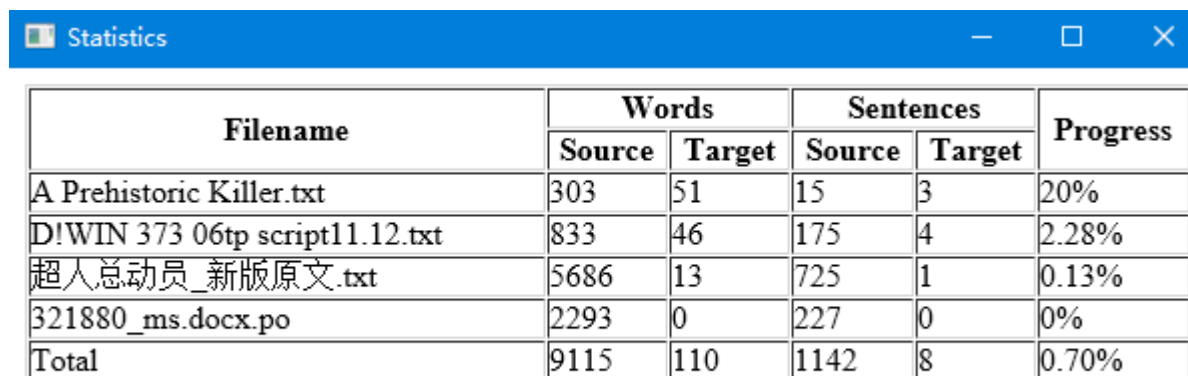
BasicCAT bude zaznamenávat historii segmentů. Kliknutím na nabídku Upravit-> Zobrazit historii segmentů zobrazíte historii, kde uživatelské jméno je uživatelské jméno přidané v nastavení správy verzí.



The screenshot shows a window titled "HistoryViewer" with a blue header bar. Inside the window, there is a list of segment history entries. Each entry consists of a timestamp and a text area. The first entry is: "xulihang 12/14/2018 15:45:36" followed by a text area containing "介绍". The second entry is: "xulihang 12/14/2018 15:39:21" followed by a text area containing "本书介绍".

Statistika ¶

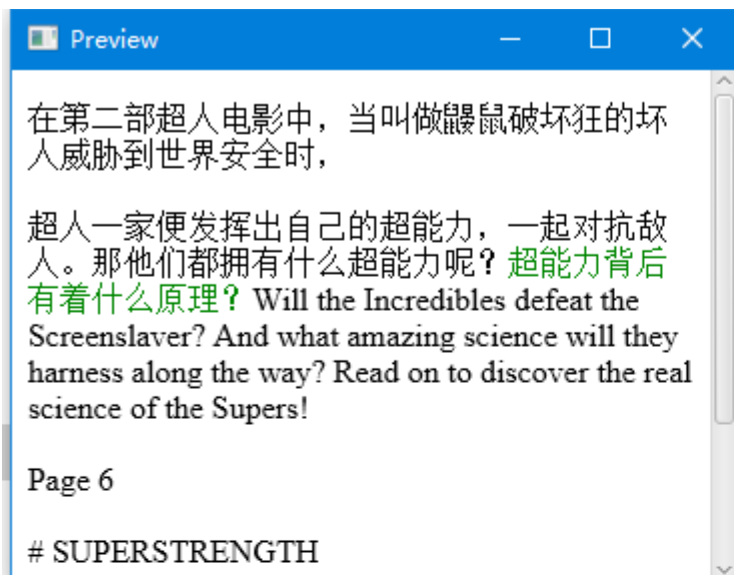
Klikněte na „Statistika“ v oblasti projektu, můžete vidět statistiky jako počet slov a procento dokončeno.



Filename	Words		Sentences		Progress
	Source	Target	Source	Target	
A Prehistoric Killer.txt	303	51	15	3	20%
D!WIN 373 06tp script11.12.txt	833	46	175	4	2.28%
超人总动员_新版原文.txt	5686	13	725	1	0.13%
321880_ms.docx.po	2293	0	227	0	0%
Total	9115	110	1142	8	0.70%

Náhled

Kliknutím na „Náhled“ v oblasti projektu zobrazíte náhled textu. Přeložený zdrojový text bude nahrazen překladem.



Page 5	第5页
In the second film, a villain named the Screenslaver threatens all Supers.	在第二狂的场
The family has to summon all their powers and work together.	超人一对抗敌
What exactly are their superpowers?	那他们
How do they work?	超能力
Will the Incredibles defeat the Screenslaver?	

Generovat cílové soubory ¶

Při překladu použijte menu „File->Generovat cílové soubory“ vytvořit přeložené soubory v cílové složce.



Pokročilé funkce

Strojový překlad ¶

BasicCAT má vestavěnou podporu pro 12 strojových překladatelských služeb:

- [Baidu](#)
- [Google](#)
- [Microsoft Bing](#)
- [Niutrans](#)

- [Youdao](#)
- [Yandex](#)
- [MyMemory](#)
- [Sogou](#)
- [Sogou DeepI](#)
- [Tencent Cloud](#)
- [IBM Watson](#)
- [Amazon](#)

Výsledek se ukáže ve spodní části.

A Prehistoric Killer, Buried In Muck (Excerpt)	淤泥下的史前杀手 (节选)
On a sunny January morning outside Richmond Hill, Georgia, Bill Eberlein, a fifty-two-year-old former I.T. specialist, went diving in a local creek.	
He wore a wetsuit, a drysuit, a dive hood, and an orange kayaking helmet affixed with a waterproof headlamp, whose ten thousand lumens would afford him six inches of visibility under the murky water.	
Eberlein—soft-spoken, blue-eyed, big-bellied, with curly dark hair and a long face, had worked this stretch of	

一月一个阳光明媚的早晨，在佐治亚州里士满山外，52岁的前IT专家比尔·艾伯莱恩（Bill Eberlein）在当地的小溪里潜水。	<table border="1"> <thead> <tr> <th>Simil...</th> <th>Source</th> <th>Target</th> <th>M</th> </tr> </thead> <tbody> <tr> <td></td> <td></td> <td>一月一个阳光明媚的早晨，在...</td> <td>baidu</td> </tr> <tr> <td></td> <td></td> <td>一月的一个阳光明媚的早晨，乔</td> <td>niutran</td> </tr> <tr> <td></td> <td></td> <td>在乔治亚州里士满山外面一个...</td> <td>mymen</td> </tr> </tbody> </table>	Simil...	Source	Target	M			一月一个阳光明媚的早晨，在...	baidu			一月的一个阳光明媚的早晨，乔	niutran			在乔治亚州里士满山外面一个...	mymen
Simil...	Source	Target	M														
		一月一个阳光明媚的早晨，在...	baidu														
		一月的一个阳光明媚的早晨，乔	niutran														
		在乔治亚州里士满山外面一个...	mymen														

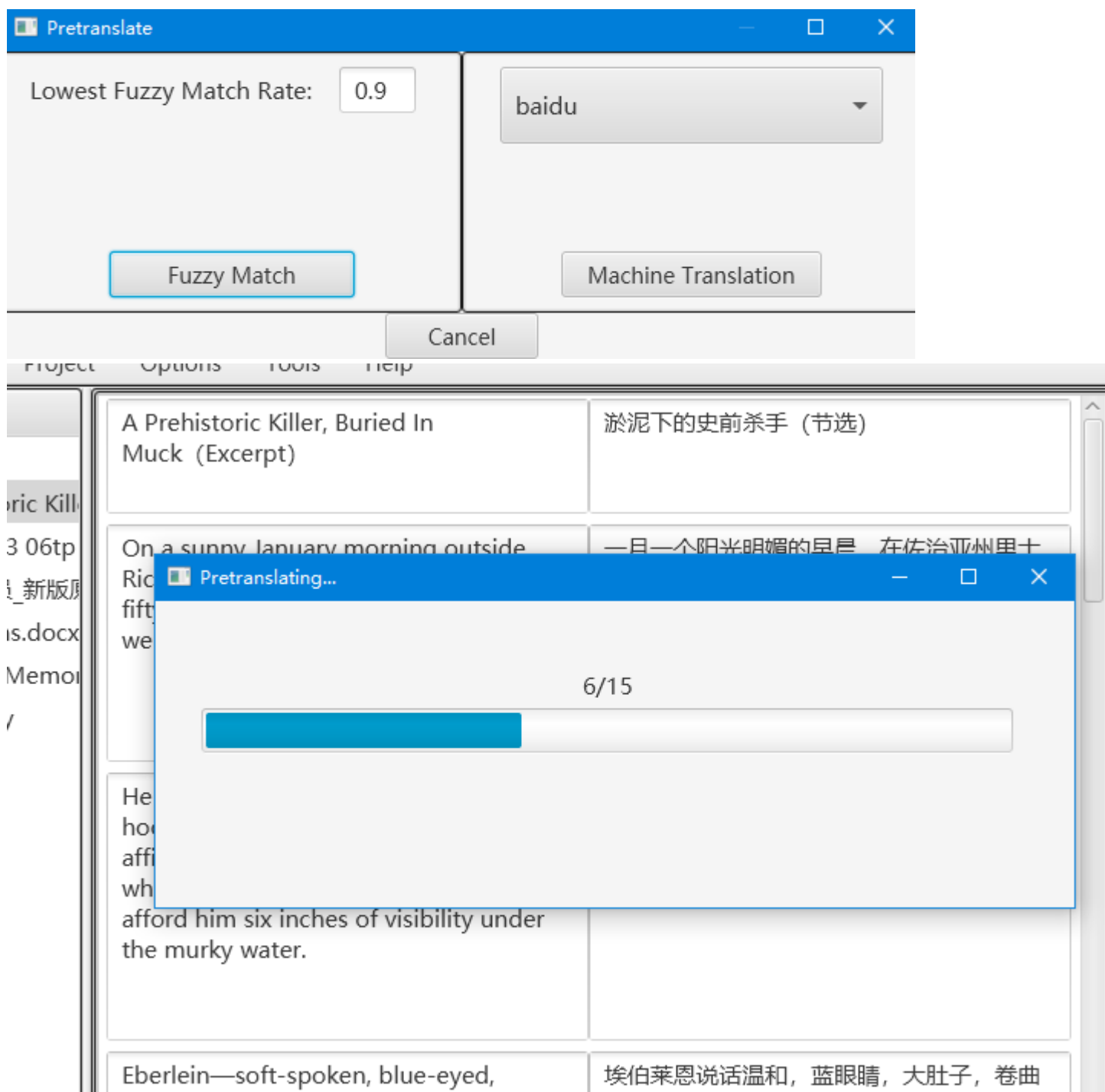
TM Match Segment Search LanguageCheck

:26

Abyste mohli tyto překladatelské služby používat, musíte požádat o klíče API. Odkazy k podání žádosti byly zahrnuty do výše uvedeného seznamu MT. MyMemory nevyžaduje klíč API a je třeba uvést pouze e-mailovou adresu. Chcete-li použít výběr slov k získání významů a funkcí automatického doplňování, musíte nastavit alespoň jeden strojový překlad.

Použijte menu „Options->Preferences“ pro zadání nastavení rozhraní.

nebo MT. Můžete nastavit nejnižší míru shody a jakou službu strojového překladu použít.



Vyberte slova, abyste získali význam

BasicCAT může zobrazit významy vybraného textu na základě strojového překladu a online slovníku. Tuto funkci musíte zkontrolovat v nastavení a povolit alespoň jednu službu strojového překladu.

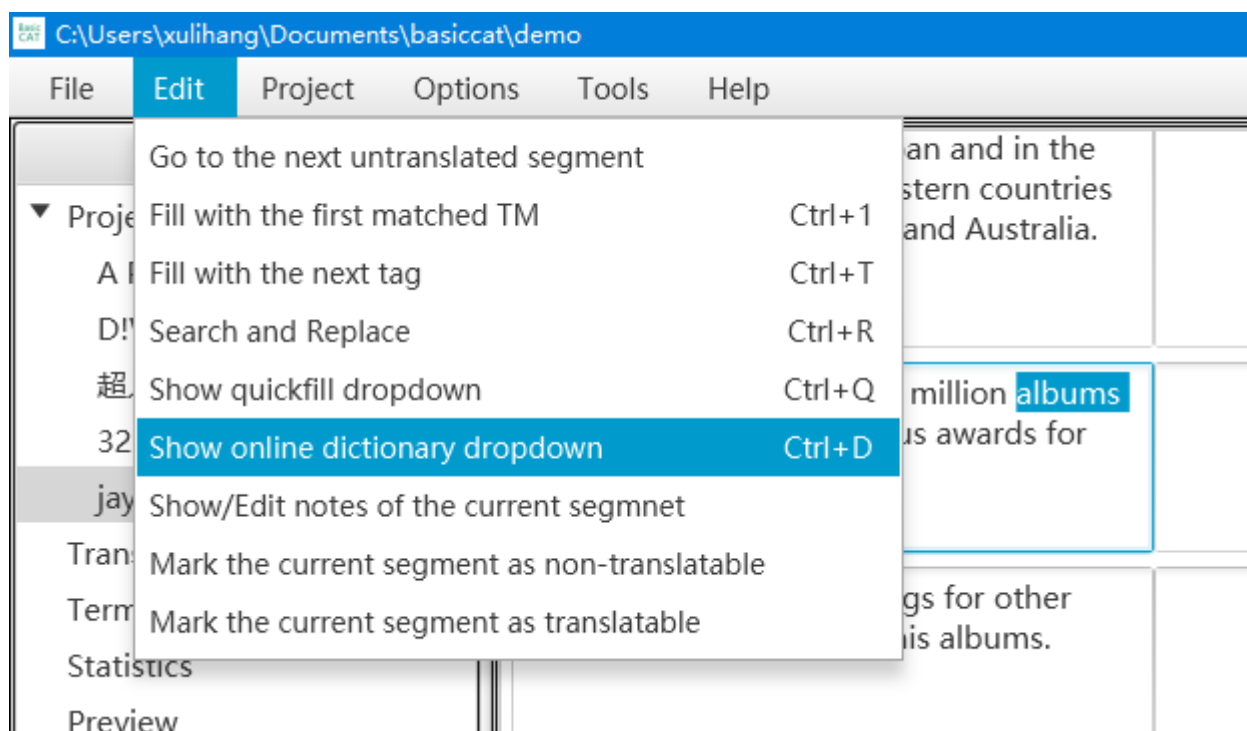
Jay Chou is a Taiwanese musician, singer, songwriter 周杰伦 and producer, film producer, actor, and director.

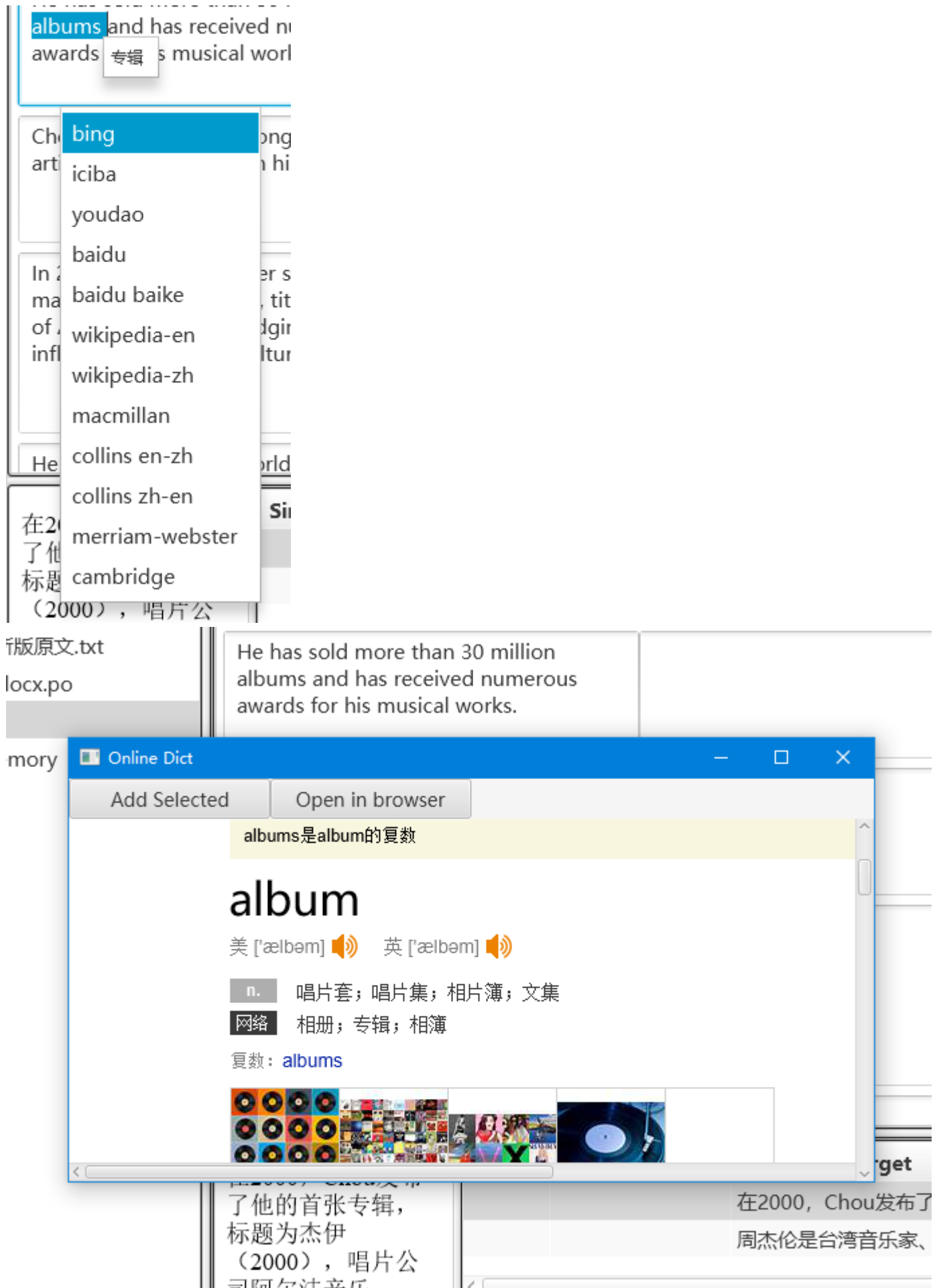
Výsledek do překladu můžete vyplnit kliknutím na položku v rozbalovacím seznamu.

Online slovník ¶

BasicCAT integruje online slovníky jako offline slovníky mají problémy s autorskými právy.

Vyberte slovo ve zdrojovém textu a klikněte na nabídku „Upravit-> Zobrazit rozevřací seznam online slovníku“ nebo použijte klávesovou zkratku CTRL + D. Zobrazí se rozevřací seznam jmen online slovníku. Otevře se okno procházení.





Pomocí tlačítka „Přidat vybrané“ vyplňte vybraný text v prohlížeči do textové oblasti překladu. Stisknutím tlačítka „Otevřít v prohlížeči“ otevřete stránku v místních prohlížečích.

Další online slovníky můžete přidat úpravou dictList.txt v konfigurační složce projektu.

Kontrola jazyka ¶

BasicCAT používá Language Tool ke kontrole jazykových chyb. Language Tool je otevřený zdroj kontroly pravopisu a gramatiky.

Když je segment přeložen a stisknutím klávesy „Enter“ přepnete na další, zkontroluje se předchozí segment. Ve spodní části se zobrazí chyby. Pod textovou oblastí překladu se také zobrazí rozevírací seznam obsahující možné opravy. Kliknutím na náhradní položku v dolní části nebo položkou rozevíracího seznamu budou chyby nahrazeny tou pravou.

The screenshot shows the BasicCAT interface with a translation of a Chinese text about raising flowers. The source text is in Chinese, and the target text is in English. A dropdown menu is open over the word 'raissing', showing 'raising' and 'Riesling'. A message box at the bottom indicates a spelling mistake found in the target text.

我可还没成为养花专家，因为没有工夫去研究和试验。	I don't have time to be an expert in raising flowers, as I have no time to study and experiment.
我只把养花当做生活中的一种乐趣，花开得大小好坏都不计较，只要开花，我就高兴。	raising Riesling
在我的小院子里，一到夏天满是花草，小猫只好上房去玩，地上没有它们的运动场。	
花虽然多，但没有奇花异草。	
珍贵的花草不易养活，看着一棵好花生病要死是件难过的事。	

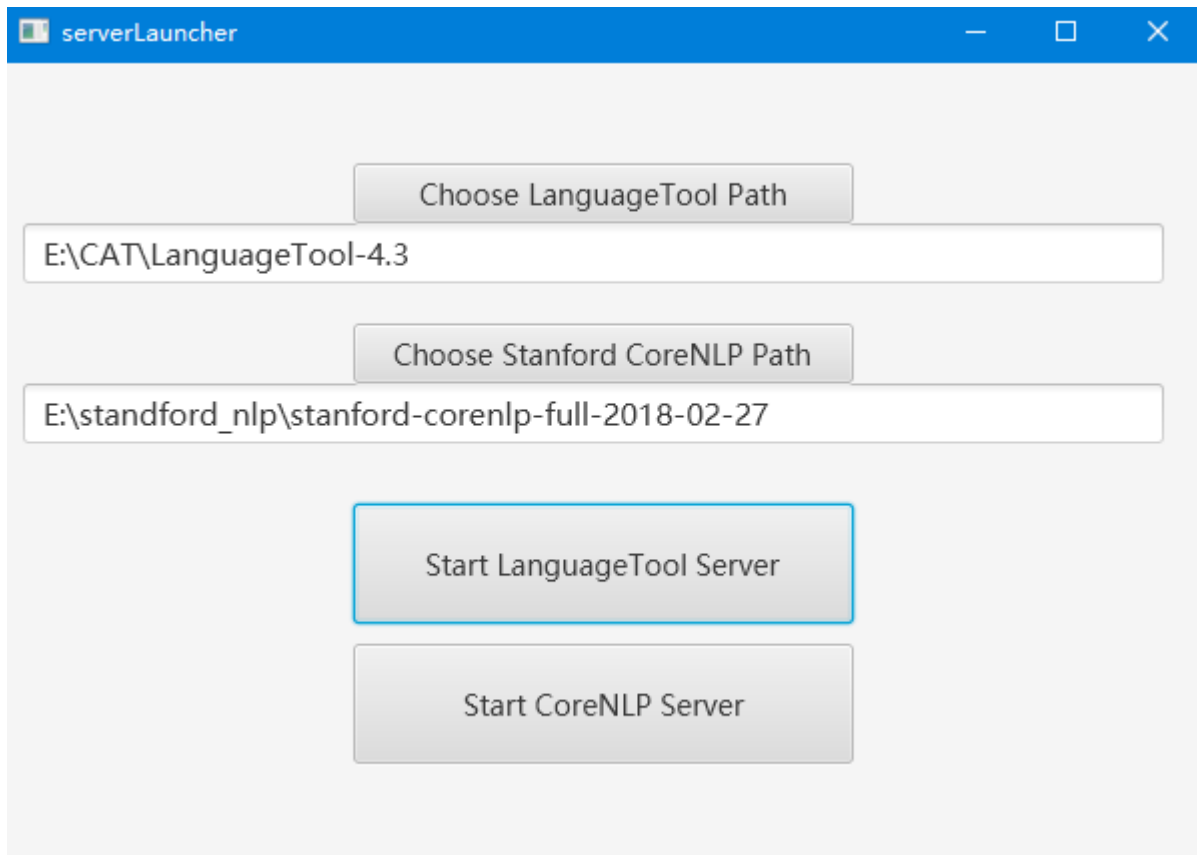
message: Possible spelling mistake found
context: I don't have time to be an expert in **raissing** flowers, as I have no time to study and experiment...
replacement: raising
replacement: Riesling

TM Match Segment Search LanguageCheck

Chcete-li použít jazykovou kontrolu, můžete buď použít API poskytované přímo LanguageTool (výchozí adresa: <https://languagetool.org/api/v2/check>), nebo stáhnout LanguageTool a spustit jej lokálně.

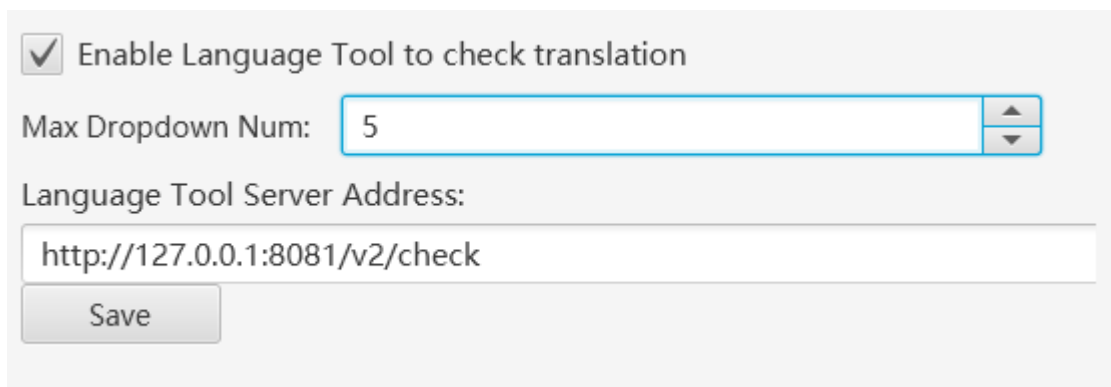
[LanguageTool 4.3 Stáhnout](#)

Rozbalte stažený soubor, otevřete BasicCAT, klikněte na nabídku „Nástroje-> Spouštěč serveru“ a můžete nastavit cestu ke složce obsahující soubor s názvem languagetool-server.jar. Stisknutím tlačítka „Spustit server LanguageTool Server“ spustíte server lokálně.



Můžete navštívit stránku <http://127.0.0.1:8081/v2/check?text=a%20example&language=en-US> a zkontrolovat, zda je server spuštěn.

Kromě toho je třeba zaškrtnout políčko „povolitelný jazykový nástroj“ v nastavení.



Můžete také nastavit maximální rozbalovací návrhy, aby se zabránilo zobrazování příliš mnoho návrhů.

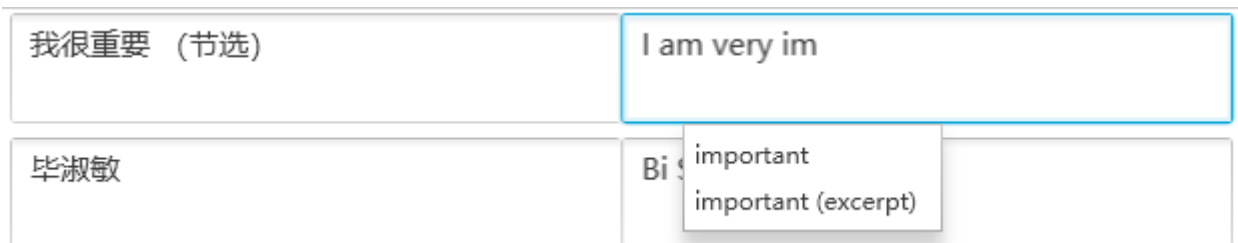
Autokompletní (Interaktivní strojový překlad) ¶

Zde funguje automatické dokončování: Zdrojový text je tokenizován a analyzován

pomocí Stanford NLP Tools. Poté budou extrahována slova a fráze a budou strojově přeloženy. Výsledky budou uloženy do paměti. Když uživatel zadá slovo, které odpovídá začátku těchto segmentů, zobrazí se rozevírací seznam obsahující návrhy odpovídajících překladů.



Pokud jde o překlad z čínštiny do angličtiny, může být také použit k rychlému vložení anglických slov.



Chcete-li použít automatické dokončování, potřebujete Stanford CoreNLP .

Domovská stránka Stanford CoreNLP:

<https://stanfordnlp.github.io/CoreNLP/index.html>

BaiduNetdisk (backup download):

<https://pan.baidu.com/s/1LNW4IDw8Viz6RURmzqxI9Q>

CoreNLP a čínský model jar jsou potřeba. Rozbalte CoreNLP a vložte čínský model jar do stejné složky. Spusťte server pomocí spouštěče serveru.

Navštivte stránku <http://127.0.0.1:9000> a zkontrolujte, zda je server spuštěn.

Stejně jako nastavení jazykového nástroje, musíte také zaškrtnout zaškrtačkové políčko „Povolit automatické dokončení“.

Změnit odkaz na použití vzdáleného serveru.

Můžete také nastavit maximální rozevírací návrhy, aby se zabránilo zobrazování příliš mnoho návrhů.

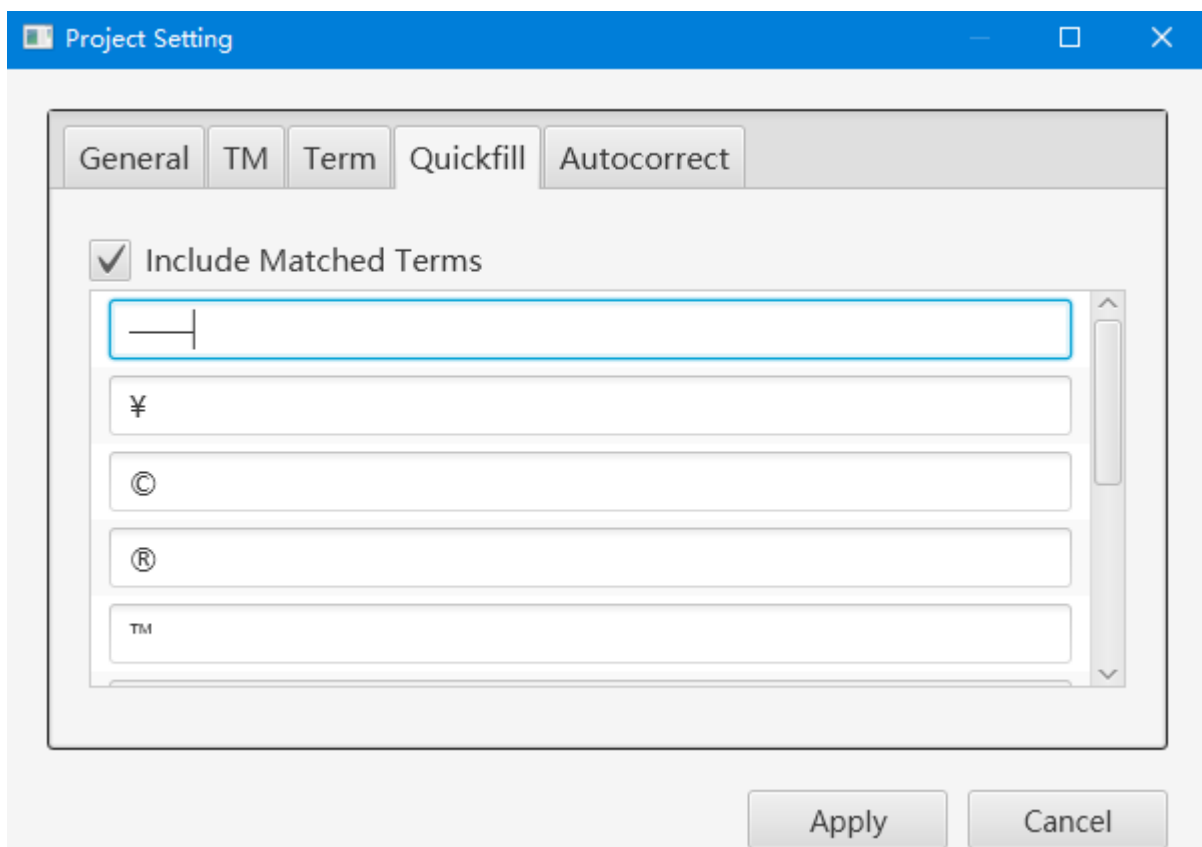
Quickfill¶

Často musíme mnohokrát zadávat speciální znaky nebo stejný text. BasicCAT podporuje rychlé vyplňování. Můžete použít klávesovou zkratku CTRL + Q nebo kliknout na nabídku „Upravit - zobrazit rozevírací seznam pro rychlé vyplnění“ a zobrazit rozevírací seznam položek s rychlým vyplněním. Lze také zahrnout

odpovídající výrazy.



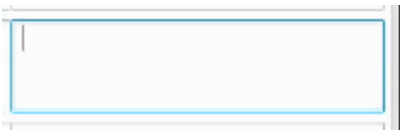
Klikněte na menu „Project->Nastavení projektu“ pro nastavení rychlého naplnění.



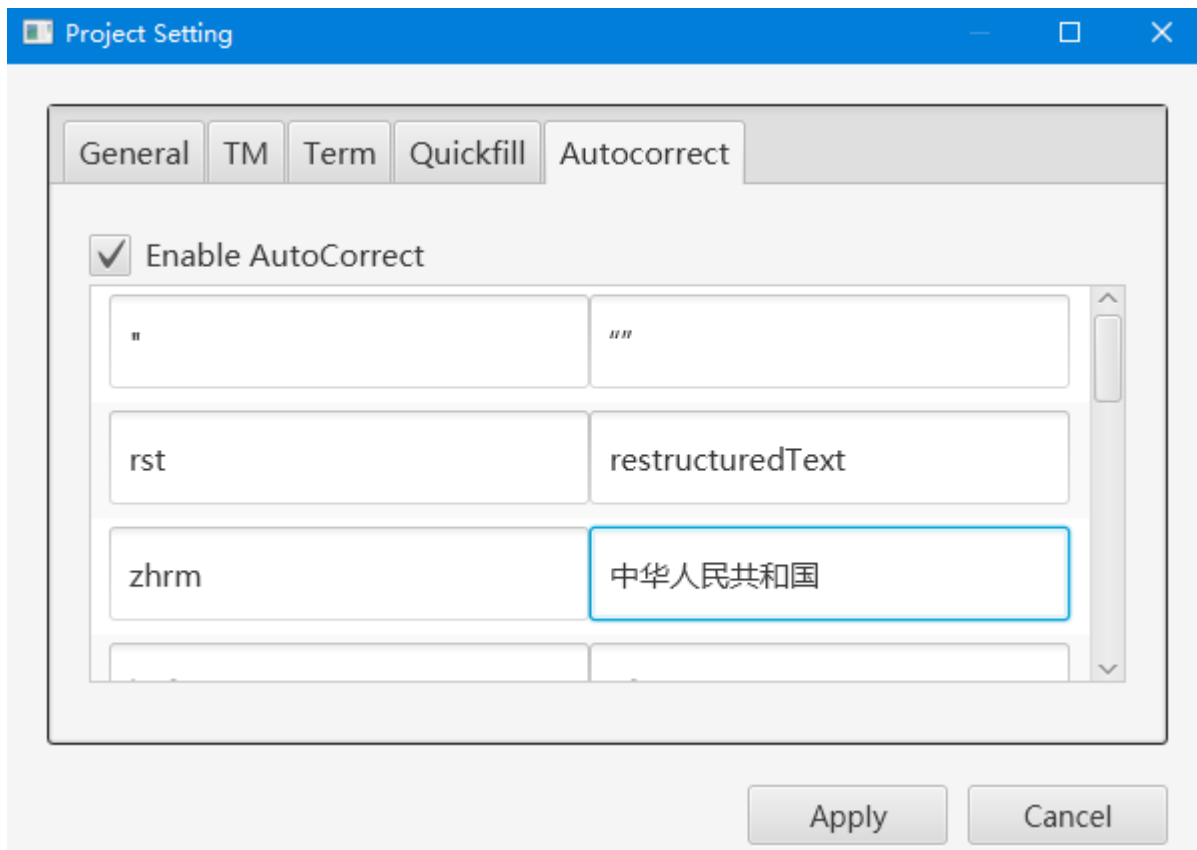
Automatická korekce

Automatické opravy jsou inspirovány funkcí aplikace Microsoft Word. Dokáže detekovat vložený text a opravit pravopisné chyby. Například při překladu z angličtiny do čínštiny musíme zadávat čínskou interpunkci a automatická oprava může nahradit anglickou interpunkci čínskými interpunkcemi. Můžeme jej také použít k rychlému zadávání obsahu. Například „rst“ je zkratka „restructuredText“.

Když je nastavena funkce AutoCorrect, můžeme zadat „rst“ get „restructuredText“.



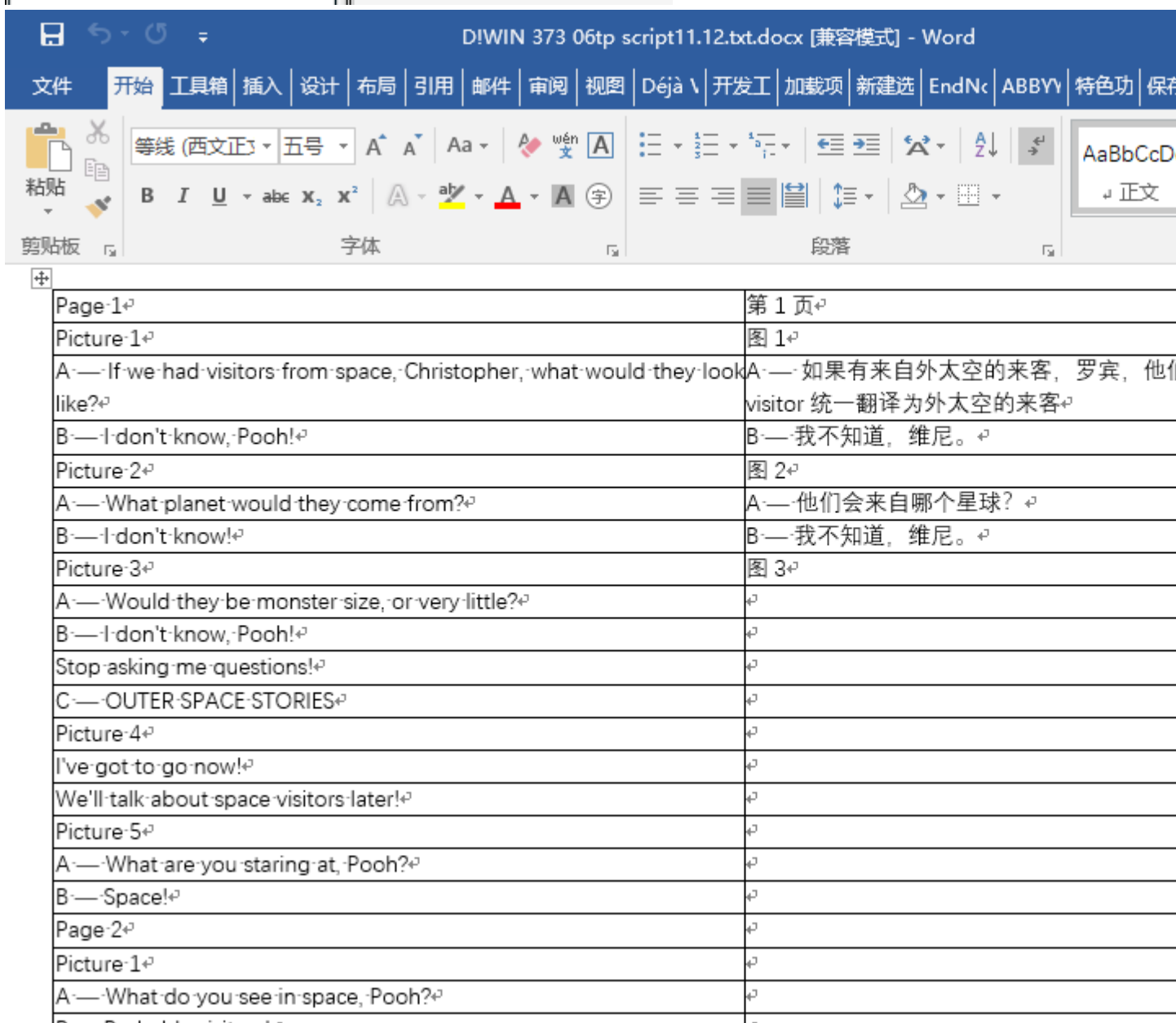
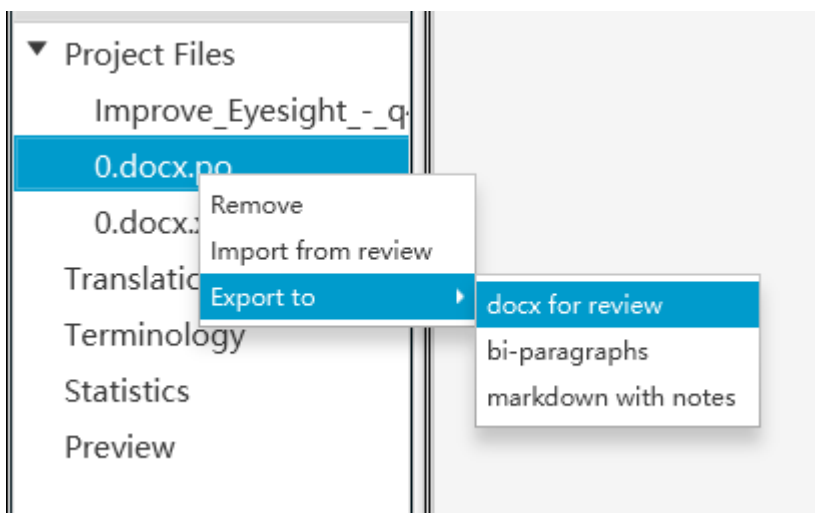
Klikněte na menu „Project->Nastavení projektu“ pro nastavení automatické korekce.



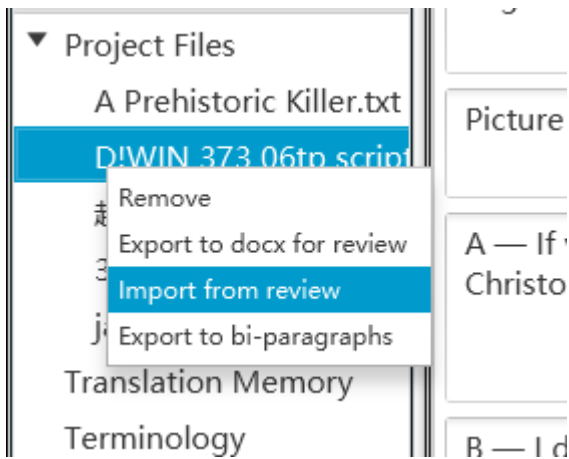
Exportovat Slovo pro recenzi ¶

BasicCAT může exportovat přeložený obsah docx souborů a používat Word ke kontrole.

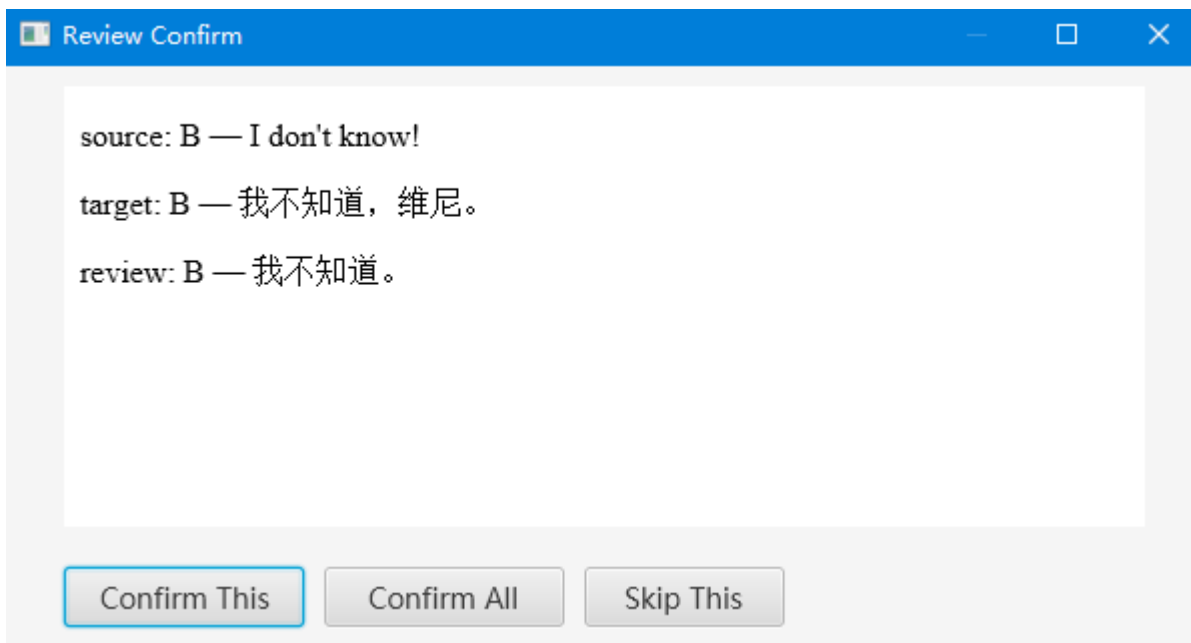
Klikněte pravým tlačítkem myši na soubor a klikněte na „Export pro docx pro kontrolu“ exportovat soubor docx.



Po dokončení kontroly můžete výsledek importovat zpět. Klikněte pravým tlačítkem na soubor a klikněte na „Importovat z recenze“.

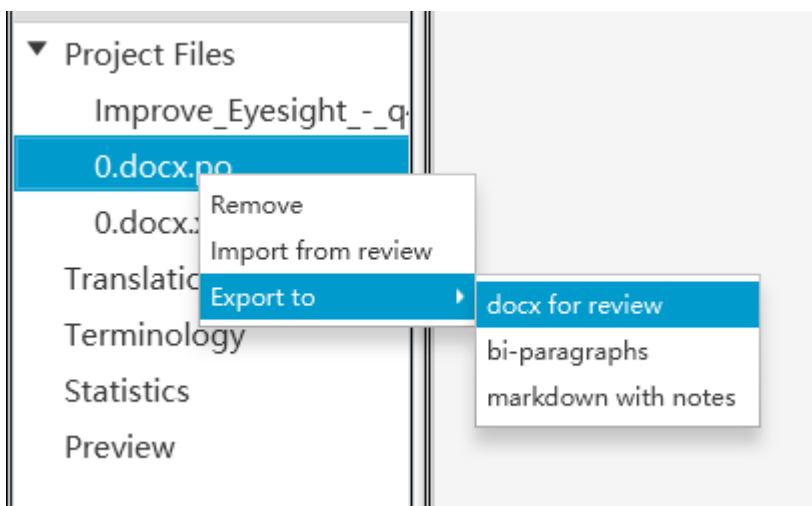


Můžete zkontrolovat importovaný výsledek jeden po druhém nebo nahradit původní překlad s přezkumem přímo.



Exportovat dvojjazyčné soubory odstavce ¶

Přímo na soubor klikněte na „Exportovat do bi-paragraphů“, abyste to udělali.



Export do markdownu s poznámkami ¶

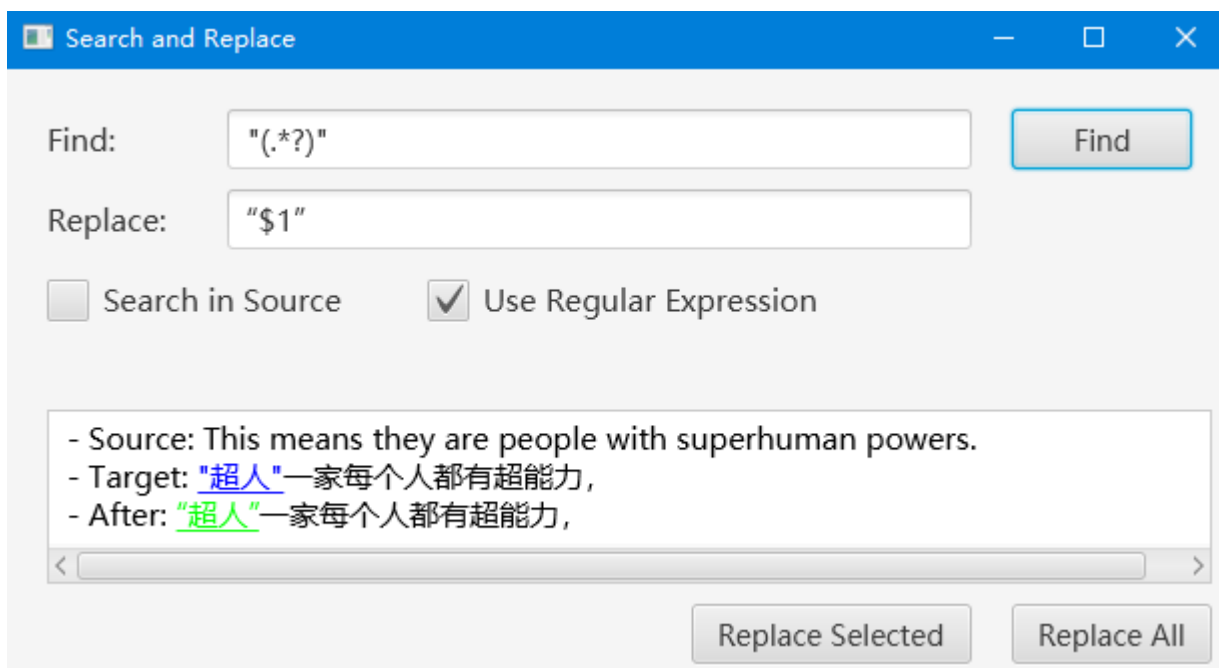
Stejně jako předchozí export, klikněte pravým tlačítkem myši na název souboru a klikněte na export na-> markdown s poznámkami získat soubor markdown.

Markdown soubory lze převést na soubory slov později pomocí Pandoc.

Hledat a nahradit ¶

Kliknutím na nabídku „Upravit-> Hledat a nahradit“ otevřete dialogové okno Hledat a nahradit. Můžete hledat ve zdrojovém nebo cílovém textu. Regulární výrazy jsou podporovány.

Níže je uveden příklad nahrazení anglických uvozovek čínskými uvozovkami.



O regulárních výrazech se můžete dozvědět více na těchto webových stránkách.

Překlady všech druhů souborů

Soubory txt

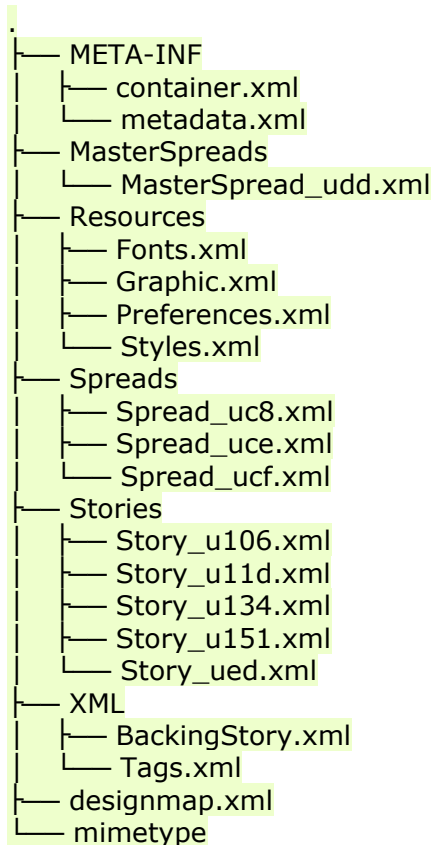
Soubory TXT jsou čistě textové soubory. Neexistují žádné další značky a všechny segmenty lze sloučit a rozdělit.

IDML soubory ¶

IDML je formát souboru založený na XML používaný aplikací Adobe InDesign. Dokumenty vytvořené nejnovější verzí aplikace InDesign musí být převedeny do IDML, aby mohly být použity ve starých verzích aplikace InDesign. Výchozí formát souboru aplikace InDesign je indd, což je proprietární formát. Nelze jej otevřít aplikacemi třetích stran. Aby bylo možné přeložit dokumenty InDesignu, je třeba

soubory Indd převést na soubory IDML.

Soubor IDML je přirozeně komprimovaný soubor. Jeho struktura je následující.



Co je pro nás důležité, jsou designmap.xml, rozložení složky, složky příběhů a zdroje složky.

designmap.xml definuje základní strukturu dokumentu. Soubor rozšíření obsahuje strukturu jedné stránky nebo protilehlých stránek. Soubor příběhu obsahuje text zobrazený v textovém rámečku. Fonts.xml ukládá informace o písmu a Styles.xml ukládá informace o stylech.

BasicCAT čte designmap.xml a šíří soubory, aby získal pořadí příběhů v dokumentu a extrahuje texty z příběhových souborů.

Níže je, jak příběh soubor vypadá:

```
<?xml version="1.0" encoding="UTF-8"?>
<idPkg:Story xmlns:idPkg="http://ns.adobe.com/AdobeInDesign/idml/1.0/packaging"
DOMVersion="13.0">
  <Story Self="u19caf" AppliedTOCStyle="n" UserText="true" IsEndnoteStory="false"
TrackChanges="false" StoryTitle="$ID/" AppliedNamedGrid="n">
  <StoryPreference OpticalMarginAlignment="false" OpticalMarginSize="12"
FrameType="TextFrameType" StoryOrientation="Horizontal"
StoryDirection="LeftToRightDirection" />
  <InCopyExportOption IncludeGraphicProxies="true" IncludeAllResources="false" />
  <ParagraphStyleRange AppliedParagraphStyle="ParagraphStyle/Intro Copy">
    <CharacterStyleRange AppliedCharacterStyle="CharacterStyle/$ID/[No character style]">
      <Content>"No capes!" That's what Edna says in the first film when Mr. Incredible wants
a cape on his new Supersuit. She knows that capes can be dangerous for Supers. A cape caused one
Super to get pulled into a jet turbine, and another was sucked into a spinning </Content>
    </CharacterStyleRange>
    <CharacterStyleRange AppliedCharacterStyle="CharacterStyle/Intro Copy Bold">
      <Content>vortex</Content>
```

```

</CharacterStyleRange>
<CharacterStyleRange AppliedCharacterStyle="CharacterStyle/$ID/[No character style]">
  <Content>. Pláště mohou způsobit i další problémy. Uvidíme, co jsou zač.</Content>
</CharacterStyleRange>
<CharacterStyleRange AppliedCharacterStyle="CharacterStyle/$ID/[No character style]"
FillColor="Color/Black" FontStyle="300" PointSize="12">
  <Properties>
    <Leading type="unit">20</Leading>
  </Properties>
  <Br />
</CharacterStyleRange>
</ParagraphStyleRange>
<ParagraphStyleRange AppliedParagraphStyle="ParagraphStyle/Body Copy">
  <CharacterStyleRange AppliedCharacterStyle="CharacterStyle/$ID/[No character style]" />
</ParagraphStyleRange>
</Story>
</idPkg:Story>

```

Informace o odstavci jsou uloženy ve značce OdstavecStyleRange. Jedna značka SectionStyleRange může obsahovat mnoho značek CharacterStyleRange, které obsahují texty. V IDML existují dva typy stylu. Jeden je globální styl a druhý je místní styl. Globální styly jsou uloženy v souboru Styles.xml. Soubory příběhu používají k označení globálních stylů atributy AppliedParagraphStyle a AppliedCharacterStyle. Místní styly jsou definovány v atributech a prvku vlastností v souborech příběhu, jako je PointSize a FontStyle.

BasicCAT převede soubor s příběhem výše do označeného textu níže.

"Žádné pláštěnky!" To říká Edna v prvním filmu, když Mr. Incredible chce na svém novém Supersuitu pláštěnku. Ví, že peleríny mohou být pro Supers nebezpečné. Plášť způsobil, že jeden Super byl zatažen do tryskové turbíny a další byl nasáván do rotace vír
. Pláštěnky mohou způsobit i další problémy. Uvidíme, co jsou zač.

```

<c0 id="3"><p16><c0></c0><c3></c3><c0></c0></p16><p3><c0></c0></p3>

```

Čísla v p16, c0 odpovídají hodnoti nějakého globálního stylu ve style.xml. Atribut id<c0 id="3"> odpovídá hodnoti řady znaků v souboru příběhu. Číslo se používá ke čtení místních stylů při generování cílových souborů. c0 je výchozí styl, který nemá žádné speciální formáty. BasicCAT ji skryje v zobrazeném zdrojovém textu. Značky odstavců a značky bez textů lze také skrýt.

Když příklad otevře BasicCAT, výsledek vypadá takto:

"No capes!"	没有披风!
That's what Edna says in the first film when Mr. Incredible wants a cape on his new Supersuit.	当超能先生想给他新的超人装要一个披风时，衣夫人是这么回答的。
She knows that capes can be dangerous for Supers.	她知道披风可能会威胁到超人的安全。
A cape caused one Super to get pulled into a jet turbine, and another was sucked into a spinning vortex.	披风可能会将超人卷入飞机的<c3>涡轮引擎</c3>,
Capes could cause other problems, too.	还可能造成其它的问题。
Let's see what they are.	让我们看一下它们是什么吧!

Značka c3 přidá do textu tučný styl písma. Anglická písma mají často mnoho vah písma, jako Extra Light, Light, Normal a Heavy, zatímco čínská písma mají často pouze jednu normální váhu písma. Tučného stylu čínských znaků ve Wordu je dosaženo softwarovým algoritmem, ale IDML k tomu používá váhu písma. Abychom zajistili, že značky budou mít efekt na překladové projekty z angličtiny do čínštiny, používá BasicCAT jako zdrojové písmo IDML čínské písmo Source Han Serif. Zdroj Han Serif má 7 vah písma. Názvy vah písma u anglických písem budou převedeny na názvy v Source Han Serif následujícím způsobem.

- 100-> ExtraLight
- 200-> ExtraLight
- 300-> Světlo
- 400-> Pravidelné
- 500-> střední
- 600-> semibold
- 700-> Bold
- 800-> Těžké
- 900-> Těžké
- Normální-> Pravidelné
- Black-> Heavy

Takže musíte nainstalovat zdroj Han Serif pro zobrazení čínštiny v InDesign.

[Stáhnout zdroj Han Serif](#)

Překlad značek do IDML není nutný. Pokud v cílových textech nejsou žádné odpovídající značky, budou vyplněny značky c0, aby se zajistilo, že překlad nebude vynechán.

Soubory XLIFF ¶

XLIFF je formát XML pro zpracování extrahovaného textu k překladu. Nástroje CAT často používají tento formát k ukládání extrahovaného textu ze souborů, jako jsou docx, html a idml. Po dokončení překladu vygenerují nástroje CAT cílové soubory na základě souborů XLIFF.

XLIFF je standard formulovaný organizací OASIS a může být použit jako přechodný formát mezi různými překladatelskými programy.

XLIFF používá abstraktní placeholder tagy zděděné z opentag a zapouzdřovací značky zděděné po TMX, aby reprezentovaly speciální formáty.

Například, existují dva způsoby, jak reprezentovat „To je tučné.“ v XLIFF.

Abstraktní zástupce:

```
<trans-unit id="1">
<source>This is <g id="1">bold</g>.</source>
</trans-unit>
```

Zapouzdření:

```
<trans-unit id="1">
<source>This is <bpt id="1">\b</bpt>bold<ept id="1">\b0</ept>.</source>
</trans-unit>
```

Zástupný způsob abstrahuje značky na své přednastavené zástupné značky. Pro označení tučného písma používá HTML a RTF \ b. Všichni se stali<g>v XLIFF. Zobrazuje také méně informací o značkách. Ale po této abstrakci nemůžeme říci, jakou funkci má tato značka.

Při překladu souborů XLIFF budou ve zdrojovém textu značky. Pokud tagy nejsou vloženy do cílového textu, BasicCAT je vyplní na konci cílového textu. To však může vést k neúplnému obsahu.

Pozor!

Pokud tagy nejsou vloženy do cílového textu, BasicCAT je vyplní na konci cílového textu. To však může vést k neúplnému obsahu.

Níže je XLIFF extrahován okapi z příkladu souboru v IDML části.

```
<?xml version="1.0" encoding="UTF-8"?>
<xliff version='1.2'
  xmlns='urn:oasis:names:tc:xliff:document:1.2'>
<file original="Stories/Story_u19caf.xml" source-language="en-US" target-language="zh-CN"
  datatype="xml">
<body>
<trans-unit id="NB085C0FA-tu1" xml:space="preserve">
  <source xml:lang="en-US"><g id="1">"Žádné pláštěnky!" To říká Edna v prvním filmu, když Mr.
```

Incredible chce na svém novém Supersuitu pláštěnku. Ví, že peleríny mohou být pro Supers nebezpečné. Plášť způsobil, že jeden Super byl zatažen do tryskové turbíny a další byl nasáván do rotace</g><g id="2"> vír</g><g id="3"> . Pláštěnky mohou způsobit i další problémy. Uvidíme, co jsou zač.</g></source>

<target xml:lang="zh-CN"><g id="1">"Žádné pláštěnky!" To říká Edna v prvním filmu, když Mr. Incredible chce na svém novém Supersuitu pláštěnku. Ví, že peleríny mohou být pro Supers nebezpečné. Plášť způsobil, že jeden Super byl zatažen do tryskové turbíny a další byl nasáván do rotace</g><g id="2"> vír</g><g id="3"> . Pláštěnky mohou způsobit i další problémy. Uvidíme, co jsou zač.</g></target>

</trans-unit>

</body>

</file>

</xliff>

Když soubor XLIFF otevře BasicCAT, výsledek vypadá níže. Jednotlivá značka a spárované značky na obou stranách segmentu budou skryty.

"No capes!"	
That's what Edna says in the first film when Mr. Incredible wants a cape on his new Supersuit.	
She knows that capes can be dangerous for Supers.	
A cape caused one Super to get pulled into a jet turbine, and another was sucked into a spinning </g><g id="2">vortex</g><g id="3">.	
Capes could cause other problems, too.	
Let's see what they are.	

Soubory po

PO je formát podobný XLIFF. Je navržen tak, aby nejprve lokalizoval programy C.

Níže je obsah PO extrahován okapi z příkladu IDML story souboru.

msgctxt "okpCtx: sd = 197: tu = NB085C0FA-tu1"

msgstr "" Žádné pláštěnky! " To říká Edna v prvním filmu, když pan Incredible chce na svém novém Supersuitu pláštěnku. Ví, že pláštěnky mohou být pro Supery nebezpečné. Plášť způsobil, že

jednoho Supera vtáhli do tryskové turbíny a další byl nasáván do rotujícího vír. Mys může způsobit i další problémy. Podívejme se, co to je. "

msgstr ""<g1></g1><g2></g2><g3></g3>

msgctxt ukládá kontextové informace. msgid uloží zdrojový text a msgstr uloží cílový text. Soubory PO vytvořené okapi také používají značky k reprezentaci formátů.

Níže je uvedeno, jak to vypadá při překladu souboru PO v BasicCAT. Stejně jako XLIFF budou skryty jednoduchá značka a spárované značky na obou stranách segmentu.

"No capes!"	
That's what Edna says in the first film when Mr. Incredible wants a cape on his new Supersuit.	
She knows that capes can be dangerous for Supers.	
A cape caused one Super to get pulled into a jet turbine, and another was sucked into a spinning </g1> <g2>vortex</g2> <g3>.	
Capes could cause other problems, too.	
Let's see what they are.	

PDF Soubory ¶

Soubor PDF je obtížně zpracovatelný. Lze jej převést do formátu docx pomocí nástrojů, jako jsou Word, ABBYY, Solid Document Converter atd. Původní formát však nelze dobře zachovat. Adobe Acrobat lze použít k úpravě textu, ale existuje mnoho omezení.

Soubory PDF jsou často generovány soubory v jiném formátu, například docx, idml atd. Je lepší zacházet se zdrojovým souborem přímo. Pokud chcete zachovat formáty, ale nemáte zdrojový soubor, musíte nejprve přeložit text a znovu publikovat v počítači.

Při překládání dokumentů aplikace InDesign vydavatelství často předávají dokument PDF překladateli. Překladač zadá překlad do aplikace Word a poté jej dá sazeči, aby v InDesignu nahradil zdrojové texty cílovými texty. V takovém případě musí překladatelé extrahovat text z PDF.

BasicCAT má nástroj PDF to Text. Je přístupný z nabídky „Nástroje-> PDF2TXT“. Pokud lze text PDF extrahovat, lze jej extrahovat přímo stisknutím tlačítka „strip“. Pokud ne, použije se open source OCR software Tesseract.

K extrahovanému textu můžete přidat číslo stránky. Může také přidat protilehlé stránky, jako je strana 4-5.



Texty v PDF neobsahují informace o odstavcích. Jsou to jen pevné tečky na stránce. Extrahovaný text může mít na konci každého řádku nový řádek. PDF2TXT poskytuje funkci „přeformátování“ k odstranění dalších nových řádků.

O instalaci tesseractu si uživatelé Windows mohou stáhnout kopii odtud. PDF2TXT požádá o cestu k tesseract.exe.

Uživatelé systémů Linux a macOS (s homebrew) si mohou nainstalovat tesseract-ocr přímo. Možná budete muset stáhnout další jazykové modely.

Používání Okapi k překládání souborů v jiných formátech ¶

Okapi je sada překladových komponent. Má Checkmate pro kontrolu kvality překladu, Ratel pro úpravu segmentačního pravidla a Rainbow pro provádění všech druhů překladových a lokalizačních úkolů.

Chcete-li překládat soubory v jiných formátech, musíme použít Rainbow.

1. Vytvořit soubory XLIFF nebo PO ze zdrojových souborů

Otevřete Rainbow a přetáhněte zdrojové soubory do okna.



Nastavit zdrojový jazyk, cílové jazyky, kódování souborů, atd.



Klikněte na nabídku „Utilities-> Translation Kit Creation“. Můžete se rozhodnout generovat XLIFF, PO nebo jiné přechodné formáty. Výchozí výstupní cesta je složka zdrojových souborů.



Po dokončení spuštění bude ve výstupní cestě složka s názvem pack1. V pracovní složce jsou generovány soubory.

2. Generovat cílové soubory z dříve vytvořeného projektu okapi

Po dokončení překladu vložte přeložené soubory zpět do pracovní složky. Otevřete Rainbow a přetáhněte do něj manifest.rkm.



Klikněte na menu „Utilities->Translation Kit Post-processing“ pro generování cílových souborů.

Navštivte zde, abyste viděli, jaké formáty Okapi podporuje.

Od verze 1.5 má BasicCAT integrovaný Okapi Tikal. Může automaticky převádět další soubory formátu do souborů XLIFF a automaticky generovat cílové soubory z přeložených souborů XLIFF. Pokud však potřebujete upravit parametry, které Okapi používá ke zpracování formátu, stále musíte použít Rainbow.

Pomocí Felixe překládat soubory slov ¶

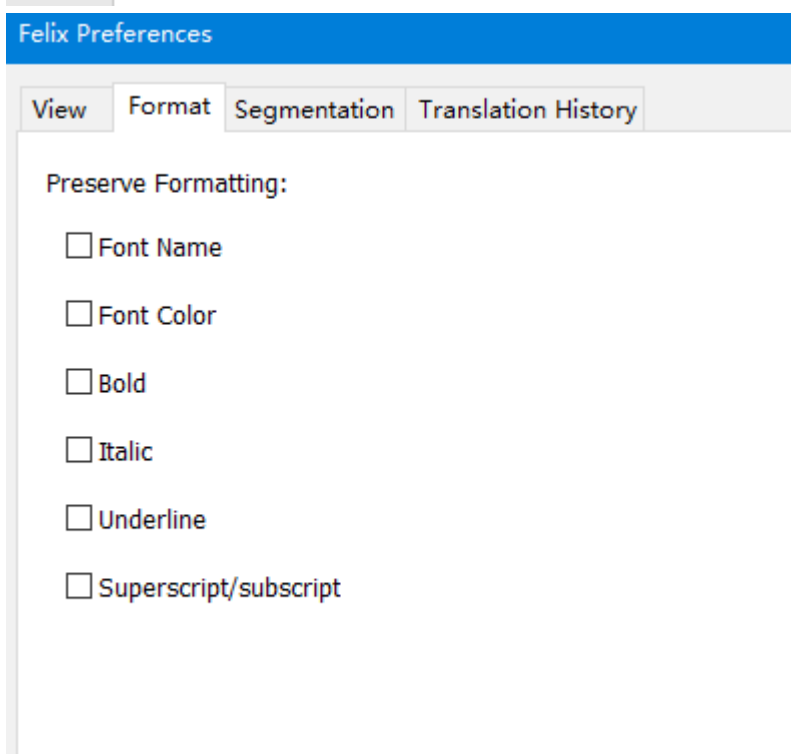
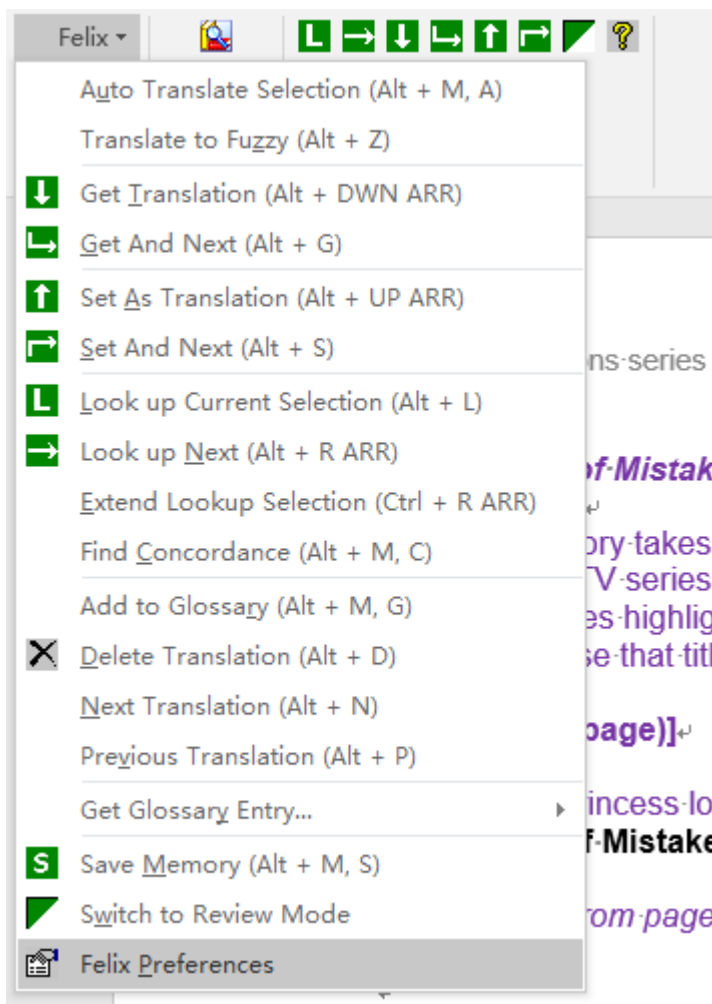
Jedním ze způsobů překladu souborů aplikace Word je převod slova do formátu xliiff pomocí okapi. Dalším způsobem je použití Felix CAT.

Felix je sada maker běžících na Microsoft Office. Můžete překládat přímo do aplikace Word a používat překladovou paměť.

O tom, jak používat Felix, navštivte jeho domovskou stránku a dozvědět se více.

Zde je, jak používáme BasicCAT přeložit text docx souborů a použít Felix k vyplnění překladu v aplikaci Word.

1. Exportovat Word do TXT pro překlad do BasicCAT.
2. Po provedení překladu exportujte překlad do TMX a importujte ji do Felixe.
3. Na kartě Doplněk aplikace Word máte přístup k rozhraní Felix. V nastavení Felixu zrušte zaškrtnutí políček na kartě Formát. Tímto způsobem, když Felix nahradí zdrojový text cílovým textem, zůstanou zachovány původní formáty.



4. Můžete použít klávesu zkratka ALT+G pro rychlé vyplnění překladů.

The screenshot displays two windows. The top window is Microsoft Word, showing a document with Chinese text and a highlighted English sentence: "This story takes place shortly after the events of the film, of the TV series." The bottom window is a translation memory tool, showing a match for the highlighted sentence with the Chinese translation "这个故事发生在...".

Spolupráce ¶

Sdílení překladové paměti a termínu ¶

Serverový program BasicCAT je potřebný k nastavení služby překladové paměti a sdílení termínů. Program je k dispozici na stránce stahování BasicCAT.

Na straně klienta jsou nutná příslušná nastavení.



Jak spustit program serveru (vyžaduje java8+):

```
$ nohup java -jar CloudKVS_Server.jar &
```

Můžete upravit key.txt ve složce programu a nastavit přístupový klíč. Klíčem může být libovolný jednořádkový text.

Používat Git ke spolupráci ¶

BasicCAT může vytvořit místní úložiště git a nahrát jej do vzdáleného úložiště, když jsou nastaveny vzdálené URI a informace o účtu. Při žádosti o git push program načte nejnovější změny a aktualizuje místní soubory. Poté budou nahrány nové změny v místním repo. BasicCAT automaticky řeší konflikty podle doby vytvoření překladu segmentu. Skoro nenarazíte na žádné konflikty.

Musíte nastavit vzdálené URI v nastavení projektu a zvolit, zda se mají změny načít na vzdálené při každém provedení operace uložení. Nejprve musíte nastavit účty git v nastavení předvoleb.

Můžete také použít menu pro ruční provádění akcí git.

Spolupráce pomocí git nebude synchronizovat překladovou paměť a termíny. Synchronizuje pouze pracovní soubory.

Použití GitHub ¶

GitHub je populární hostitelská platforma Git. Můžeme na tom spolupracovat.

Nejprve vytvořte prázdné úložiště.

Zobrazí se tato stránka nastavení:

Pomocí BasicCAT otevřete projekt a nastavte vzdálený repozitář.

Pak můžete projekt zabalit jiným překladatelům.

Ostatní překladatelé musí mít svůj vlastní GitHub účet a mít autorizované oprávnění pro odesílání dat do úložiště.

V nastavení repozitáře klikněte na Spolupracující vlevo a přidejte další překladatelský účet GitHub.



Více se dozvíte v systému GitHub Help.

Ostatní ¶

Nastavení vzhledu

Kliknutím na nabídku „Možnosti-> Předvolby“ otevřete okno Předvolby. Vlevo klikněte na „Vzhled“.

Pro tuto chvíli je podporováno pouze nastavení fontu editační oblasti.



Automatické zálohování

BasicCAT podporuje automatické zálohování. Můžete nastavit jeho časový interval. Záložní soubory budou uloženy ve složce bak projektu.



Ovládání verzí pomocí systému Git ¶

BasicCAT přichází s git a může provádět příkaz git add a git commit pokaždé, když je projekt uložen.

Informace o uživateli git commit můžete nastavit v Předvolbách. Pokud potřebujete synchronizovat se vzdálenými úložišti, můžete nastavit heslo.



Pomocí git show můžete zjistit, jaké jsou nejnovější změny. Chcete-li to provést, musíte si do počítače nainstalovat Git. Pomocí git reset se můžete vrátit k předchozí verzi.

```

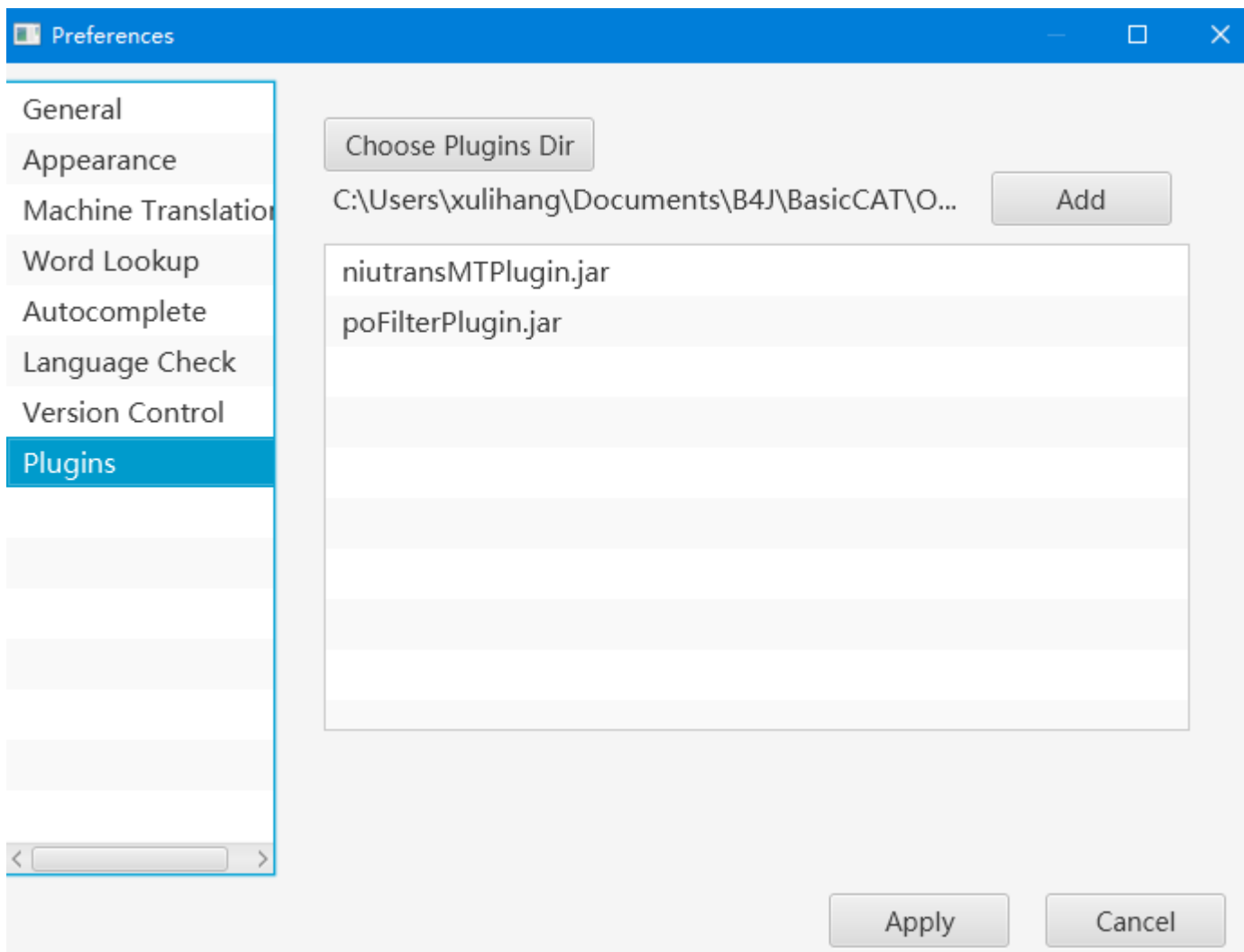
@@ -40,21 +40,21 @@
    ],
    [
        "A - What planet would they come from?",
-       "A - 他们会来自那个星球?",
+       "A - 他们会来自哪个星球?",
        "A - What planet would they come from?\r\n",
        "D!WIN 373 06tp script11.12.txt",
        {}
    ],
    [
-       "B - I don't know!",
+       "B - 我不知道, 维尼。",
        "B - I don't know!\r\n",
        "D!WIN 373 06tp script11.12.txt",
        {}
    ],
    [
-       "Picture 3",
+       "图3",
        "Picture 3 \r\n",
        "D!WIN 373 06tp script11.12.txt",
        {}
    ]

```

Další funkce lze realizovat instalací klienta git desktop.

Řízení Plugins ¶

BasicCAT v současné době podporuje dva druhy pluginů. Jedním je plugin pro strojový překlad a druhým je plugin pro filtrování. V předvolbách můžete nastavit složku s doplňky a přidat nebo odebrat doplňky.



Jak přispět

BasicCAT je open source projekt s licencí pod GPLv2. Oceníme příspěvky přispívající k růstu BasicCAT.

Pokud jste vývojář, můžete vyvinout pluginy nebo požádat o stažení pro zlepšení kódu BasicCAT.

Pokud jste uživatel a narazíte na některé problémy nebo máte nějaké návrhy, můžete otevřít problém v github repozitáři BasicCAT.

Chcete-li přeložit webové stránky BasicCAT, dokumentaci nebo její rozhraní, můžete také otevřít problém aplikovat.

© Copyright 2018, Xu Lihang Revize 98a0e09a.

Postavena se Sphinx pomocí tématu poskytovaném Read the Docs.
Přečtěte si Docs v: [nejnovější](#)

Verze
[nejnovější](#)

Soubory ke stažení
[HTML](#)

[epub](#)

On Read the Docs

[Domov projektu](#)

[Sestavení](#)

Hosting dokumentů zdarma od Read the Docs.