Machine Learning System Architecture

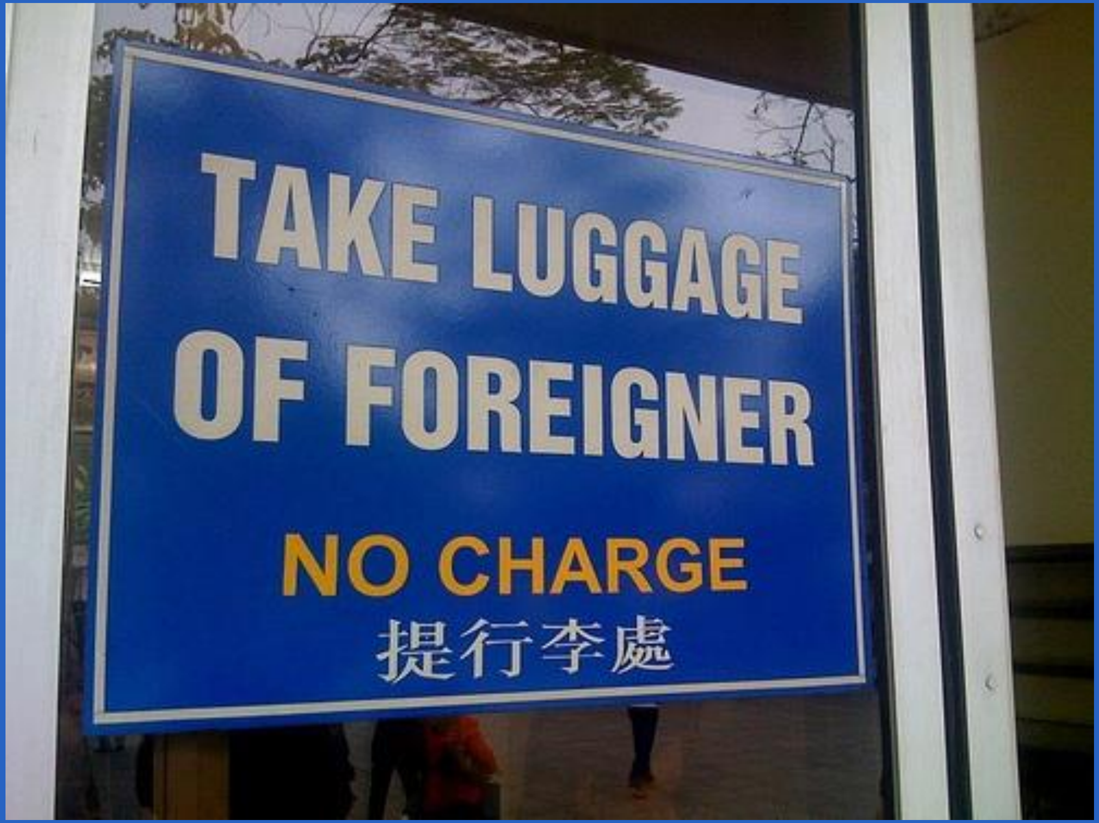
Microsoft Translator, a Case Study.

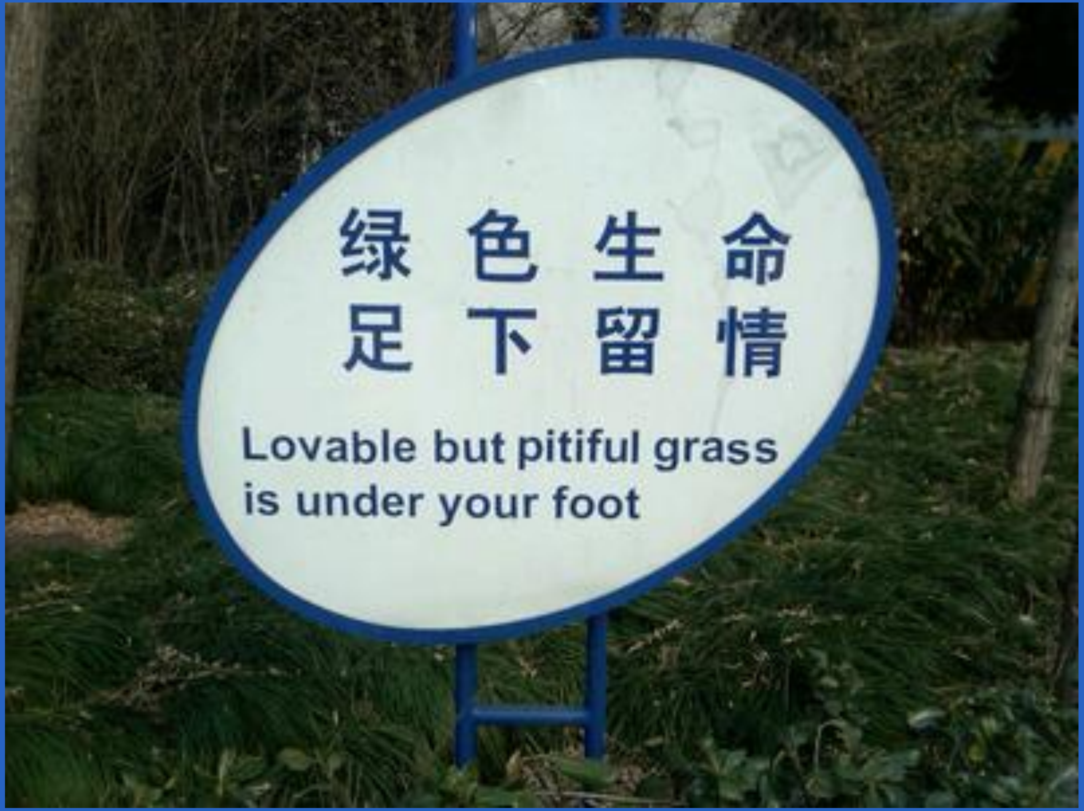
Vishal Chowdhary (@talktovishal, vishalc@microsoft)

What do we want to do?

Microsoft® Translator

What do you think we do?





Microsoft® Translator

Why is Machine Translation difficult?

Ambiguous Input

Kids make nutritious snacks

A million people live on water

Variable Input

why is this complicated

why is this difficult

Divergence between languages

word-order (syntactic structure), morphology, idiomatic expressions

Evolving

just 4get abt it!

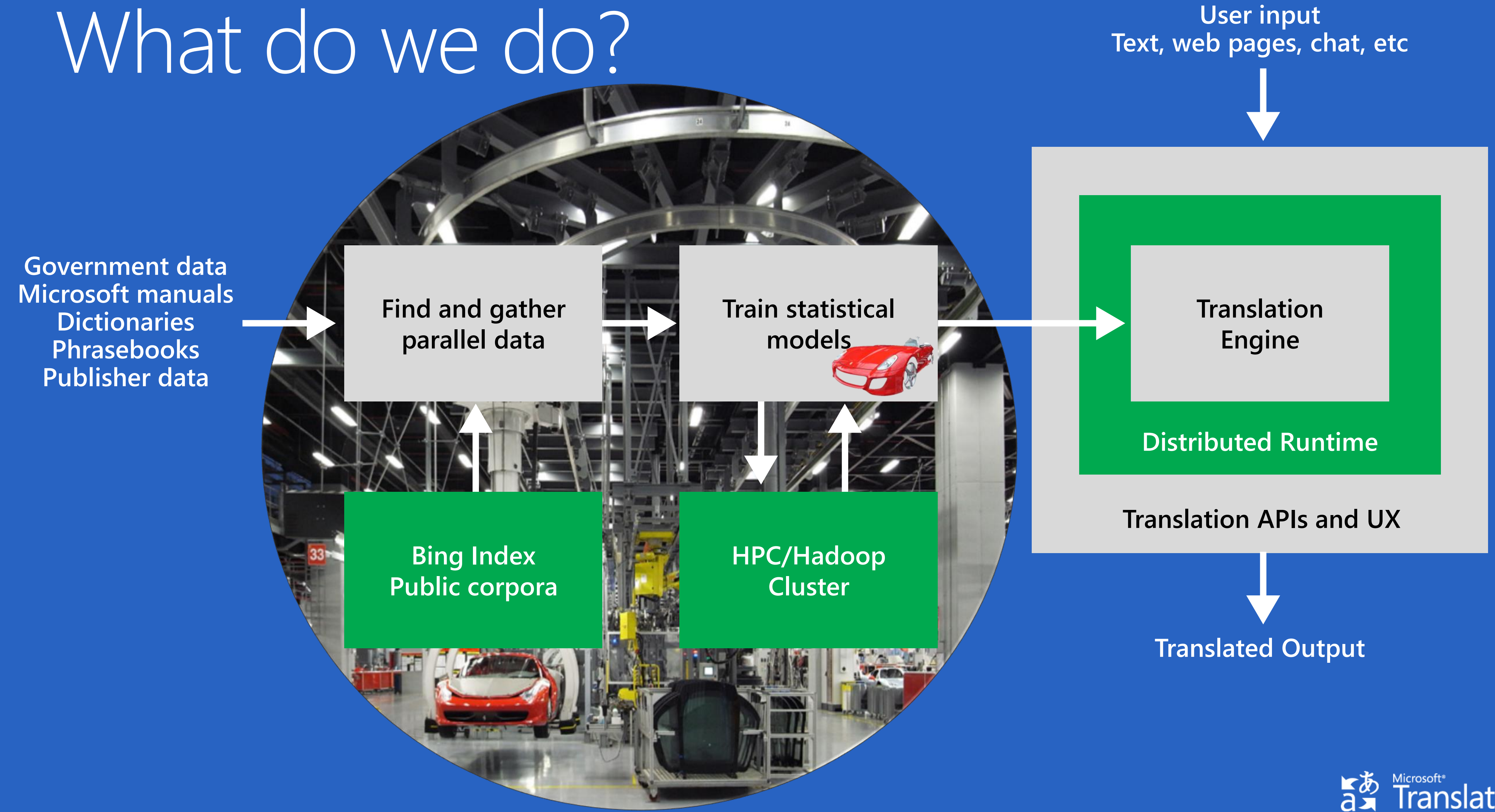
idk how did he reach b4 me

Outline

VIETATO FUMARE

33

What do we do?



User input
Text, web pages, chat, etc

Government data
Microsoft manuals
Dictionaries
Phrasebooks
Publisher data

Find and gather parallel data

Train statistical models

Bing Index
Public corpora

HPC/Hadoop Cluster

Translation Engine

Distributed Runtime

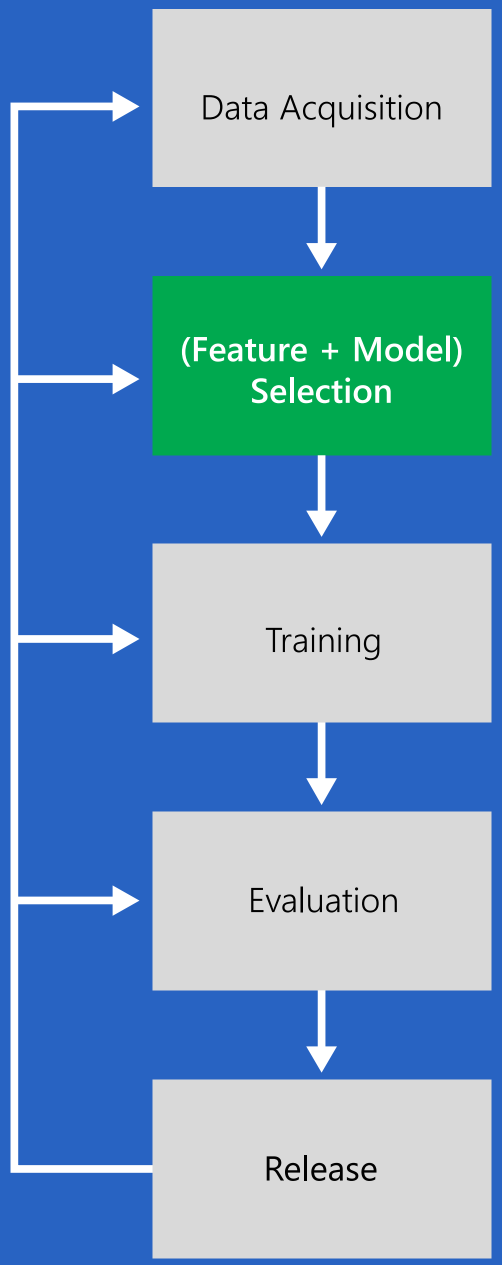
Translation APIs and UX

Translated Output

Microsoft® Translator

Outline: For-each

- Common Problems
- Our Solution
- Examples\Demo
- Summarize



Data Acquisition

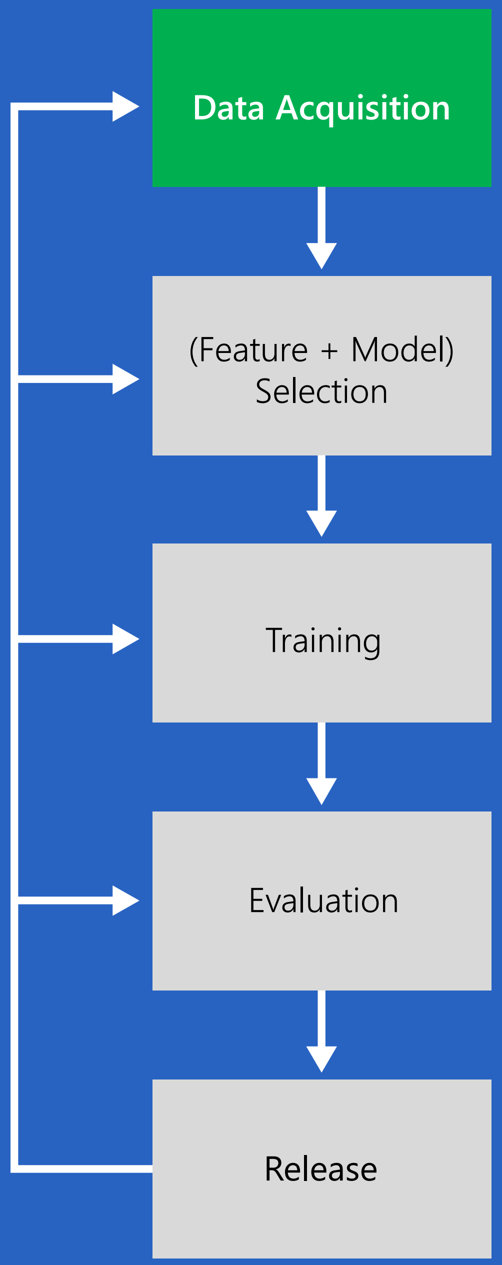
(Feature + Model) Selection

Training

Evaluation

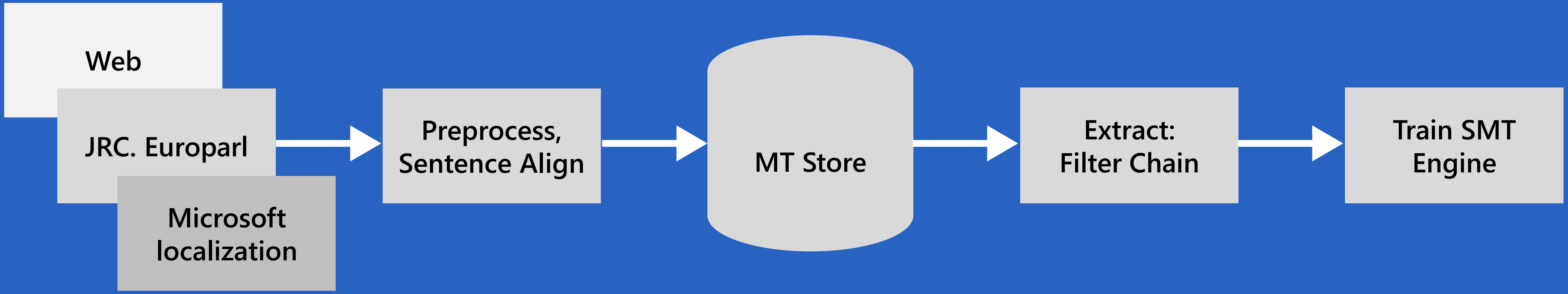
Release

Data Acquisition



Microsoft Translator

Data Pipeline



Web

JRC. Europarl

Microsoft localization

Preprocess, Sentence Align

MT Store

Extract: Filter Chain

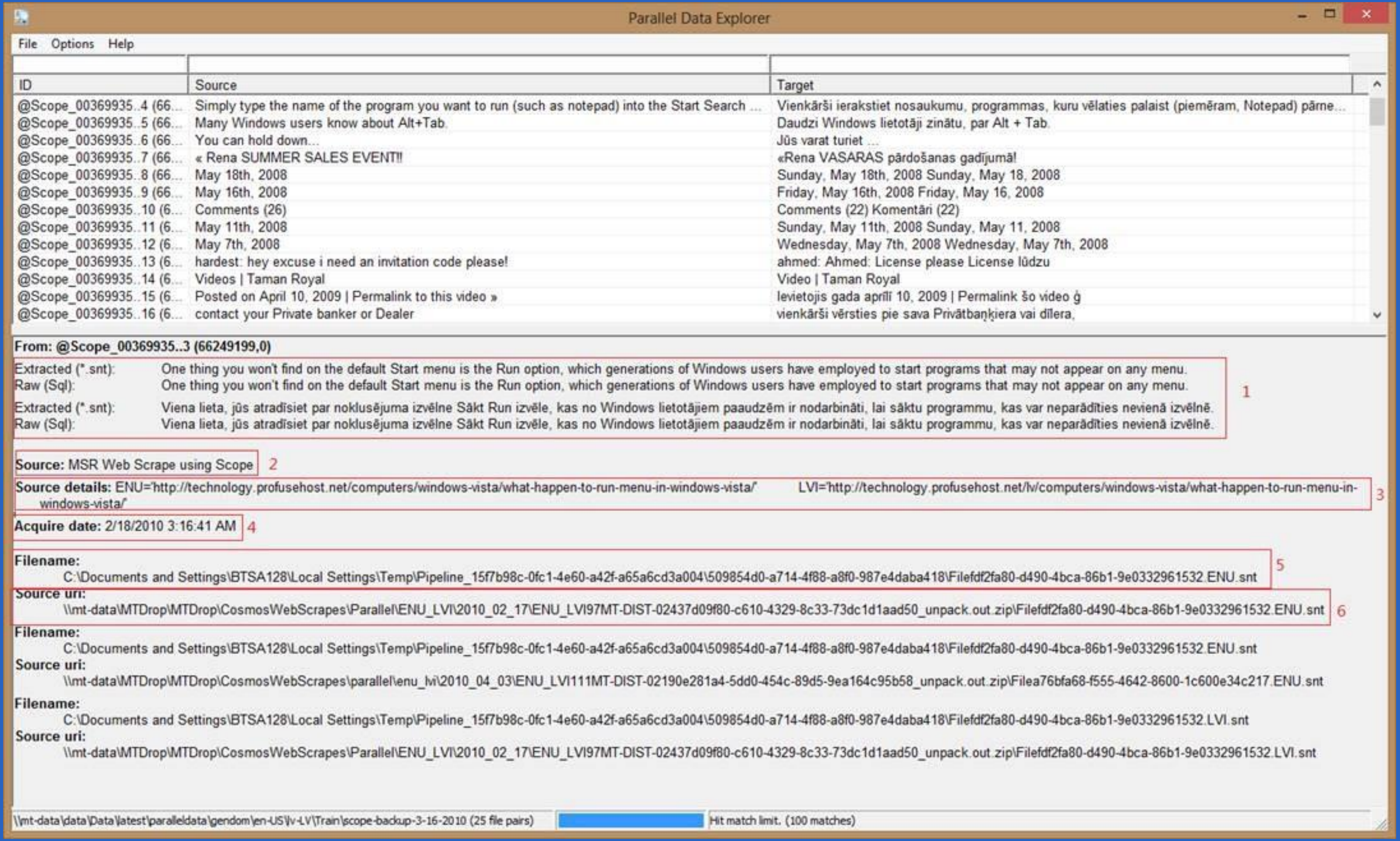
Train SMT Engine

Microsoft® Translator

Pre & Post processing data

```
UnicodeNormalization_Filter
Length_Filter
EscapeSequenceFilter
SentenceBreak_Filter
LanguageDetect_Filter
HtmlParser_Filter
LatinScriptRatio_Filter
WordCount_Filter
Dedupe_Filter
VocabularySaturation_Filter
```

Data Explorer



ID	Source	Target
@Scope_00369935..4 (66...	Simply type the name of the program you want to run (such as notepad) into the Start Search ...	Vienkārši ierakstiet nosaukumu, programmas, kuru vēlaties palaist (piemēram, Notepad) pārne...
@Scope_00369935..5 (66...	Many Windows users know about Alt+Tab.	Daudzi Windows lietotāji zinātu, par Alt + Tab.
@Scope_00369935..6 (66...	You can hold down...	Jūs varat turiet ...
@Scope_00369935..7 (66...	« Rena SUMMER SALES EVENT!!	«Rena VASARAS pārdošanas gadījumā!
@Scope_00369935..8 (66...	May 18th, 2008	Sunday, May 18th, 2008 Sunday, May 18, 2008
@Scope_00369935..9 (66...	May 16th, 2008	Friday, May 16th, 2008 Friday, May 16, 2008
@Scope_00369935..10 (6...	Comments (26)	Comments (22) Komentāri (22)
@Scope_00369935..11 (6...	May 11th, 2008	Sunday, May 11th, 2008 Sunday, May 11, 2008
@Scope_00369935..12 (6...	May 7th, 2008	Wednesday, May 7th, 2008 Wednesday, May 7th, 2008
@Scope_00369935..13 (6...	hardest: hey excuse i need an invitation code please!	ahmed: Ahmed: License please License lūdzu
@Scope_00369935..14 (6...	Videos | Taman Royal	Video | Taman Royal
@Scope_00369935..15 (6...	Posted on April 10, 2009 | Permalink to this video »	Ievietojis gada aprīlī 10, 2009 | Permalink šo video ģ
@Scope_00369935..16 (6...	contact your Private banker or Dealer	vienkārši vērsties pie sava Privātbaņķiera vai dīlera,

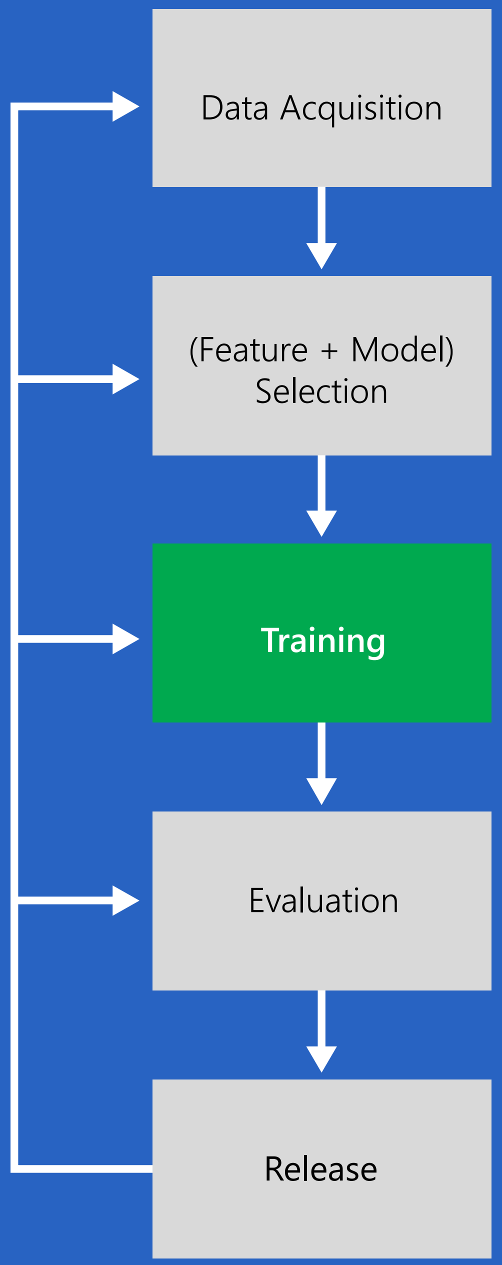
```
From: @Scope_00369935..3 (66249199,0)
Extracted (*.snt):    One thing you won't find on the default Start menu is the Run option, which generations of Windows users have employed to start programs that may not appear on any menu.
Raw (Sql):            One thing you won't find on the default Start menu is the Run option, which generations of Windows users have employed to start programs that may not appear on any menu.
Extracted (*.snt):    Viena lieta, jūs atradīsiet par noklusējuma izvēlne Sākt Run izvēle, kas no Windows lietotājiem paaudzēm ir nodarbināti, lai sāktu programmu, kas var neparādīties nevienā izvēlnē.
Raw (Sql):            Viena lieta, jūs atradīsiet par noklusējuma izvēlne Sākt Run izvēle, kas no Windows lietotājiem paaudzēm ir nodarbināti, lai sāktu programmu, kas var neparādīties nevienā izvēlnē.

Source: MSR Web Scrape using Scope
Source details: ENU='http://technology.profusehost.net/computers/windows-vista/what-happen-to-run-menu-in-windows-vista/'        LVI='http://technology.profusehost.net/lv/computers/windows-vista/what-happen-to-run-menu-in-
    windows-vista/'
Acquire date: 2/18/2010 3:16:41 AM

Filename:
        C:\Documents and Settings\BTSA128\Local Settings\Temp\Pipeline_15f7b98c-0fc1-4e60-a42f-a65a6cd3a004\509854d0-a714-4f88-a8f0-987e4daba418\Filefdf2fa80-d490-4bca-86b1-9e0332961532.ENU.snt
Source uri:
        \\mt-data\MTDrop\MTDrop\CosmosWebScrapes\Parallel\ENU_LVI\2010_02_17\ENU_LVI97MT-DIST-02437d09f80-c610-4329-8c33-73dc1d1aad50_unpack.out.zip\Filefdf2fa80-d490-4bca-86b1-9e0332961532.ENU.snt
Filename:
        C:\Documents and Settings\BTSA128\Local Settings\Temp\Pipeline_15f7b98c-0fc1-4e60-a42f-a65a6cd3a004\509854d0-a714-4f88-a8f0-987e4daba418\Filefdf2fa80-d490-4bca-86b1-9e0332961532.ENU.snt
Source uri:
        \\mt-data\MTDrop\MTDrop\CosmosWebScrapes\parallel\enu_lvi\2010_04_03\ENU_LVI111MT-DIST-02190e281a4-5dd0-454c-89d5-9ea164c95b58_unpack.out.zip\Filea76bfa68-f555-4642-8600-1c600e34c217.ENU.snt
Filename:
        C:\Documents and Settings\BTSA128\Local Settings\Temp\Pipeline_15f7b98c-0fc1-4e60-a42f-a65a6cd3a004\509854d0-a714-4f88-a8f0-987e4daba418\Filefdf2fa80-d490-4bca-86b1-9e0332961532.LVI.snt
Source uri:
        \\mt-data\MTDrop\MTDrop\CosmosWebScrapes\Parallel\ENU_LVI\2010_02_17\ENU_LVI97MT-DIST-02437d09f80-c610-4329-8c33-73dc1d1aad50_unpack.out.zip\Filefdf2fa80-d490-4bca-86b1-9e0332961532.LVI.snt

\\mt-data\data\Data\latest\paralleldata\gendom\en-US\lv-LV\Train\scope-backup-3-16-2010 (25 file pairs)    Hit match limit. (100 matches)
```

Data: Solutions

Summary

- Order from chaos – store data & metadata
- Automate thy steps
- Invest heavily in data pre-processing
  - Aggregate data to reduce training time without sacrificing quality
  - Junk data removal
- Once again, great data is the key! Both quantity and quality matter!

Training


Data Acquisition

(Feature + Model) Selection

Training

Evaluation

Release

Microsoft® Translator

Training: Common Issues

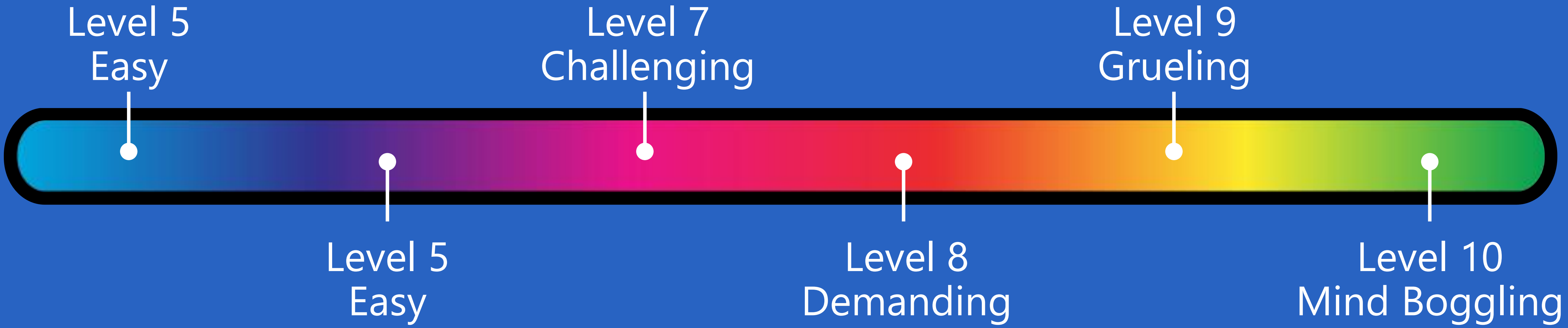
Easy, Quick?

Reproducible?

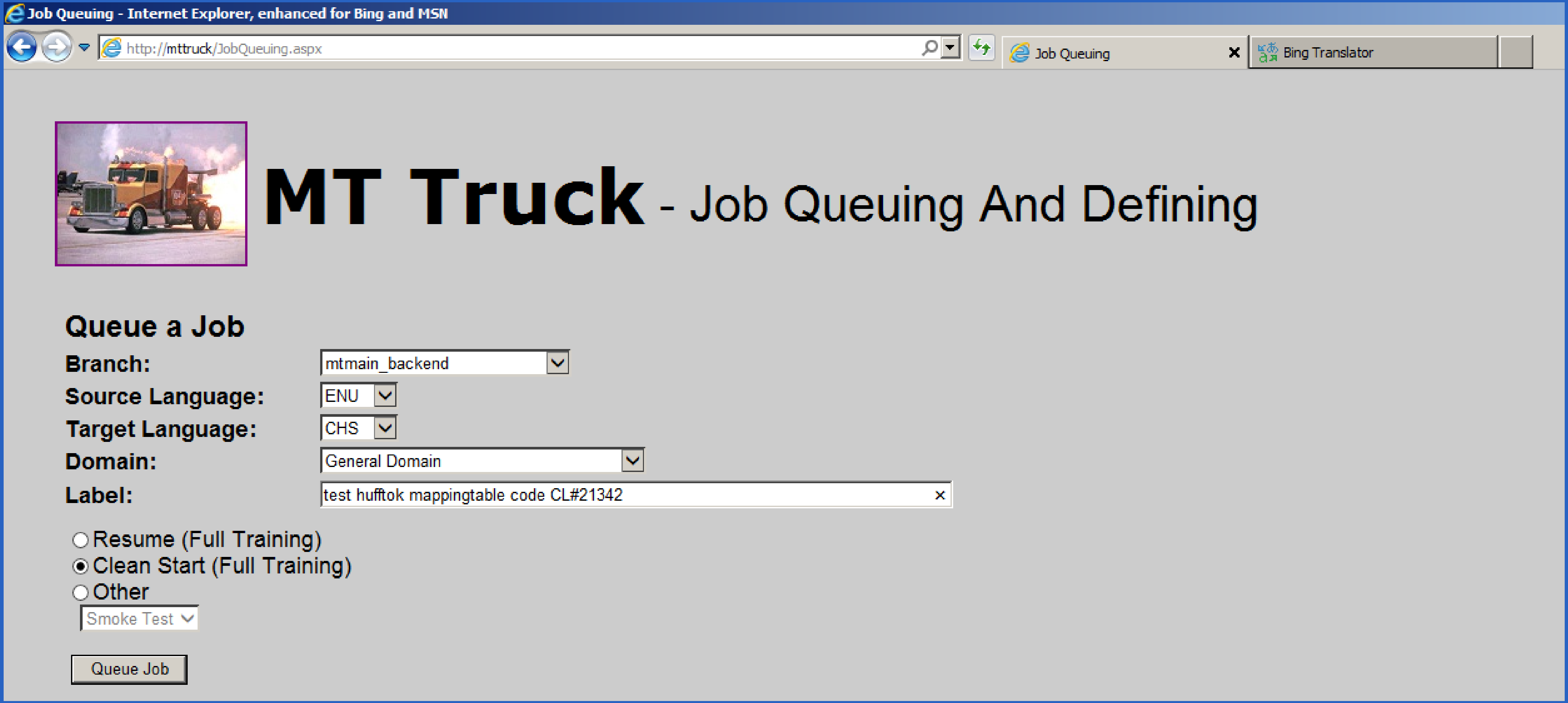
Stale?

Scaling?

Scaling out vs. Scaling up



Push button training



Job Queuing - Internet Explorer, enhanced for Bing and MSN
http://mttruck/JobQueuing.aspx
Job Queuing
Bing Translator

MT Truck - Job Queuing And Defining

Queue a Job

Branch:	mtmain_backend
Source Language:	ENU
Target Language:	CHS
Domain:	General Domain
Label:	test hufftok mappingtable code CL#21342

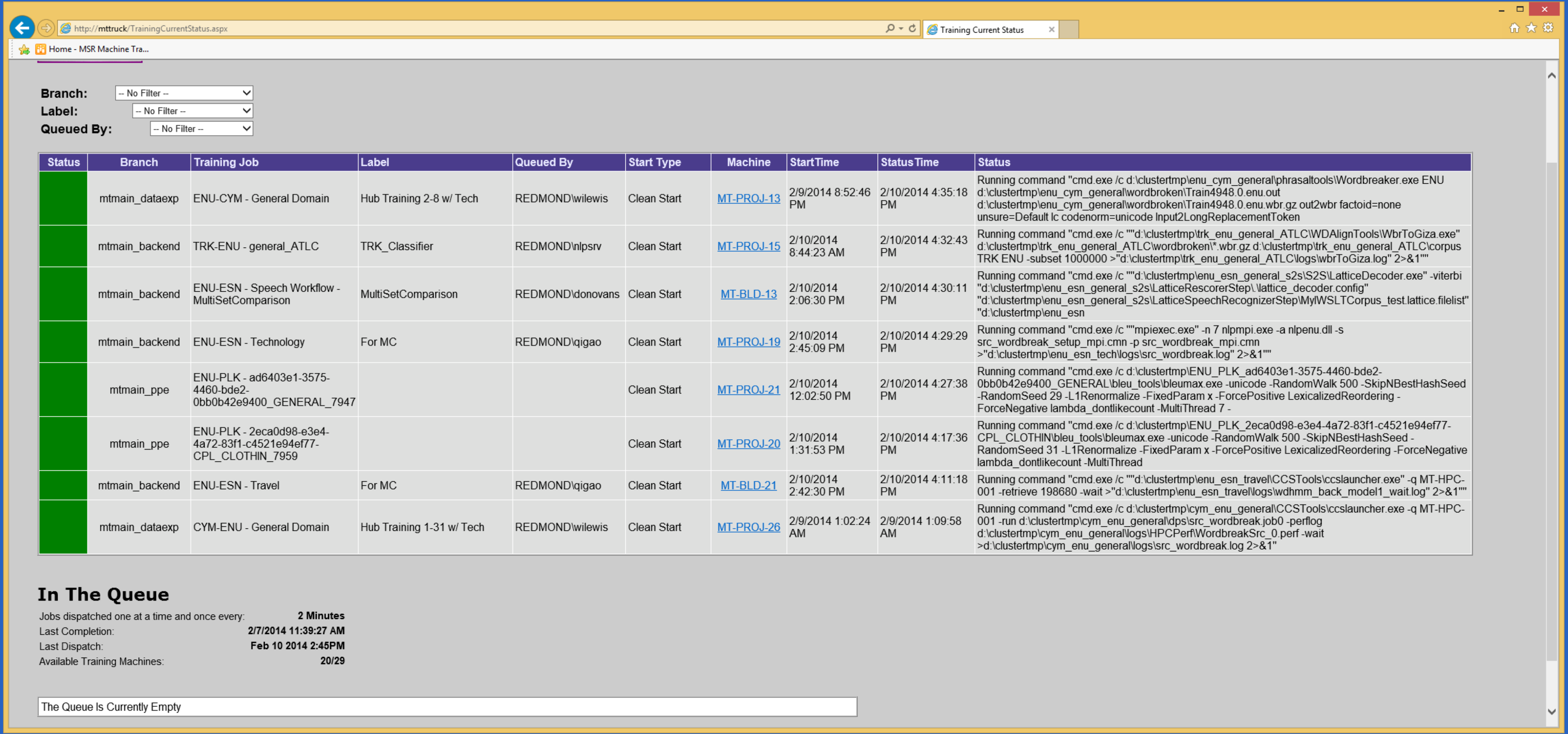
○ Resume (Full Training)
● Clean Start (Full Training)
○ Other

Smoke Test

Queue Job

Microsoft® Translator

Monitoring


http://mttruck/TrainingCurrentStatus.aspx
Training Current Status
Home - MSR Machine Tra...

Branch: -- No Filter --
Label: -- No Filter --
Queued By: -- No Filter --

Status	Branch	Training Job	Label	Queued By	Start Type	Machine	StartTime	StatusTime	Status
🟩	mtmain_dataexp	ENU-CYM - General Domain	Hub Training 2-8 w/ Tech	REDMOND\wilewis	Clean Start	MT-PROJ-13	2/9/2014 8:52:46 PM	2/10/2014 4:35:18 PM	Running command "cmd.exe /c d:\clustertmp\enu_cym_general\phrasaltools\Wordbreaker.exe ENU d:\clustertmp\enu_cym_general\wordbroken\Train4948.0.enu.out d:\clustertmp\enu_cym_general\wordbroken\Train4948.0.enu.wbr.gz out2wbr factoid=none unsure=Default lc codenorm=unicode Input2LongReplacementToken
🟩	mtmain_backend	TRK-ENU - general_ATLC	TRK_Classifier	REDMOND\nlpsrv	Clean Start	MT-PROJ-15	2/10/2014 8:44:23 AM	2/10/2014 4:32:43 PM	Running command "cmd.exe /c ""d:\clustertmp\trk_enu_general_ATLC\WDAlignTools\WbrToGiza.exe" d:\clustertmp\trk_enu_general_ATLC\wordbroken\*.wbr.gz d:\clustertmp\trk_enu_general_ATLC\corpus TRK ENU -subset 1000000 >"d:\clustertmp\trk_enu_general_ATLC\logs\wbrToGiza.log" 2>&1""
🟩	mtmain_backend	ENU-ESN - Speech Workflow - MultiSetComparison	MultiSetComparison	REDMOND\donovans	Clean Start	MT-BLD-13	2/10/2014 2:06:30 PM	2/10/2014 4:30:11 PM	Running command "cmd.exe /c ""d:\clustertmp\enu_esn_general_s2s\S2S\LatticeDecoder.exe" -viterbi "d:\clustertmp\enu_esn_general_s2s\LatticeRescorerStep\.\lattice_decoder.config" "d:\clustertmp\enu_esn_general_s2s\LatticeSpeechRecognizerStep\MyIWSLTCorpus_test.lattice.filelist" "d:\clustertmp\enu_esn
🟩	mtmain_backend	ENU-ESN - Technology	For MC	REDMOND\qigao	Clean Start	MT-PROJ-19	2/10/2014 2:45:09 PM	2/10/2014 4:29:29 PM	Running command "cmd.exe /c ""mpiexec.exe" -n 7 nlpmpi.exe -a nlpenu.dll -s src_wordbreak_setup_mpi.cmn -p src_wordbreak_mpi.cmn >"d:\clustertmp\enu_esn_tech\logs\src_wordbreak.log" 2>&1""
🟩	mtmain_ppe	ENU-PLK - ad6403e1-3575-4460-bde2-0bb0b42e9400_GENERAL_7947			Clean Start	MT-PROJ-21	2/10/2014 12:02:50 PM	2/10/2014 4:27:38 PM	Running command "cmd.exe /c d:\clustertmp\ENU_PLK_ad6403e1-3575-4460-bde2-0bb0b42e9400_GENERAL\bleu_tools\bleumax.exe -unicode -RandomWalk 500 -SkipNBestHashSeed -RandomSeed 29 -L1Renormalize -FixedParam x -ForcePositive LexicalizedReordering -ForceNegative lambda_dontlikecount -MultiThread 7 -
🟩	mtmain_ppe	ENU-PLK - 2eca0d98-e3e4-4a72-83f1-c4521e94ef77-CPL_CLOTHIN_7959			Clean Start	MT-PROJ-20	2/10/2014 1:31:53 PM	2/10/2014 4:17:36 PM	Running command "cmd.exe /c d:\clustertmp\ENU_PLK_2eca0d98-e3e4-4a72-83f1-c4521e94ef77-CPL_CLOTHIN\bleu_tools\bleumax.exe -unicode -RandomWalk 500 -SkipNBestHashSeed -RandomSeed 31 -L1Renormalize -FixedParam x -ForcePositive LexicalizedReordering -ForceNegative lambda_dontlikecount -MultiThread
🟩	mtmain_backend	ENU-ESN - Travel	For MC	REDMOND\qigao	Clean Start	MT-BLD-21	2/10/2014 2:42:30 PM	2/10/2014 4:11:18 PM	Running command "cmd.exe /c ""d:\clustertmp\enu_esn_travel\CCSTools\ccslauncher.exe" -q MT-HPC-001 -retrieve 198680 -wait >"d:\clustertmp\enu_esn_travel\logs\wdhmm_back_model1_wait.log" 2>&1""
🟩	mtmain_dataexp	CYM-ENU - General Domain	Hub Training 1-31 w/ Tech	REDMOND\wilewis	Clean Start	MT-PROJ-26	2/9/2014 1:02:24 AM	2/9/2014 1:09:58 AM	Running command "cmd.exe /c d:\clustertmp\cym_enu_general\CCSTools\ccslauncher.exe -q MT-HPC-001 -run d:\clustertmp\cym_enu_general\dps\src_wordbreak.job0 -perflog d:\clustertmp\cym_enu_general\logs\HPCPerf\WordbreakSrc_0.perf -wait >d:\clustertmp\cym_enu_general\logs\src_wordbreak.log 2>&1"

In The Queue

Jobs dispatched one at a time and once every:	**2 Minutes**
Last Completion:	**2/7/2014 11:39:27 AM**
Last Dispatch:	**Feb 10 2014 2:45PM**
Available Training Machines:	**20/29**

The Queue Is Currently Empty

Microsoft® Translator

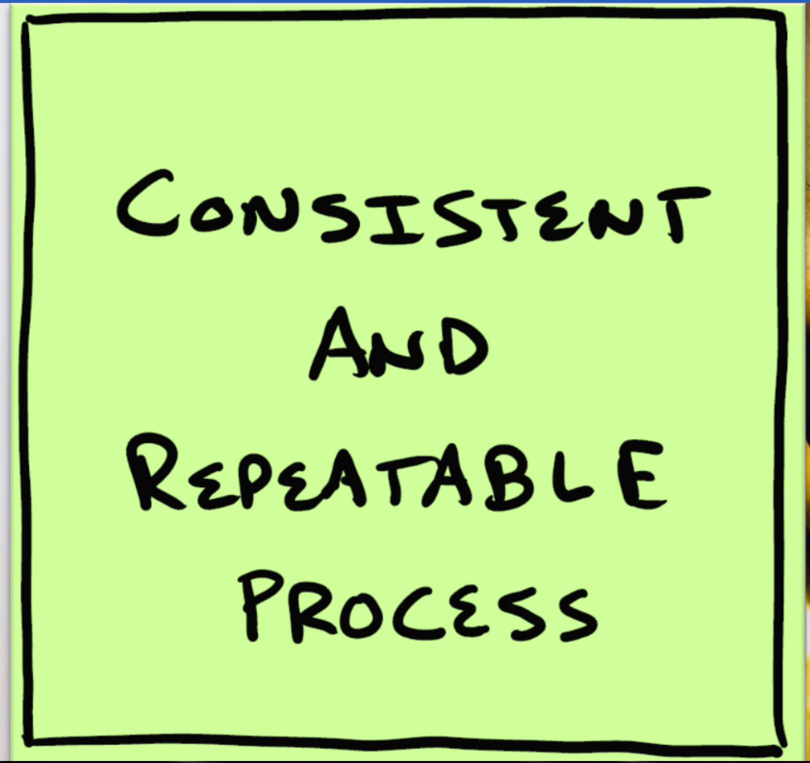
Sample Config

```
<ProjectConfig>
	<SourceLanguage>ENU</SourceLanguage>
	<TargetLanguage>HIN</TargetLanguage>
	<ParallelSentencePath>\\mt-data\data\datasink\mtmain_backend\paralleldata\general\en-us\hi-in\train</ParallelSentencePath>
	<LambdaParallelSentencePath>\\mt-data\data\datasink\mtmain_backend\paralleldata\general\en-us\hi-in\dev</LambdaTrainSourceSentences>
	<SmokeTestParallelSentencePath>\\mt-data\data\datasink\mtmain_backend\paralleldata\general\en-us\hi-in\test</LambdaTrainSourceSentences>
</ProjectConfig>
```

Expanded Config

```
<?xml version="1.0" encoding="utf-16"?>
<ProjectConfig xmlns:xsd="http://www.w3.org/2001/XMLSchema" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <DomainName>general</DomainName>
  <TrainMini>false</TrainMini>
  <SourceLanguage>ENU</SourceLanguage>
  <TargetLanguage>HIN</TargetLanguage>
  <EmailNotifyList>mttranno</EmailNotifyList>
  <ReferenceLengthSelection>Shortest</ReferenceLengthSelection>
  <SkipTranslationModelBuild>false</SkipTranslationModelBuild>
  <EuroparlBleu>false</EuroparlBleu>
  <DirectBleu>true</DirectBleu>
  <GenerateMobileModels>false</GenerateMobileModels>
  <MaxModelSize>50</MaxModelSize>
  <MobileModelsDropPath>\\mt-data\mobiletravel\mobiletravel\systems\latest_mobile</MobileModelsDropPath>
  <ParallelSentencePath>\\mt-data\data\datasink\mtmain_backend\paralleldata\general\en-us\hi-in\train</ParallelSentencePath>
  <ParallelSentencePathList>
    <ParallelSentencePath>\\mt-data\data\datasink\mtmain_backend\paralleldata\general\en-us\hi-in\train</ParallelSentencePath>
  </ParallelSentencePathList>
  <PreWordBrokenSrcDataPath />
  <PreWordBrokenTgtDataPath />
  <UseDataSelection>false</UseDataSelection>
  <GeneralData />
  <DomainData />
  <GenSelectData />
  <DomainDevData />
  <DataSelectionMethod>BiEntropyDiff</DataSelectionMethod>
  <DataSelectionRankMethod>DevDataThreshold</DataSelectionRankMethod>
  <DataSelectionLMOrder>4</DataSelectionLMOrder>
  <DataSelectionDiscount>0.7</DataSelectionDiscount>
  <SelectBestN>1000000</SelectBestN>
  <DataSelectionOutPath />
  <DataSelectionDryad />
  <DataSelectionThreshold>2</DataSelectionThreshold>
```

Training: Solutions



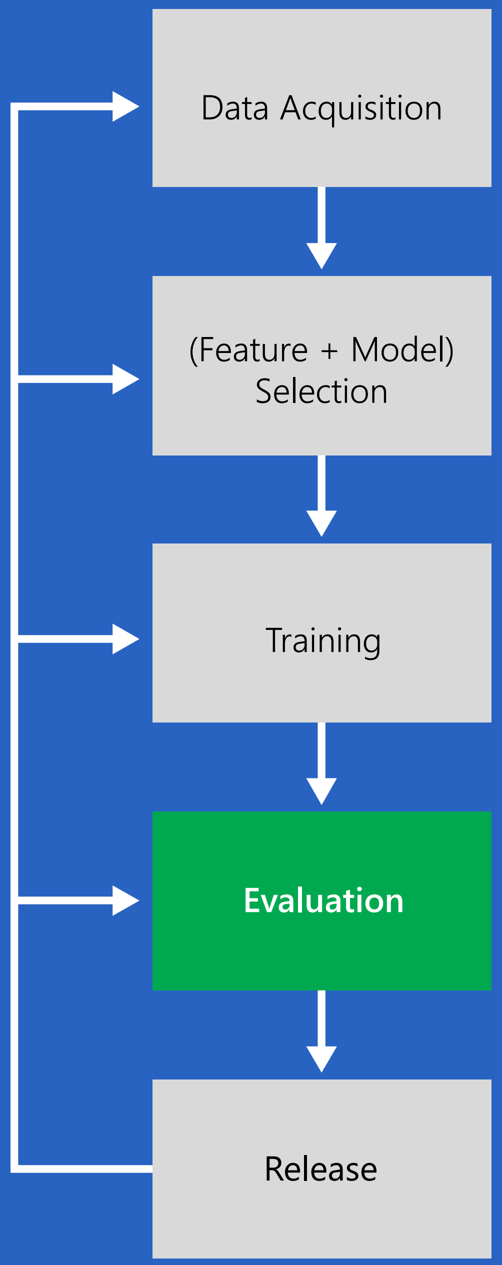
Consistent And Repeatable Process

Summary

- Speed: Distributed, multi-threaded etc.
  - As much effort into optimizing and scaling as runtime
  - Cache and re-use intermediate step results
- Reliable training
  - Every file, network access, etc
  - Auto resume – skip successful steps
- Simplify E2E training
  - Deterministic random seeds – data/code change?
  - Intelligent defaults for params, reduce human error
  - Continuously train, prevent stuff getting stale
  - Only E2E evals count for final shipping!

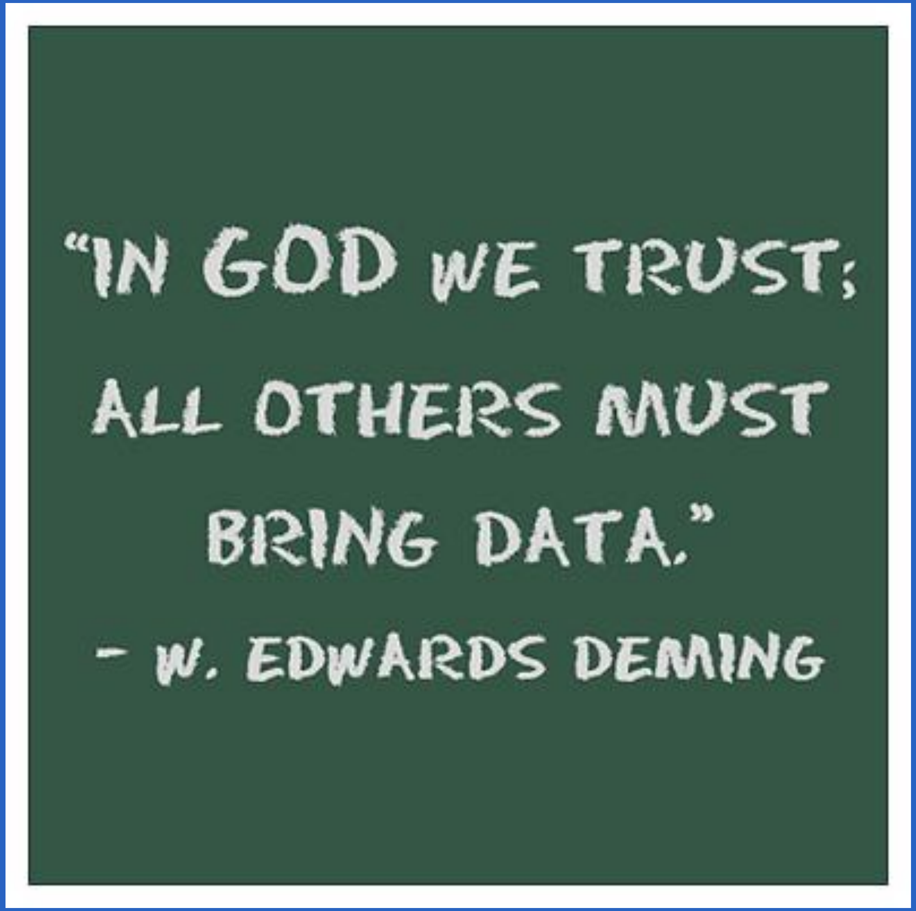
Microsoft® Translator

Debugging & Evaluation



Microsoft Translator

Evaluation: Common Issues

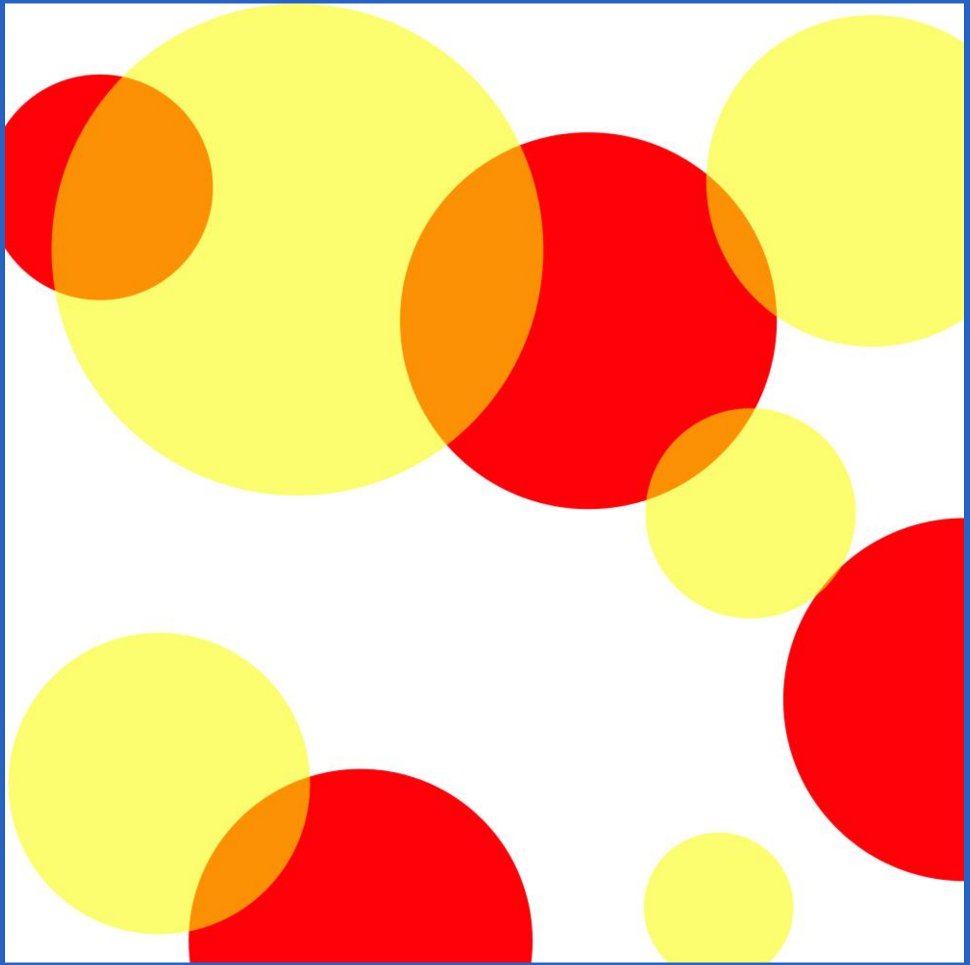


"IN GOD WE TRUST; ALL OTHERS MUST BRING DATA."

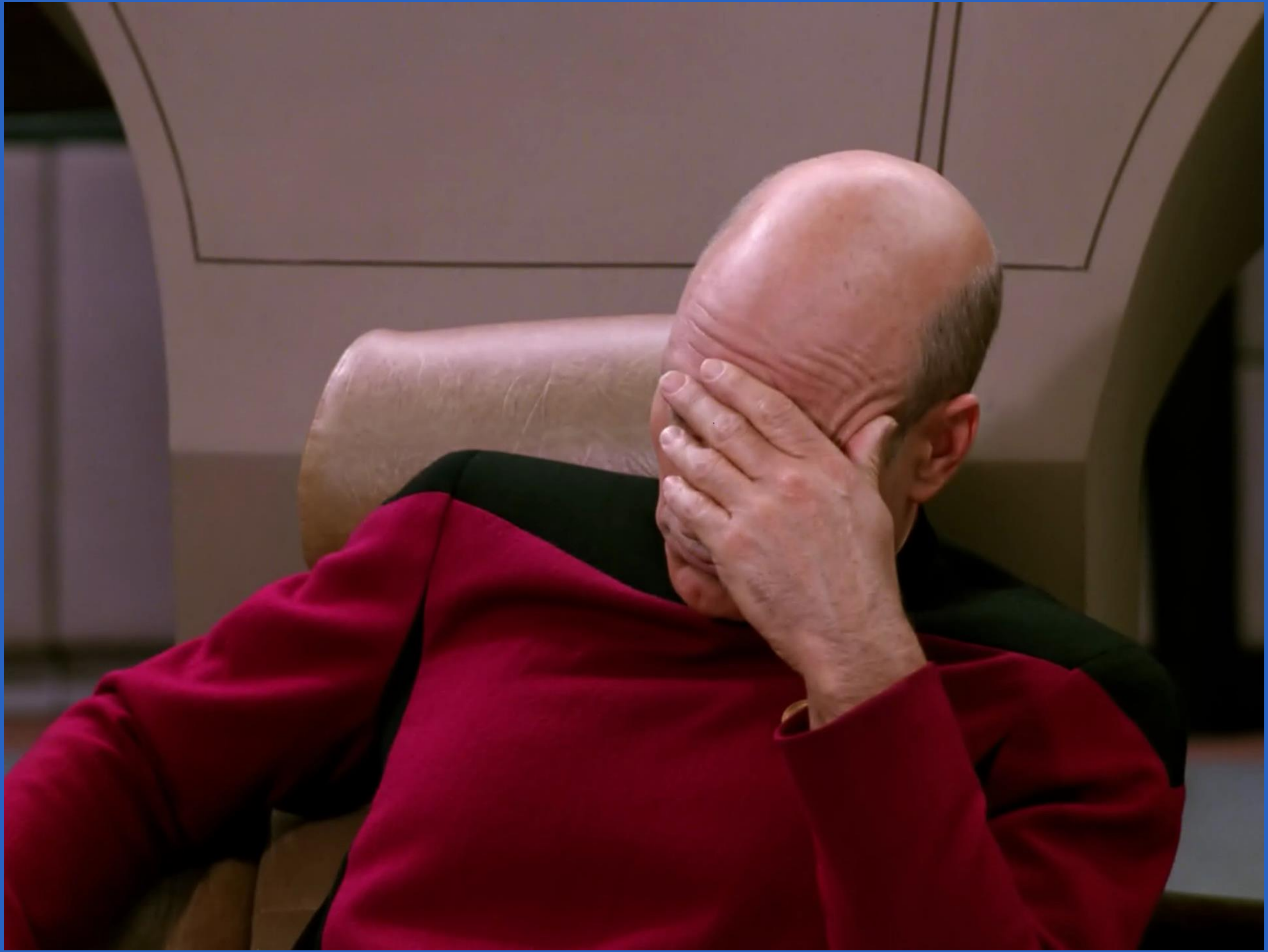
- W. EDWARDS DEMING

It works...It doesn't....It works.......
Just a hunch....maybe we do need a better way to measure results.....

Evaluation

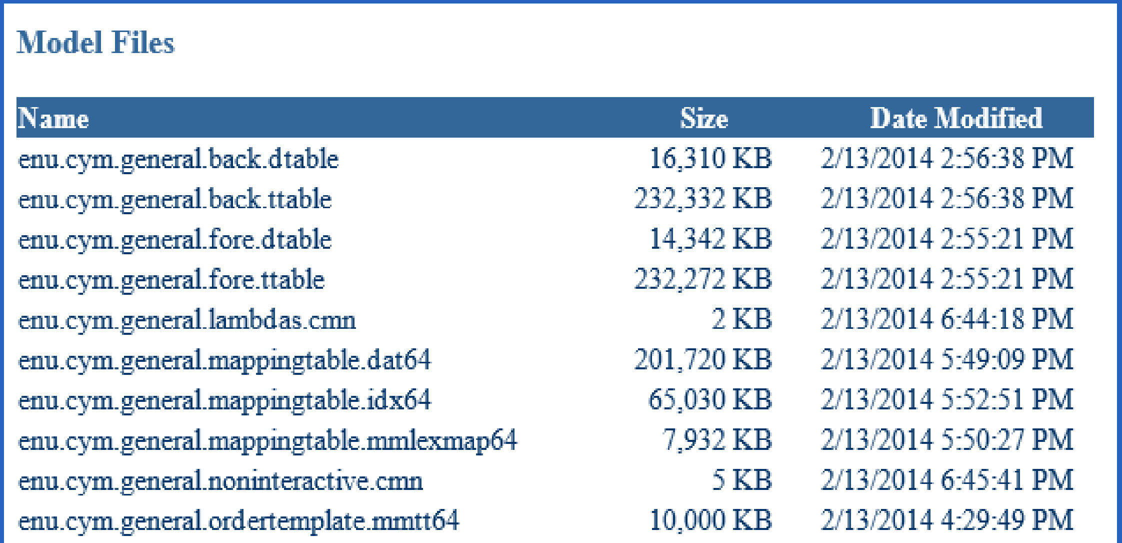


Debugging: Common Issues



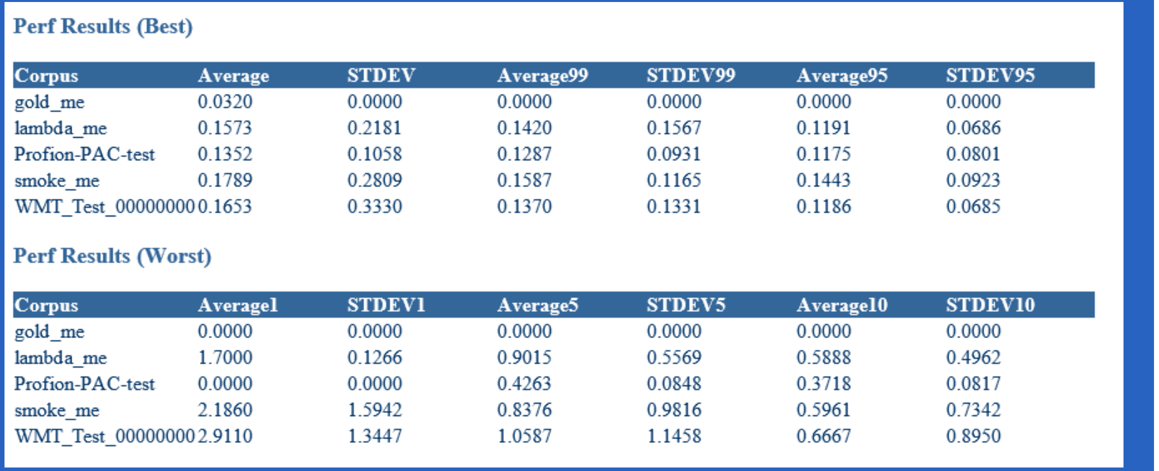
Microsoft Translator

Measure everything that matters



Model Files

Name	Size	Date Modified
enu.cym.general.back.dtable	16,310 KB	2/13/2014 2:56:38 PM
enu.cym.general.back.ttable	232,332 KB	2/13/2014 2:56:38 PM
enu.cym.general.fore.dtable	14,342 KB	2/13/2014 2:55:21 PM
enu.cym.general.fore.ttable	232,272 KB	2/13/2014 2:55:21 PM
enu.cym.general.lambdas.cmn	2 KB	2/13/2014 6:44:18 PM
enu.cym.general.mappingtable.dat64	201,720 KB	2/13/2014 5:49:09 PM
enu.cym.general.mappingtable.idx64	65,030 KB	2/13/2014 5:52:51 PM
enu.cym.general.mappingtable.mmlexmap64	7,932 KB	2/13/2014 5:50:27 PM
enu.cym.general.noninteractive.cmn	5 KB	2/13/2014 6:45:41 PM
enu.cym.general.ordertemplate.mmtt64	10,000 KB	2/13/2014 4:29:49 PM

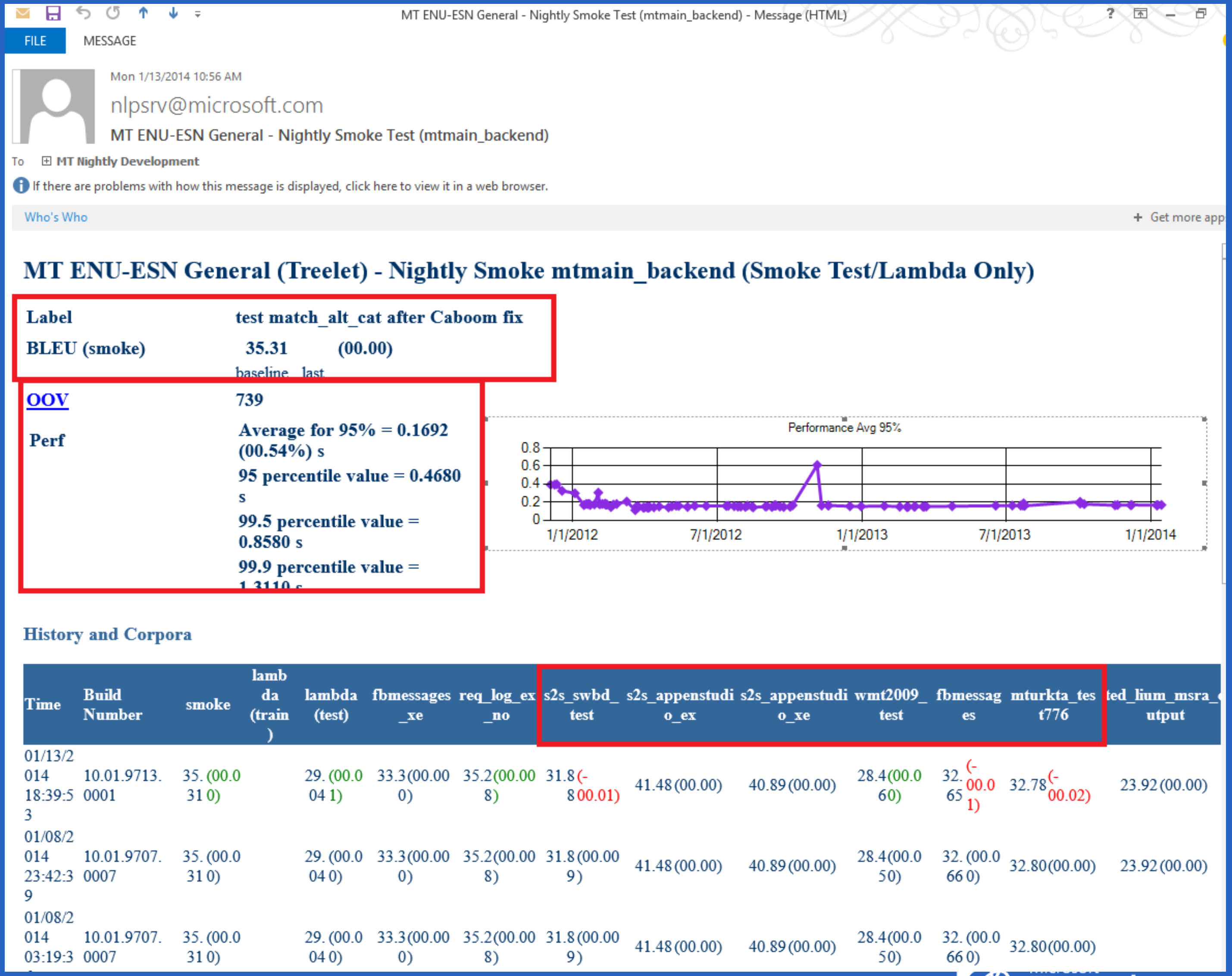


Perf Results (Best)

Corpus	Average	STDEV	Average99	STDEV99	Average95	STDEV95
gold_me	0.0320	0.0000	0.0000	0.0000	0.0000	0.0000
lambda_me	0.1573	0.2181	0.1420	0.1567	0.1191	0.0686
Profion-PAC-test	0.1352	0.1058	0.1287	0.0931	0.1175	0.0801
smoke_me	0.1789	0.2809	0.1587	0.1165	0.1443	0.0923
WMT_Test_00000000	0.1653	0.3330	0.1370	0.1331	0.1186	0.0685

Perf Results (Worst)

Corpus	Average1	STDEV1	Average5	STDEV5	Average10	STDEV10
gold_me	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
lambda_me	1.7000	0.1266	0.9015	0.5569	0.5888	0.4962
Profion-PAC-test	0.0000	0.0000	0.4263	0.0848	0.3718	0.0817
smoke_me	2.1860	1.5942	0.8376	0.9816	0.5961	0.7342
WMT_Test_00000000	2.9110	1.3447	1.0587	1.1458	0.6667	0.8950



Mon 1/13/2014 10:56 AM

nlpsrv@microsoft.com

MT ENU-ESN General - Nightly Smoke Test (mtmain_backend)

To MT Nightly Development

If there are problems with how this message is displayed, click here to view it in a web browser.

MT ENU-ESN General (Treelet) - Nightly Smoke mtmain_backend (Smoke Test/Lambda Only)

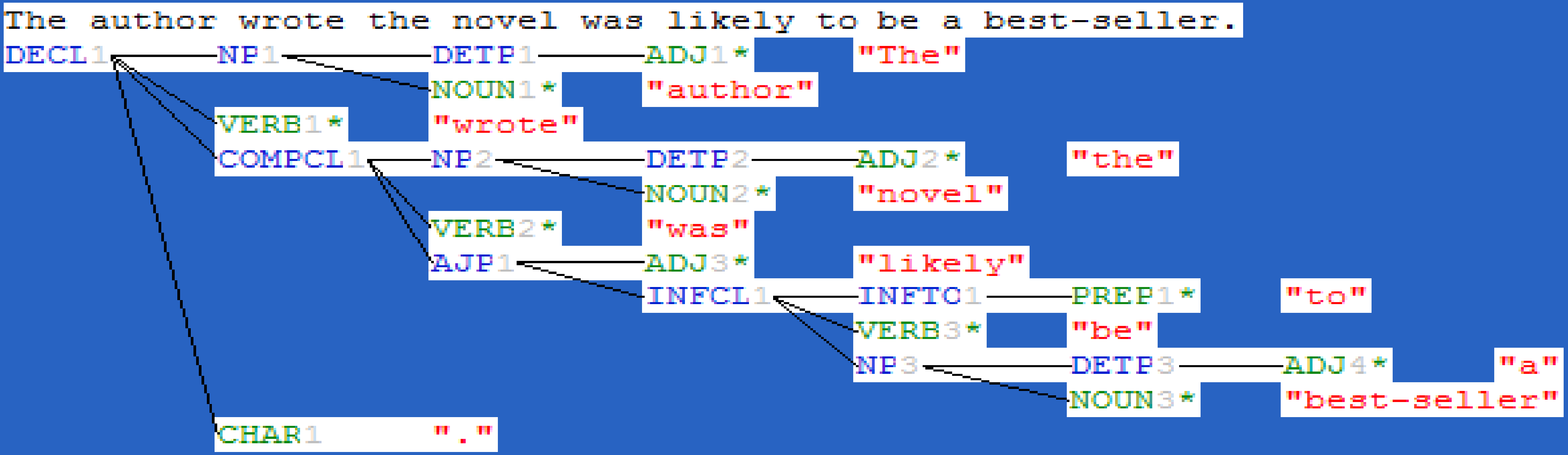
Label	test match_alt_cat after Caboom fix
BLEU (smoke)	35.31 (00.00) baseline last
OOV	739
Perf	Average for 95% = 0.1692 (00.54%) s 95 percentile value = 0.4680 s 99.5 percentile value = 0.8580 s 99.9 percentile value = 1.3110 s

History and Corpora

Time	Build Number	smoke	lambda (train)	lambda (test)	fbmessages_xe	req_log_ex_no	s2s_swbd_test	s2s_appenstudio_ex	s2s_appenstudio_xe	wmt2009_test	fbmessages	mturkta_test776	ted_lium_msra_utput
01/13/2014 18:39:53	10.01.9713.0001	35. (00.0 31 0)		29. (00.0 04 1)	33.3(00.00 0)	35.2(00.00 8)	31.8(- 8 00.01)	41.48 (00.00)	40.89 (00.00)	28.4(00.0 6 0)	32. (- 65 00.0 1)	32.78 (- 00.02)	23.92 (00.00)
01/08/2014 23:42:39	10.01.9707.0007	35. (00.0 31 0)		29. (00.0 04 0)	33.3(00.00 0)	35.2(00.00 8)	31.8 (00.00 9)	41.48 (00.00)	40.89 (00.00)	28.4(00.0 5 0)	32. (00.0 66 0)	32.80(00.00)	23.92 (00.00)
01/08/2014 03:19:3	10.01.9707.0007	35. (00.0 31 0)		29. (00.0 04 0)	33.3(00.00 0)	35.2(00.00 8)	31.8 (00.00 9)	41.48 (00.00)	40.89 (00.00)	28.4(00.0 5 0)	32. (00.0 66 0)	32.80(00.00)	

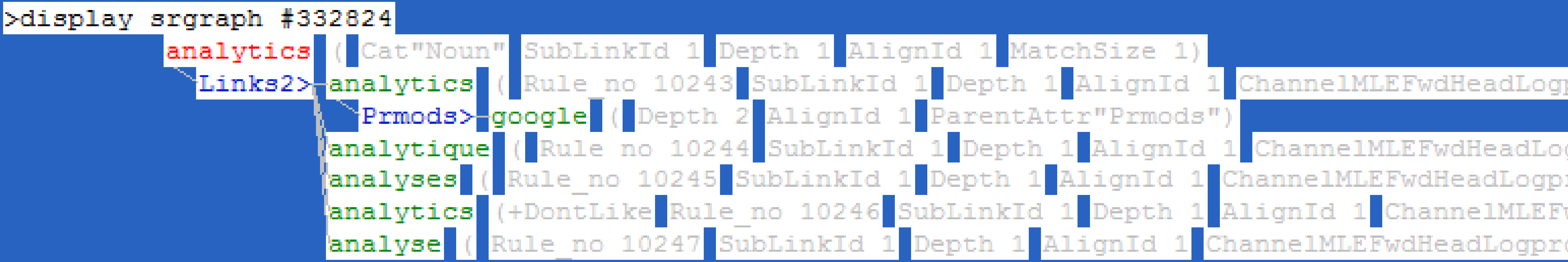
Translator

Debugging



```
The author wrote the novel was likely to be a best-seller.
DECL1     NP1       DETP1     ADJ1*               "The"
                    NOUN1*    "author"
          VERB1*    "wrote"
          COMPCL1   NP2       DETP2     ADJ2*          "the"
                              NOUN2*    "novel"
                    VERB2*    "was"
                    AJP1      ADJ3*     "likely"
                              INFCL1    INFTO1    PREP1*    "to"
                                        VERB3*    "be"
                                        NP3       DETP3     ADJ4*    "a"
                                                  NOUN3*    "best-seller"
          CHAR1     "."
```

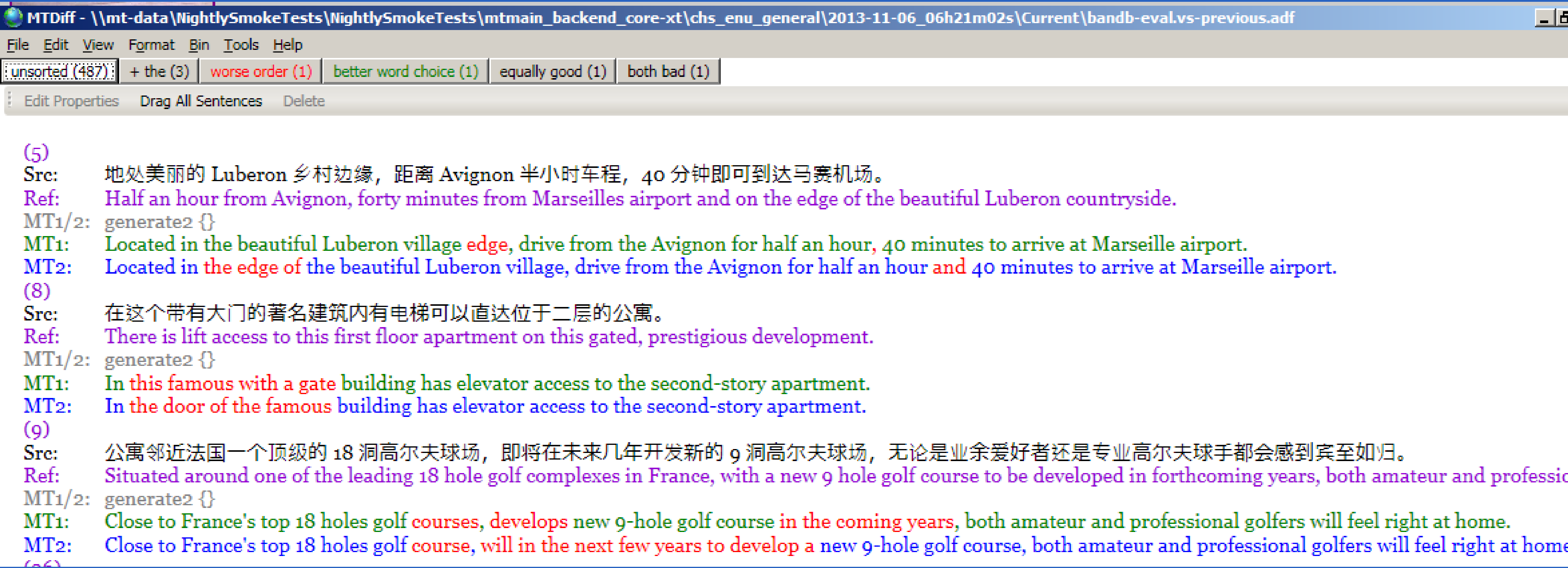
Debugging



```
>display srgraph #332824
          analytics  ( Cat"Noun" SubLinkId 1 Depth 1 AlignId 1 MatchSize 1)
            Links2>-analytics  ( Rule_no 10243 SubLinkId 1 Depth 1 AlignId 1 ChannelMLEFwdHeadLogp
                      Prmods>-google  ( Depth 2 AlignId 1 ParentAttr"Prmods")
                    analytique  ( Rule no 10244 SubLinkId 1 Depth 1 AlignId 1 ChannelMLEFwdHeadLo
                    analyses  ( Rule_no 10245 SubLinkId 1 Depth 1 AlignId 1 ChannelMLEFwdHeadLogp
                    analytics  (+DontLike Rule_no 10246 SubLinkId 1 Depth 1 AlignId 1 ChannelMLEF
                    analyse  ( Rule_no 10247 SubLinkId 1 Depth 1 AlignId 1 ChannelMLEFwdHeadLogpr
```

Microsoft® Translator

Debugging Output

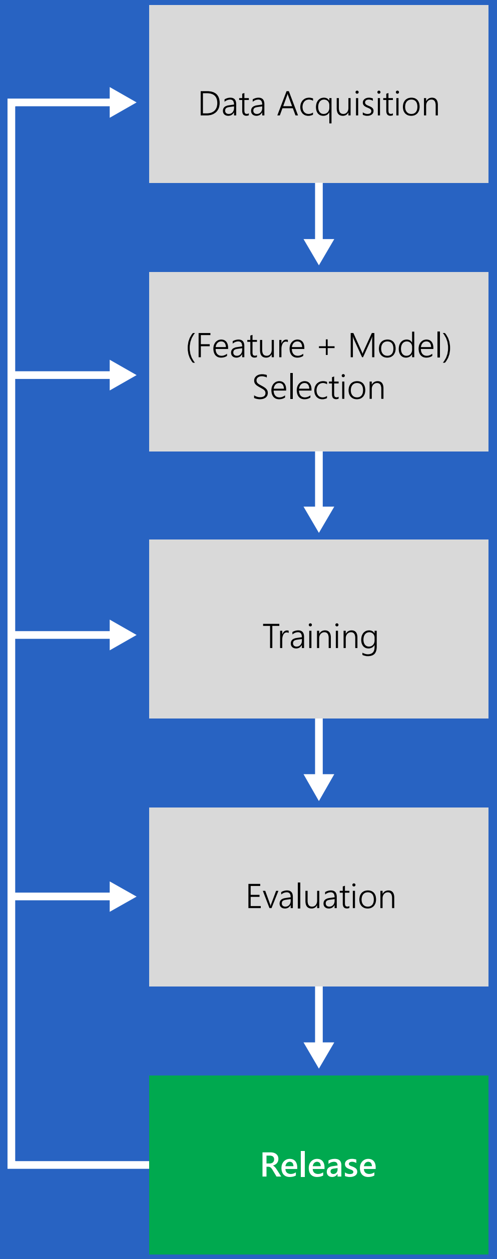


```
(5)
Src:     地处美丽的 Luberon 乡村边缘，距离 Avignon 半小时车程，40 分钟即可到达马赛机场。
Ref:     Half an hour from Avignon, forty minutes from Marseilles airport and on the edge of the beautiful Luberon countryside.
MT1/2:   generate2 {}
MT1:     Located in the beautiful Luberon village edge, drive from the Avignon for half an hour, 40 minutes to arrive at Marseille airport.
MT2:     Located in the edge of the beautiful Luberon village, drive from the Avignon for half an hour and 40 minutes to arrive at Marseille airport.
(8)
Src:     在这个带有大门的著名建筑内有电梯可以直达位于二层的公寓。
Ref:     There is lift access to this first floor apartment on this gated, prestigious development.
MT1/2:   generate2 {}
MT1:     In this famous with a gate building has elevator access to the second-story apartment.
MT2:     In the door of the famous building has elevator access to the second-story apartment.
(9)
Src:     公寓邻近法国一个顶级的 18 洞高尔夫球场，即将在未来几年开发新的 9 洞高尔夫球场，无论是业余爱好者还是专业高尔夫球手都会感到宾至如归。
Ref:     Situated around one of the leading 18 hole golf complexes in France, with a new 9 hole golf course to be developed in forthcoming years, both amateur and professi
MT1/2:   generate2 {}
MT1:     Close to France's top 18 holes golf courses, develops new 9-hole golf course in the coming years, both amateur and professional golfers will feel right at home.
MT2:     Close to France's top 18 holes golf course, will in the next few years to develop a new 9-hole golf course, both amateur and professional golfers will feel right at home
```

Summary

- Log everything with E2E tracing
- Measure everything that matters
  - Aggregate vs. deterministic testing
  - Multiple test sets, refresh test sets
  - Performance, early indication
  - Tests for individual models as well as E2E composition
- Debugging
  - UI representation of decoding path
  - Configs for turning off non-determinism

Release



Microsoft® Translator

Common Pitfalls

Microsoft® Translator

Summary

- Ship early and continuous measurement.
- Bridge expectation vs. Reality
- Models vs. code

vishalc@microsoft.com
@talktovishal

Microsoft

blogs.msdn.com/translator

twitter.com/MSTranslator

facebook.com/BingTranslator

Microsoft® Translator