Adapting taggers to Twitter using *not-so-distant* supervision

Barbara Plank, Dirk Hovy, Ryan McDonald and Anders Søgaard

Newswire bias

The CROSS-DOMAIN GULF

Newswire bias

"domain adaptation"
or "transfer learning"

The CROSS-DOMAIN GULF

Newswire bias

"domain adaptation"
or "transfer learning"

The CROSS-DOMAIN GULF

Idea

Tweet

NN	NN	JJR	.	VB	NN	.	NNP
#Localization	#job :	Supplier	/	Project	Manager	-	Localisation

NN	.	NNP	NNP	NNP	NNP	NN	NNP
Vendor	-	NY,	NY,	United	States	http://bit.ly/16KigBg	#nlppeople

Idea

Tweet

NN	NN	JJR	.	VB	NN	.	NNP
#Localization	#job :	Supplier	/	Project	Manager	-	Localisation
NN	.	NNP	NNP	NNP	NNP	NN	NNP
Vendor	-	NY,	NY,	United	States	http://bit.ly/16KigBg	#nlppeople

20%

Idea

Tweet

NN	NN	JJR	.	VB	NN	.	NNP
#Localization	#job :	Supplier	/	Project	Manager	-	Localisation
NN	.	NNP	NNP	NNP	NNP	NN	NNP
Vendor	-	NY,	NY,	United	States	http://bit.ly/16KigBg	#nlppeople

URL

NN	NN	.	NN	NN	VBZ	DET	...
The	Supplier	/	Project	Manager	performs	the	...

Idea

Tweet

NN	NN	NN	.	VB	NN	.	NNP
#Localization	#job :	Supplier	/	Project	Manager	-	Localisation
NN	.	NNP	NNP	NNP	NNP	NN	NNP
Vendor	-	NY,	NY,	United	States	http://bit.ly/16KigBg	#nlppeople

URL

NN	NN	.	NN	NN	VBZ	DET	...
The	Supplier	/	Project	Manager	performs	the	...

Idea

Tweet

NN	NN	NN	.	NN	NN	.	NNP
#Localization	#job :	Supplier	/	Project	Manager	-	Localisation
NN	.	NNP	NNP	NNP	NNP	NN	NNP
Vendor	-	NY,	NY,	United	States	http://bit.ly/16KigBg	#nlppeople

URL

NN	NN	.	NN	NN	VBZ	DET	...
The	Supplier	/	Project	Manager	performs	the	...

Same for NER

Tweet

O	O	O	O	B-PER	O	O	O
Prey	Developer	worked	with	Nintendo	on	project	http://bit.ly/17Kbsf

Same for NER

Tweet

O	O	O	O	B-PER	O	O	O
Prey	Developer	worked	with	Nintendo	on	project	http://bit.ly/17Kbsf

URL

O	O	O	O	B-ORG	O	O	...
In	a	statement	,	Nintendo	announced	that	...

Same for NER

Tweet

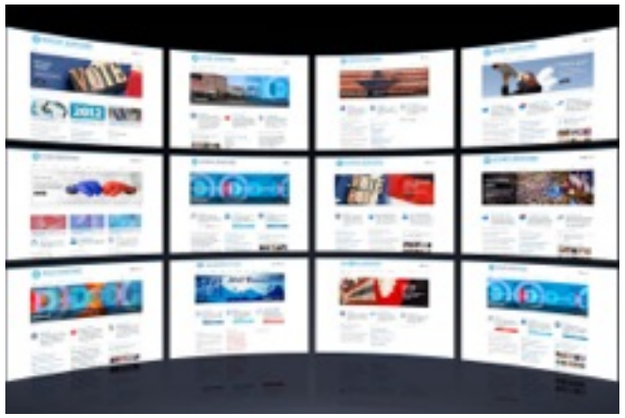
O	O	O	O	B-ORG	O	O	O
Prey	Developer	worked	with	Nintendo	on	project	http://bit.ly/17Kbsf

URL

O	O	O	O	B-ORG	O	O	...
In	a	statement	,	Nintendo	announced	that	...

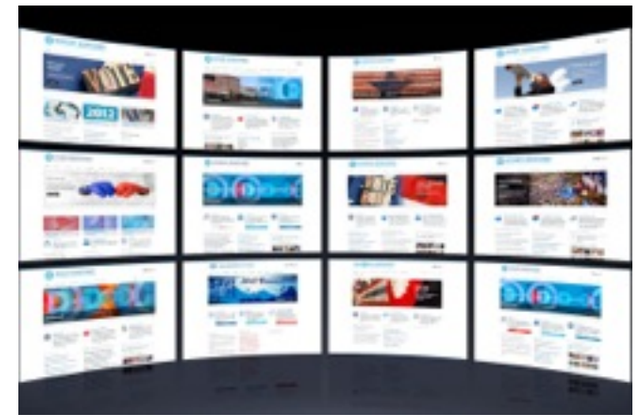
Setup


The New York Times
American Football Best Sport Ever!!!
MANY OBSERVANCES
Requiems and Flyovers, Moments of Silence, Doves on the Wing
BUSH ISSUES ORDER
Embassies Are Shut but Domestic Security Is Also Emphasized
Mental Health: The Profession Tests Its Limits
Resilience And Grief: 365 Days
PLAN TO ATTACK EMBASSIES (ITEM)
U.S. Says Qaeda Member Said of Threat to Offices in Asia
Its Ranks Depleted, A Weary Fire Dept. Is Trying to Regroup
Grieve Today, Mayor Says, And Then Grasp Tomorrow
Threat of Terrorism Is Shaping the Focus Of Bush Presidency



Setup

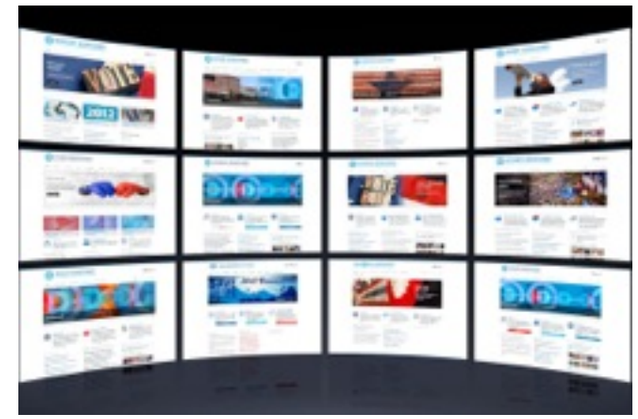

The New York Times
American Football Best Sport Ever!!!
MANY OBSERVANCES
Requiems and Flyovers, Moments of Silence, Doves on the Wing
BUSH ISSUES ORDER
Embassies Are Shut but Domestic Security Is Also Emphasized
Mental Health: The Profession Tests Its Limits
Resilience And Grief: 365 Days
PLAN TO ATTACK EMBASSIES CITED
Its Ranks Depleted, A Weary Fire Dept. Is Trying to Regroup
Grieve Today, Mayor Says, And Then Grasp Tomorrow
Threat of Terrorism Is Shaping the Focus Of Bush Presidency



tag

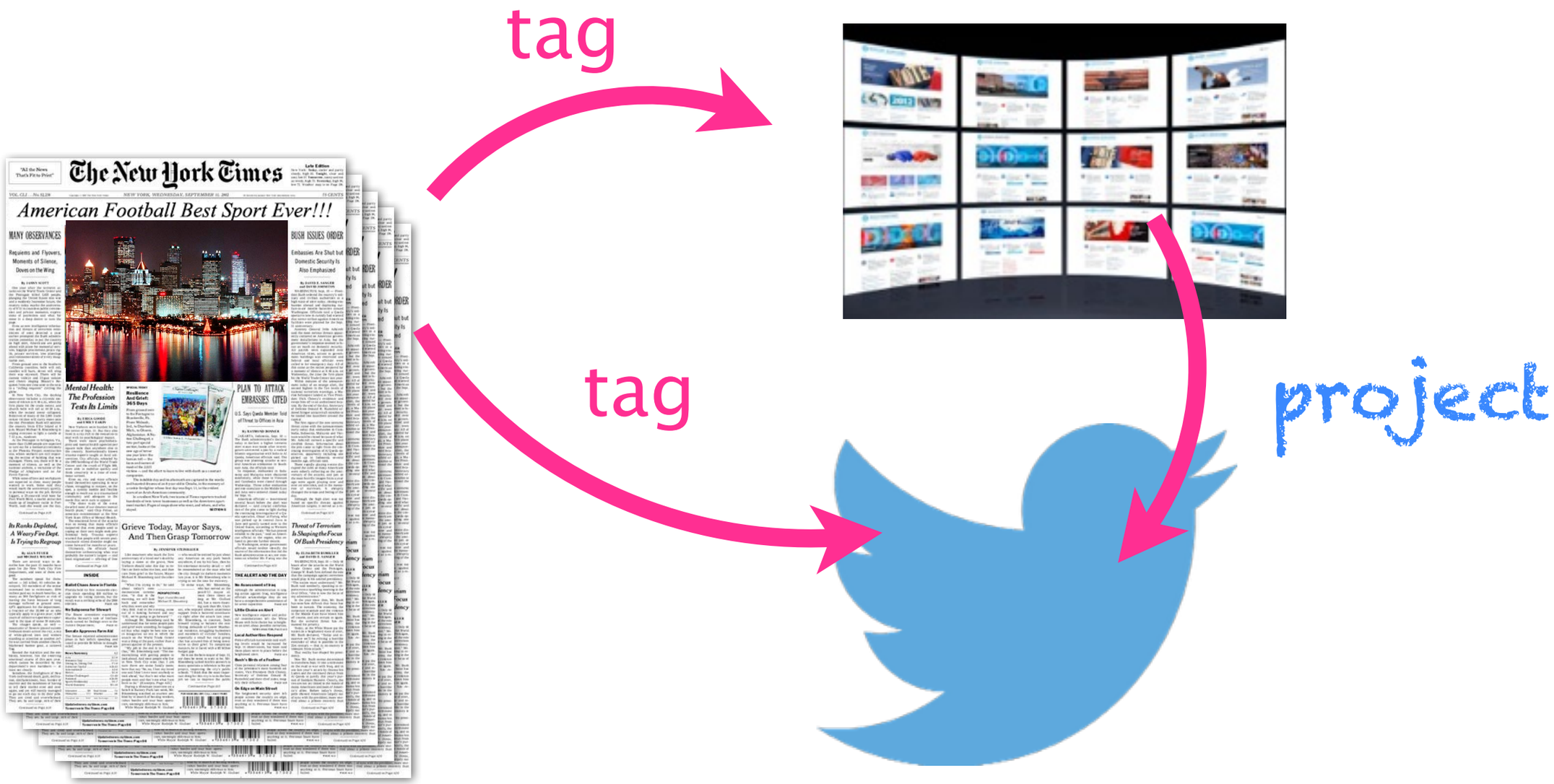
Setup

tag

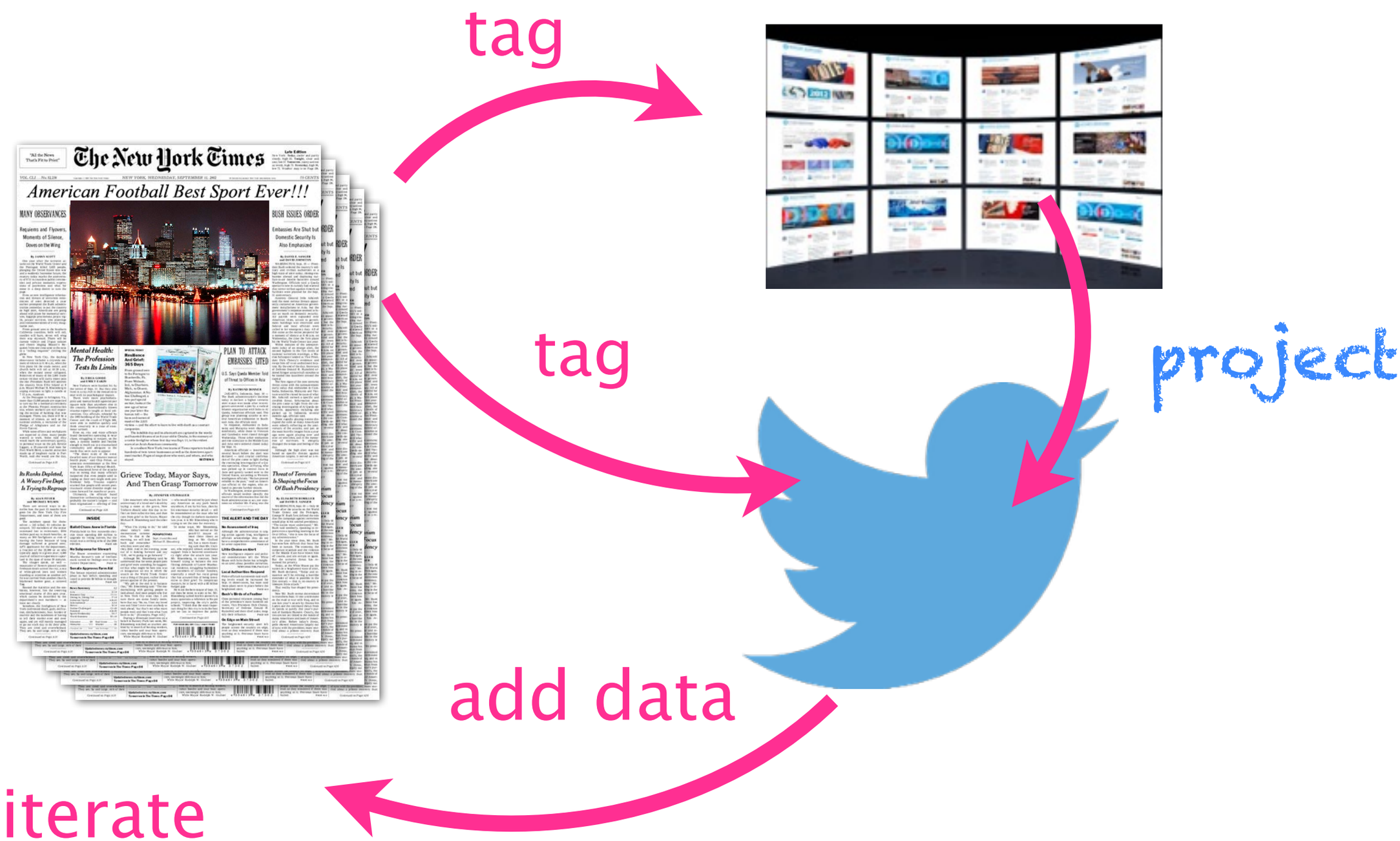


tag

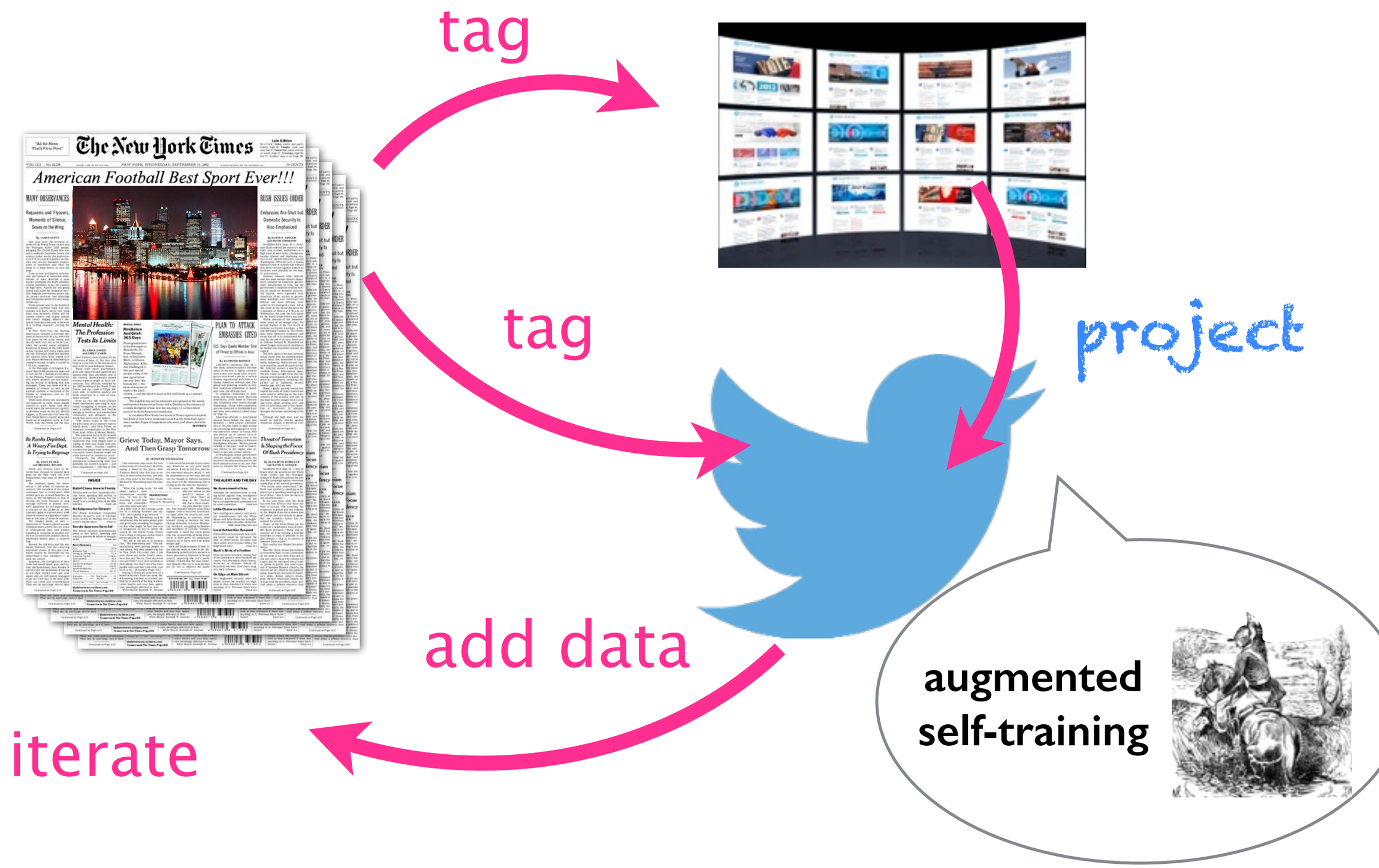
Setup


tag
tag
project
The New York Times
American Football Best Sport Ever!!!

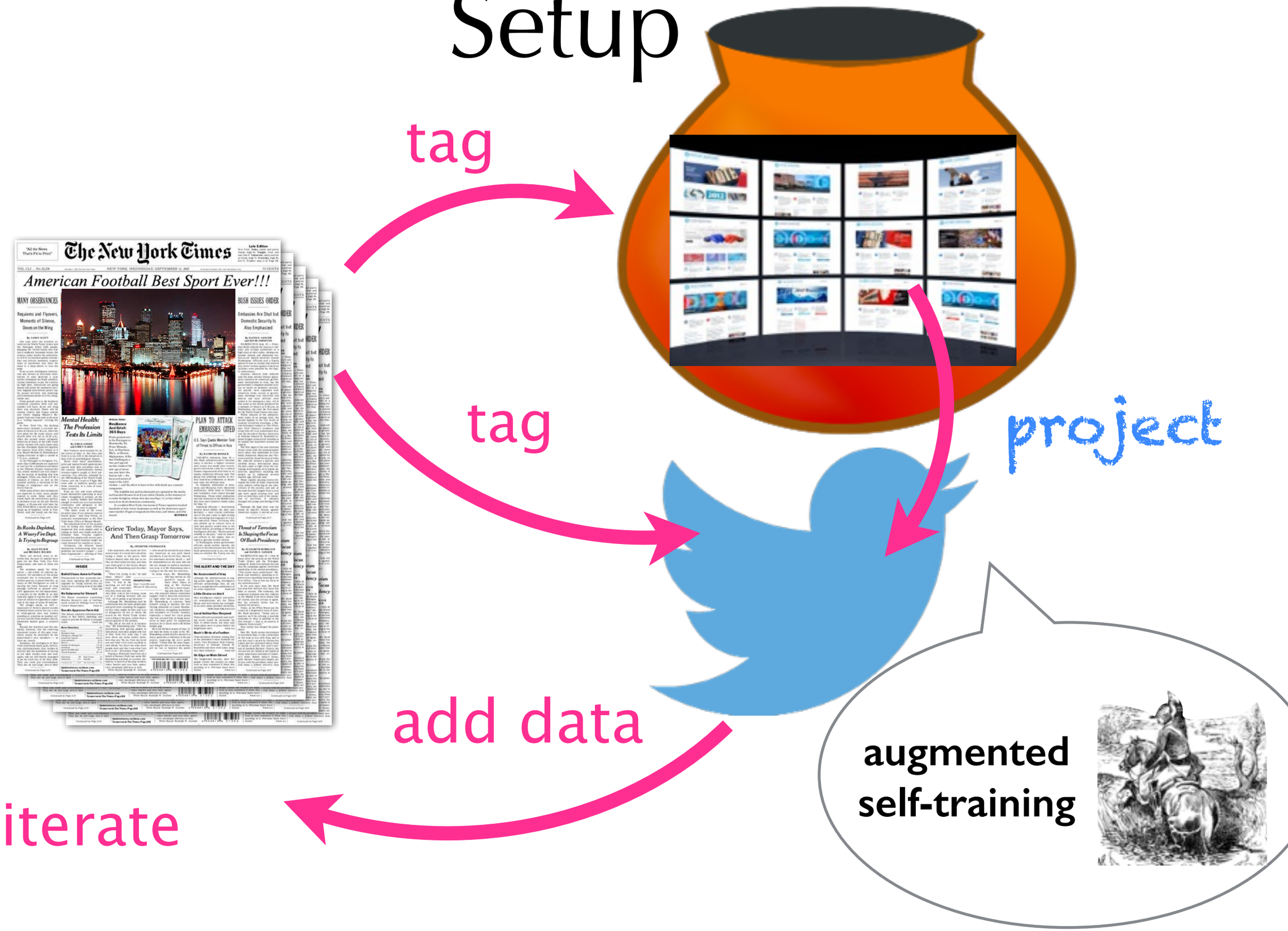
Setup



Setup

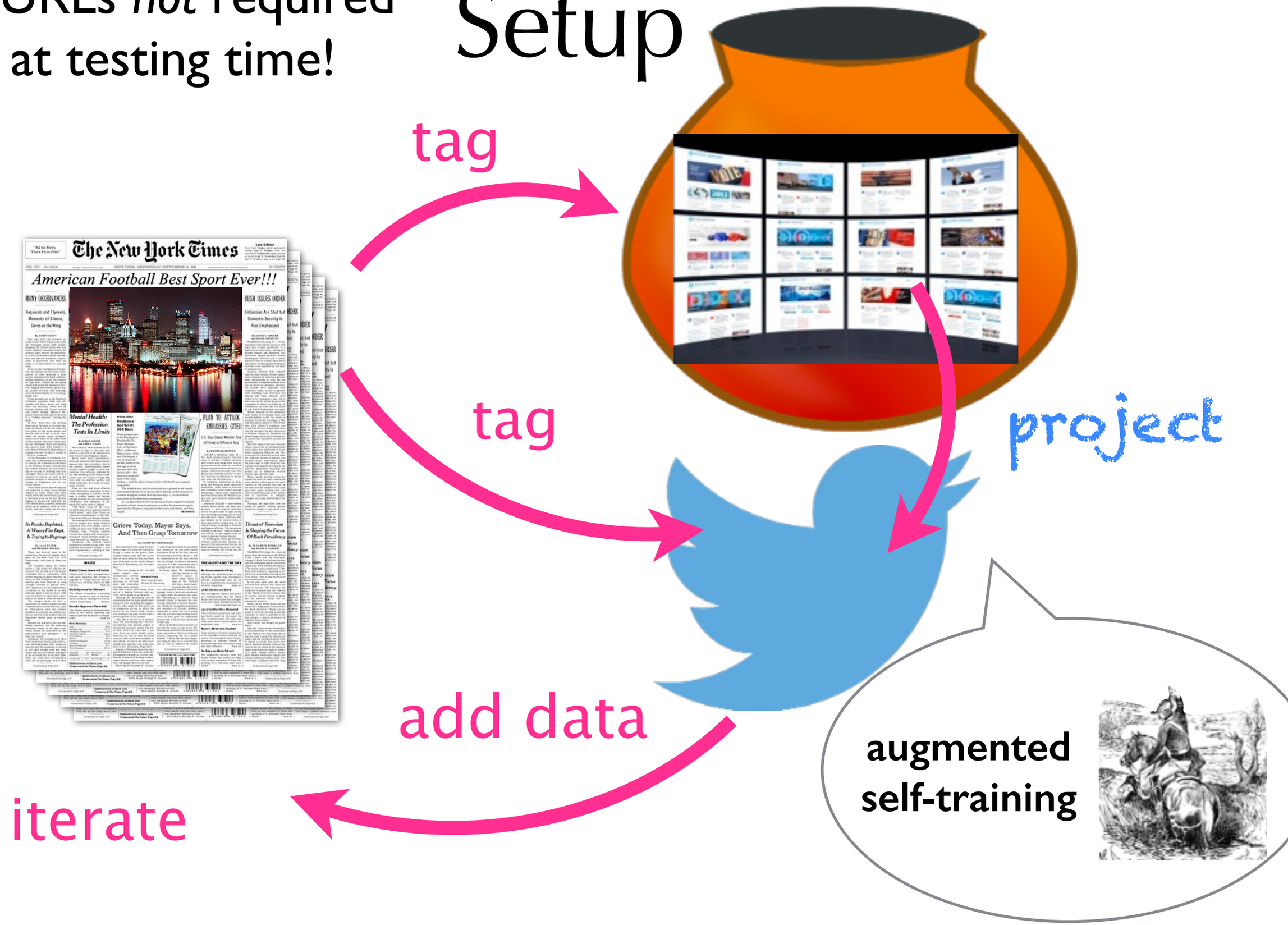


Setup



NB: URLs *not* required at testing time!

Setup



$X = \{\langle \mathbf{x}_i, y_i \rangle\}_{i=1}^{N}$ labeled tweets
$U = \{\langle \mathbf{x}_i, w_i \rangle\}_{i=1}^{M}$ unlabeled tweet-website pairs
I iterations
$k = 1000$ pool size
v=train(X) base model
for $i \in I$ **do**
  for $\langle \mathbf{x}, w \rangle \in \mathbf{pool}_k(U)$ **do**
    $\hat{y}$=**predict**$(\langle \mathbf{x}, w \rangle; \mathbf{v})$
    $X \leftarrow X \cup \{\langle \hat{y}, \mathbf{x} \rangle\}$
  end for
  v=train(X)
end for
return v

$X = \{\langle \mathbf{x}_i, y_i \rangle\}_{i=1}^{N}$ labeled tweets
$U = \{\langle \mathbf{x}_i, w_i \rangle\}_{i=1}^{M}$ unlabeled tweet-website pairs
I iterations
$k = 1000$ pool size
v=train(X) base model
for $i \in I$ **do**
  for $\langle \mathbf{x}, w \rangle \in \mathbf{pool}_k(U)$ **do**
    $\hat{y}$=**predict**($\langle \mathbf{x},$  ; **v**)
    $X \leftarrow X \cup \{\langle \hat{y}, \mathbf{x} \rangle\}$
  end for
  v=train(X)
end for
return v

additional ingredient

tweet

$X = \{\langle \mathbf{x}_i, y_i \rangle\}_{i=1}^{N}$ labeled tweets
$U = \{\langle \mathbf{x}_i, w_i \rangle\}_{i=1}^{M}$ unlabeled tweet-website pairs
I iterations
$k = 1000$ pool size
v=train(X) base model
for $i \in I$ **do**
  for $\langle \mathbf{x}, w \rangle \in$ **pool**$_k(U)$ **do**
    $\hat{y}$=**predict**$(\langle \mathbf{x},$  $; \mathbf{v})$
    $X \leftarrow X \cup \{\langle \hat{y}, \mathbf{x} \rangle\}$
  end for
  v=train(X)
end for
return v

additional ingredient

tweet

5 Approaches

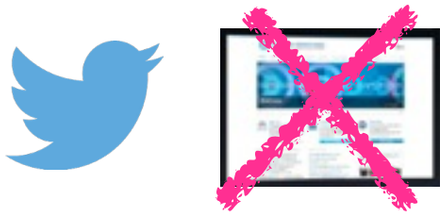
predict(<  **,**  **>)**

predict(<  ,  >)

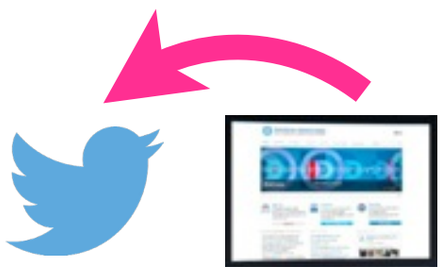
 

plain

predict(<, >)

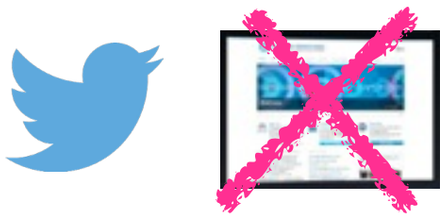


plain

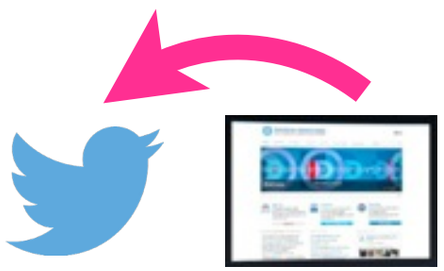


web

predict(<, >)



plain

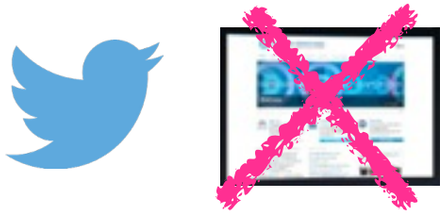


web

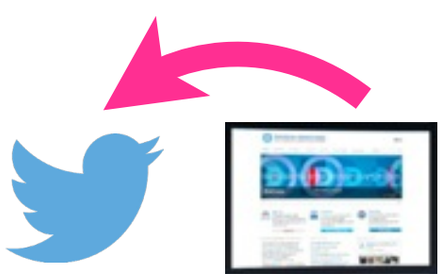
dict

93% unambiguous

predict(<, >)



plain



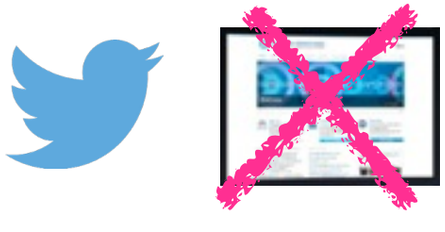
web

web<dict

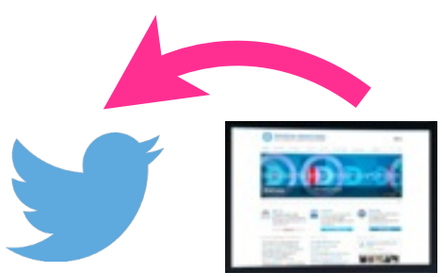
dict

93% unambiguous

predict(<  ,  >)



plain



web

web<dict

dict

dict<web

93% unambiguous

Experiments

POS

Train

+

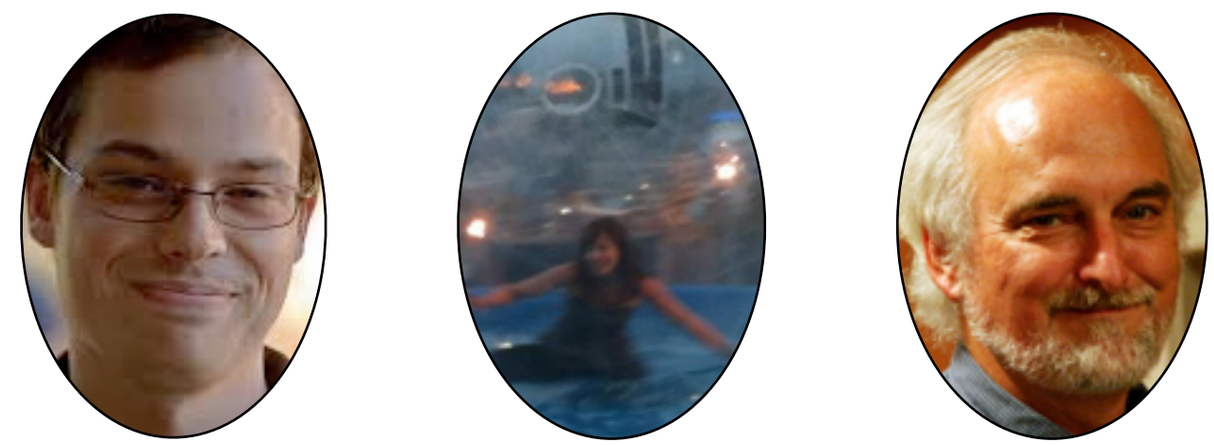
Test

NER

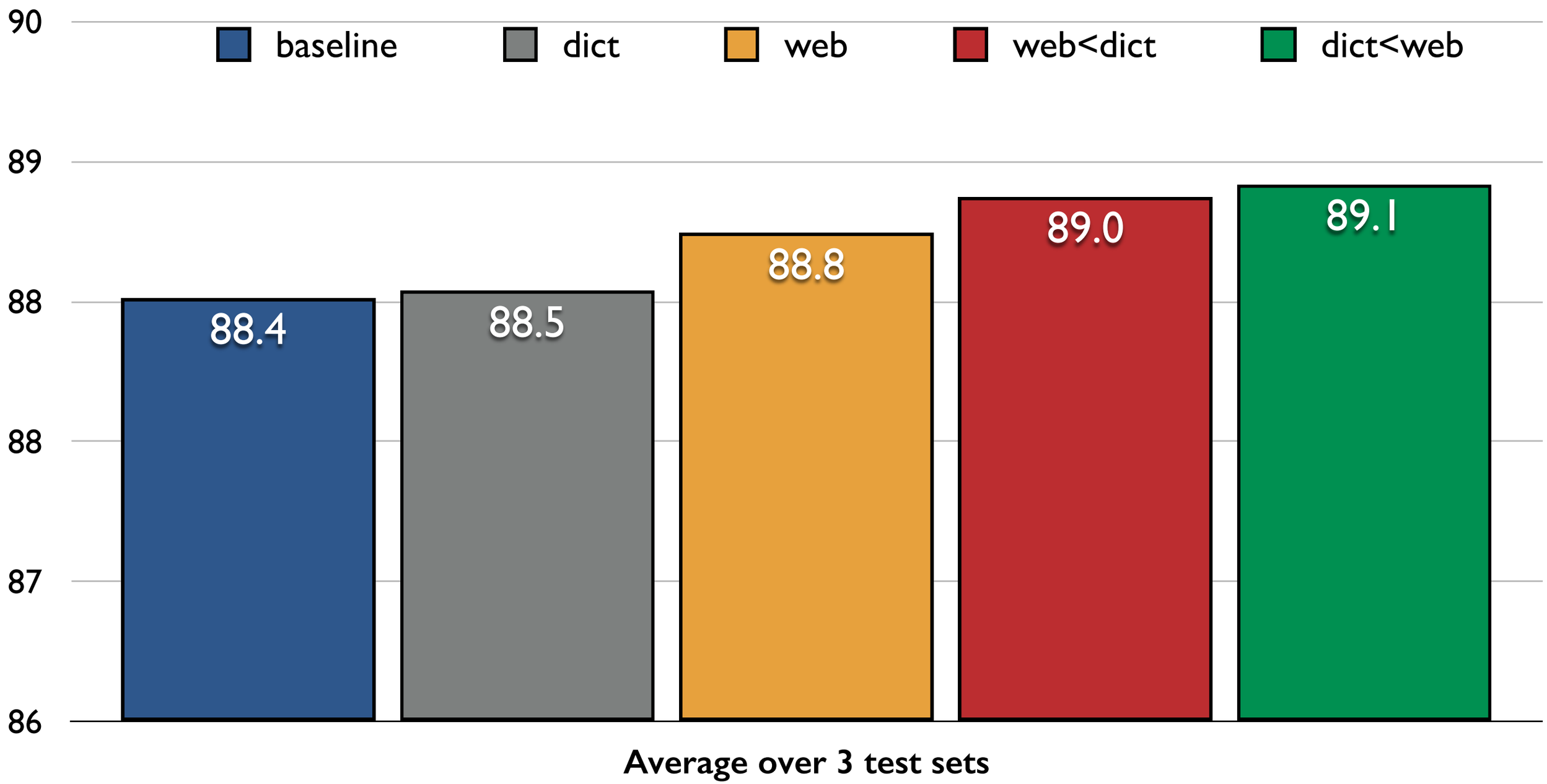
Train

+

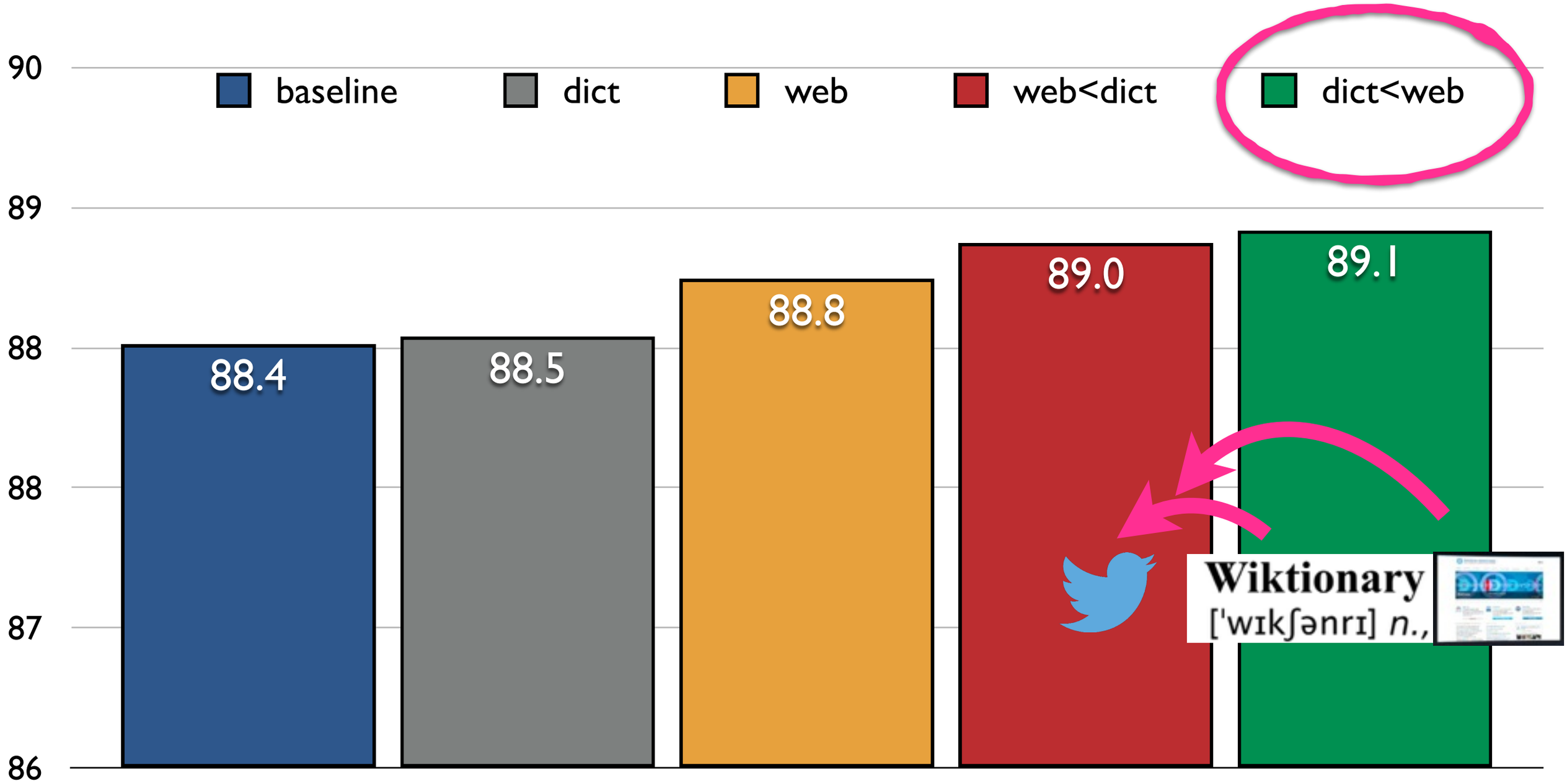
Test



POS Results: Which projection method?

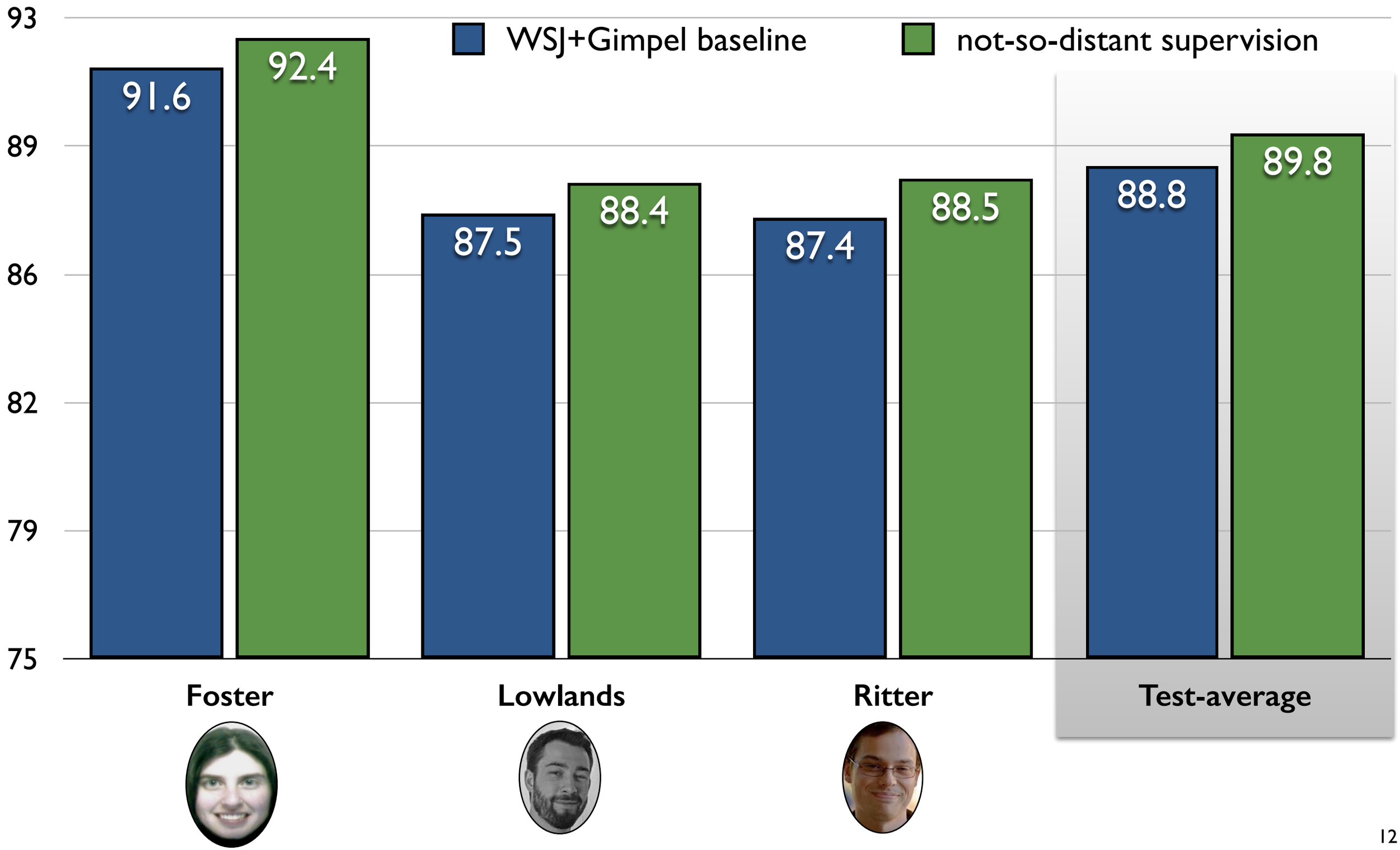


POS Results: Which projection method?

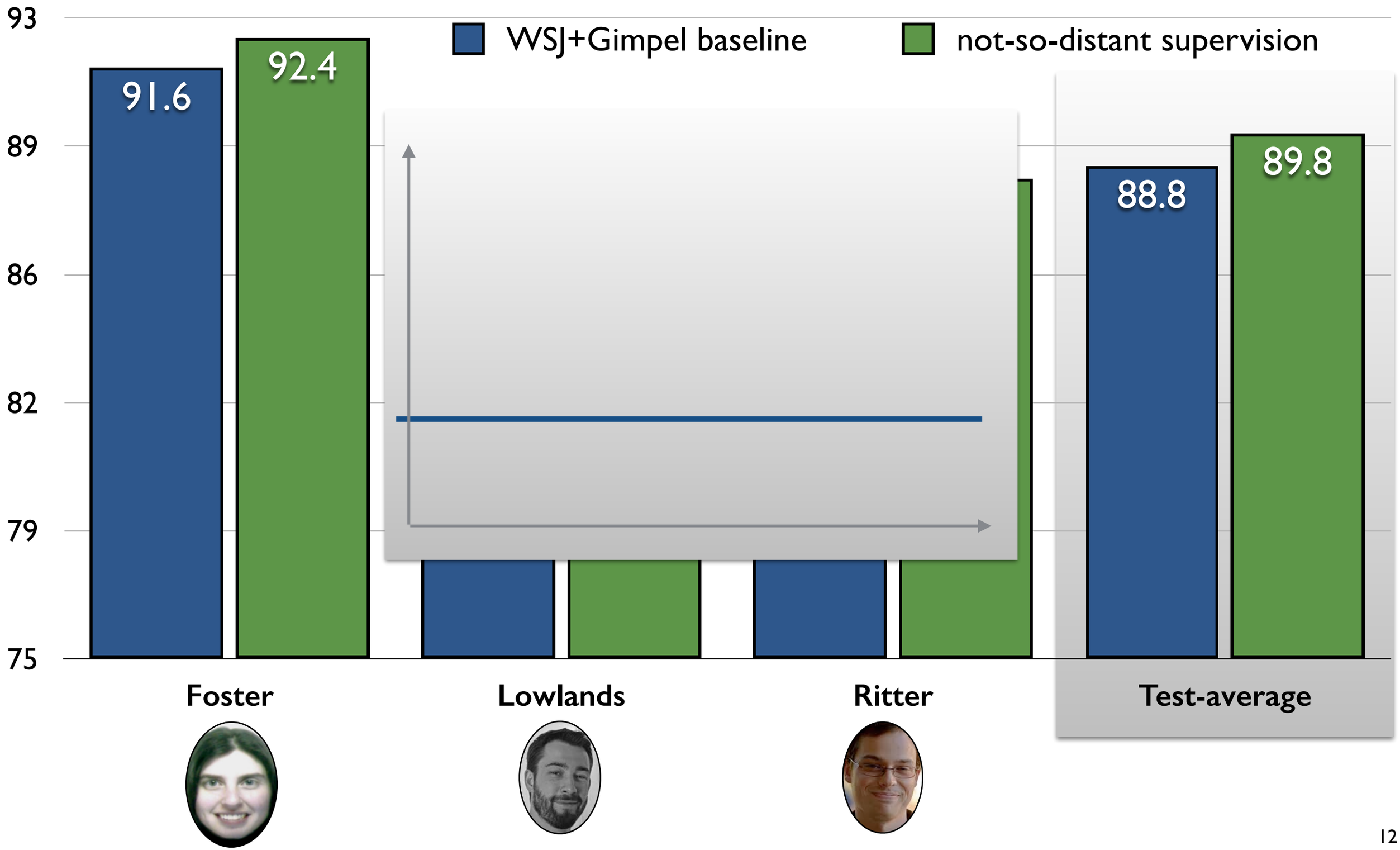


Average over 3 test sets

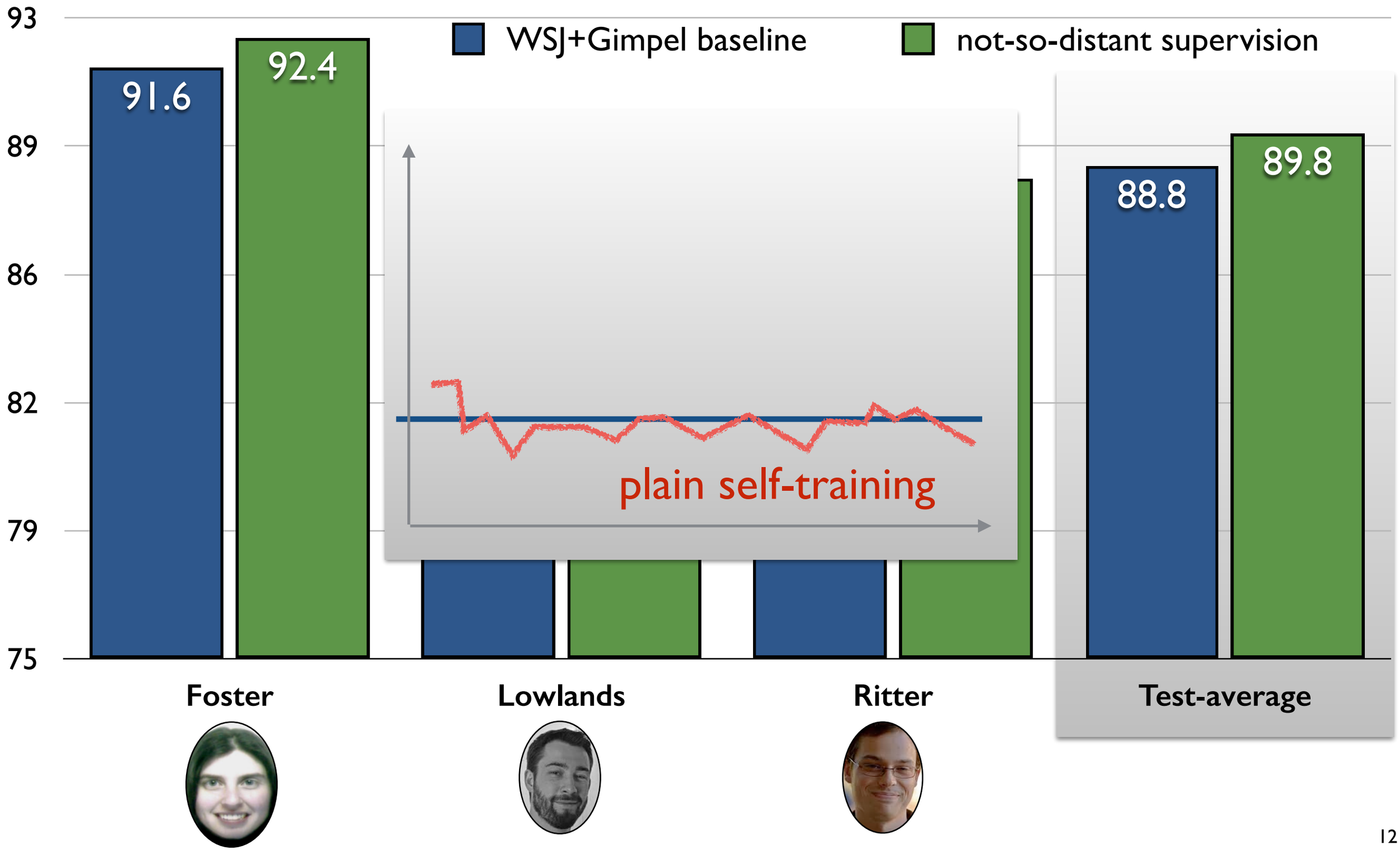
POS Results


93
89
86
82
79
75
WSJ+Gimpel baseline
not-so-distant supervision
91.6
92.4
87.5
88.4
87.4
88.5
88.8
89.8
Foster
Lowlands
Ritter
Test-average

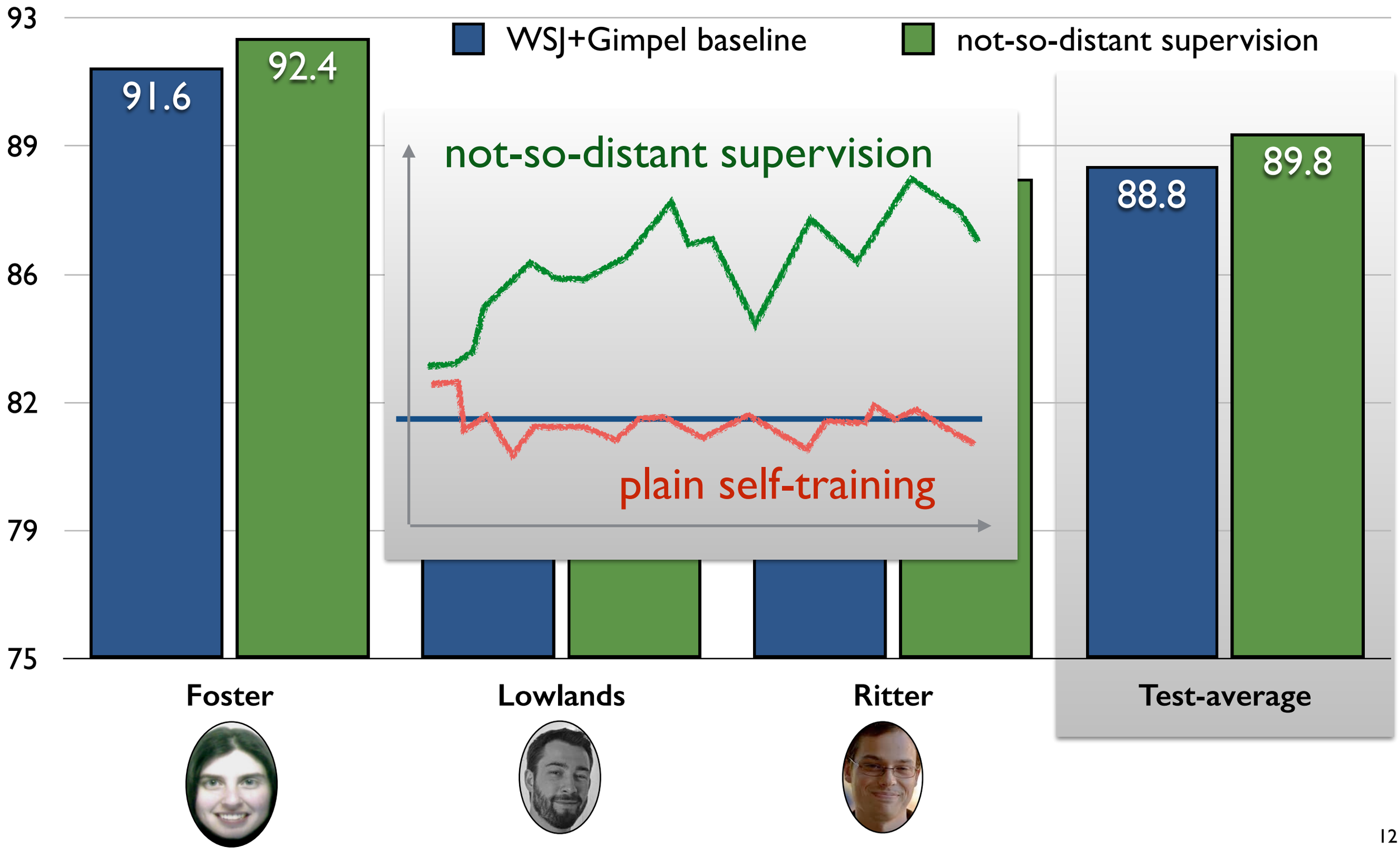
POS Results


93
89
86
82
79
75
WSJ+Gimpel baseline
not-so-distant supervision
91.6
92.4
88.8
89.8
Foster
Lowlands
Ritter
Test-average

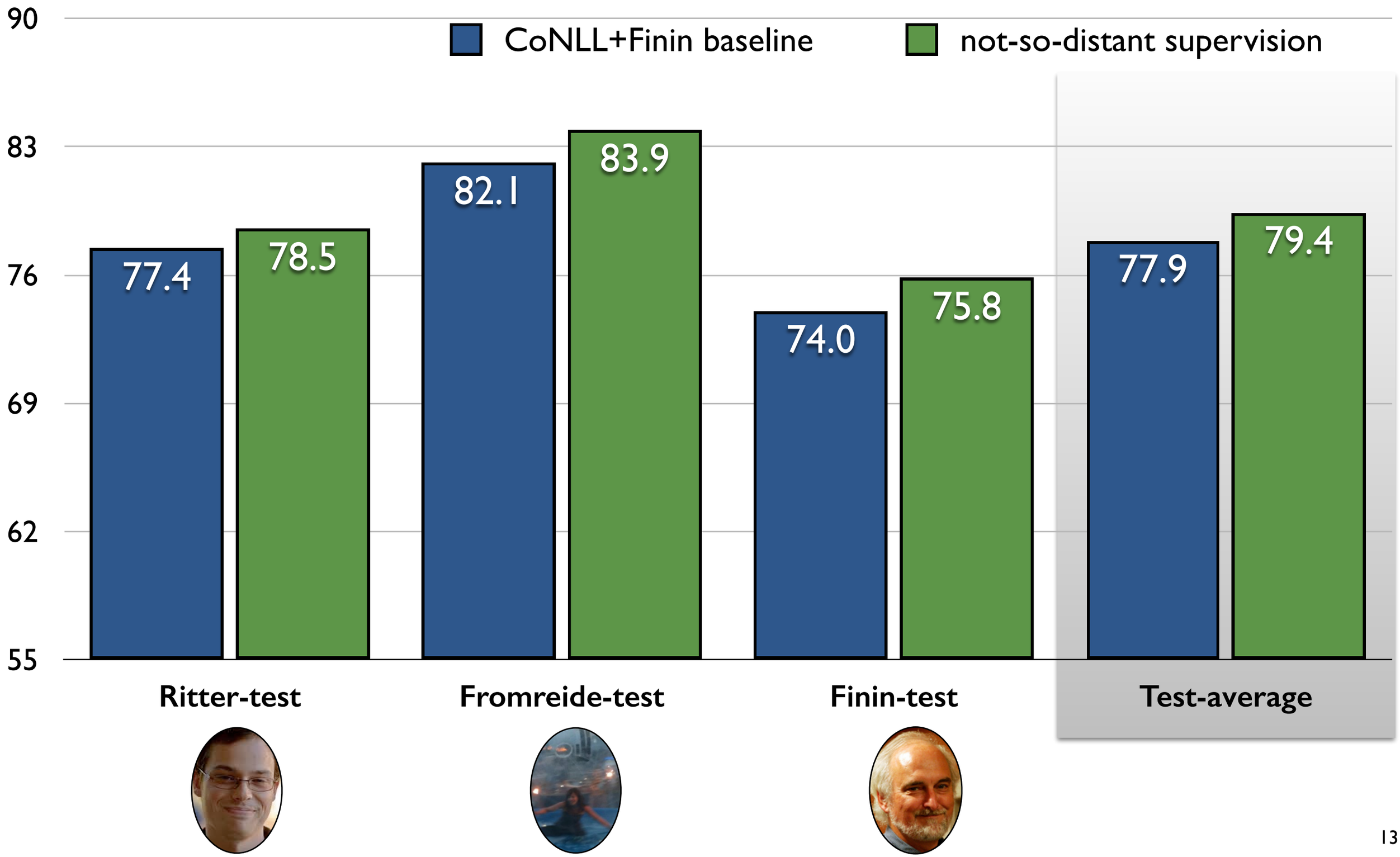
POS Results


93
89
86
82
79
75
WSJ+Gimpel baseline
not-so-distant supervision
91.6
92.4
88.8
89.8
plain self-training
Foster
Lowlands
Ritter
Test-average

POS Results


93
89
86
82
79
75
WSJ+Gimpel baseline
not-so-distant supervision
91.6
92.4
88.8
89.8
not-so-distant supervision
plain self-training
Foster
Lowlands
Ritter
Test-average

NER Results


90
83
76
69
62
55
CoNLL+Finin baseline
not-so-distant supervision
77.4
78.5
82.1
83.9
74.0
75.8
77.9
79.4
Ritter-test
Fromreide-test
Finin-test
Test-average

Analysis

Projection Examples

	Snohomish	*Bakery*	*Salmon-Safe*	*parks*
initial tag	ADJ	NOUN	NOUN	NOUN
projected	**NOUN**	NOUN	**ADJ**	NOUN

Limitation

NOUN

If I gave you one wish that will become true.

NOUN VERB

What's your wish ?... ? i wish i'll get 3 wishes

from you :p URL

Limitation

NOUN

If I gave you one wish that will become true.

NOUN **NOUN**

What's your wish ?... ? i wish i'll get 3 wishes

from you :p URL

Analysis

- improvements due to richer linguistic context

	Man	*Utd*
ARK	PRT	NOUN
Our	NOUN	NOUN

	Radio	*Edit*
ARK	NOUN	VERB
Our	NOUN	NOUN

- somewhat arbitrary differences

	Nokia	*D5000*
ARK	NOUN	NUM
Our	NOUN	NOUN

	love	*his*
ARK	VERB	DET
Our	VERB	PRON

Take-home-message

Take-home-message

✓ 20% of tweets contain URLs

Take-home-message

✓ 20% of tweets contain URLs

✓ Semi-supervised learning + not-so-distant supervision helps to bridge the gulf

Take-home-message

✓ 20% of tweets contain URLs

✓ Semi-supervised learning + not-so-distant supervision helps to bridge the gulf

✓ Models available at:

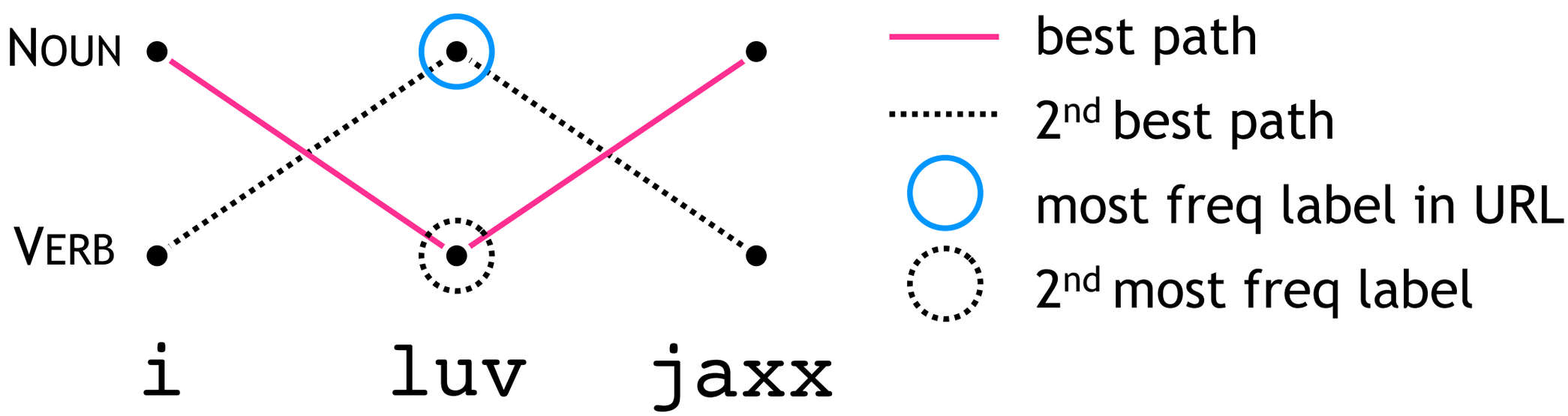
https://bitbucket.org/lowlands/ttagger-nsd

Bitbucket

Questions?

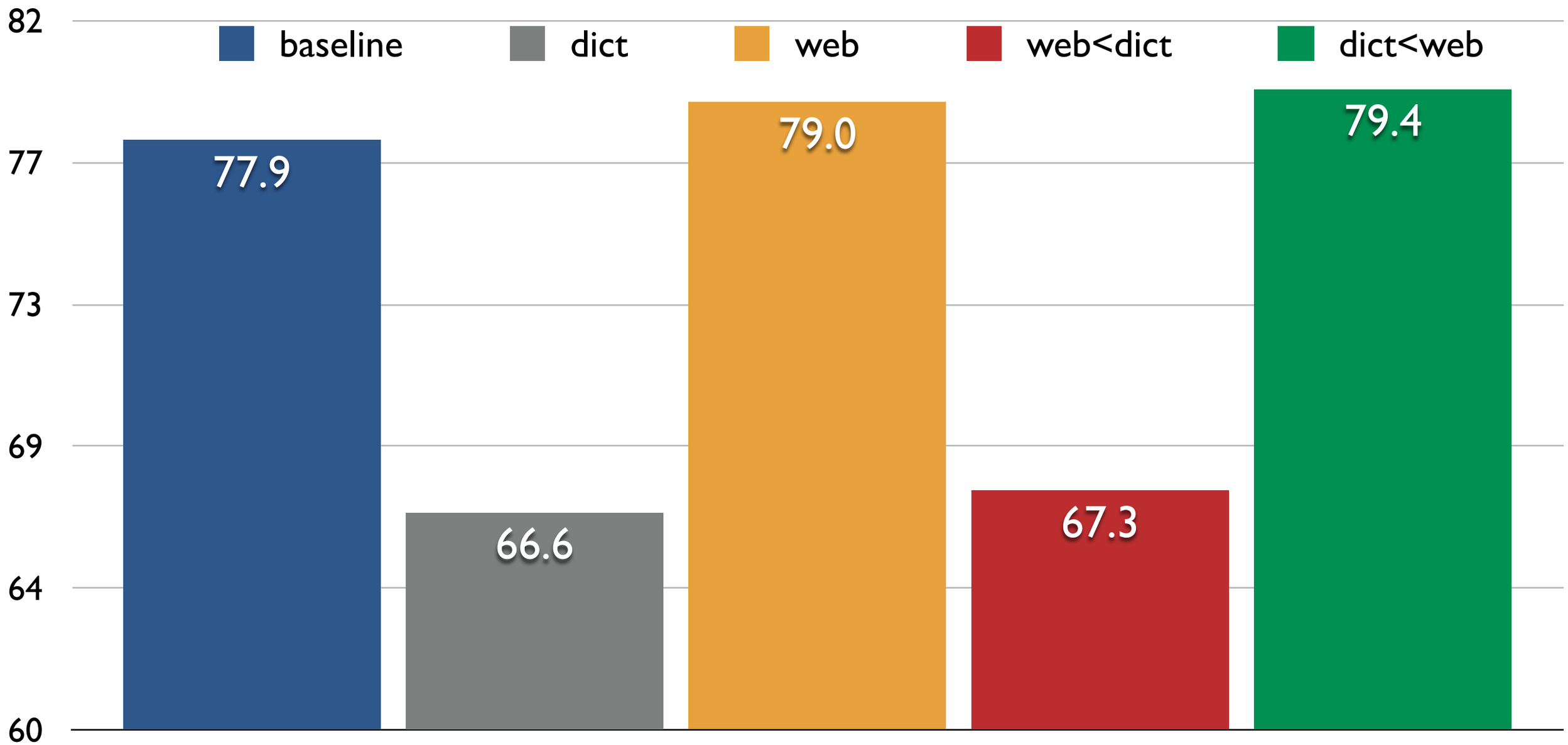
Thanks!

Alternatives

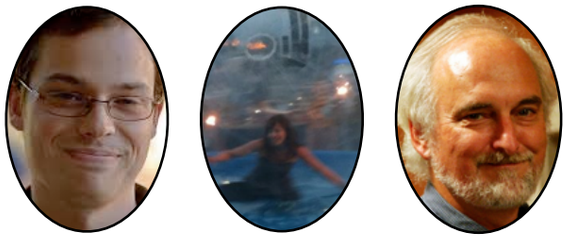


PROJECTION: NOUN–NOUN–NOUN
HARD CONSTR.: VERB–NOUN–VERB
SOFT CONSTR.: NOUN–VERB–NOUN

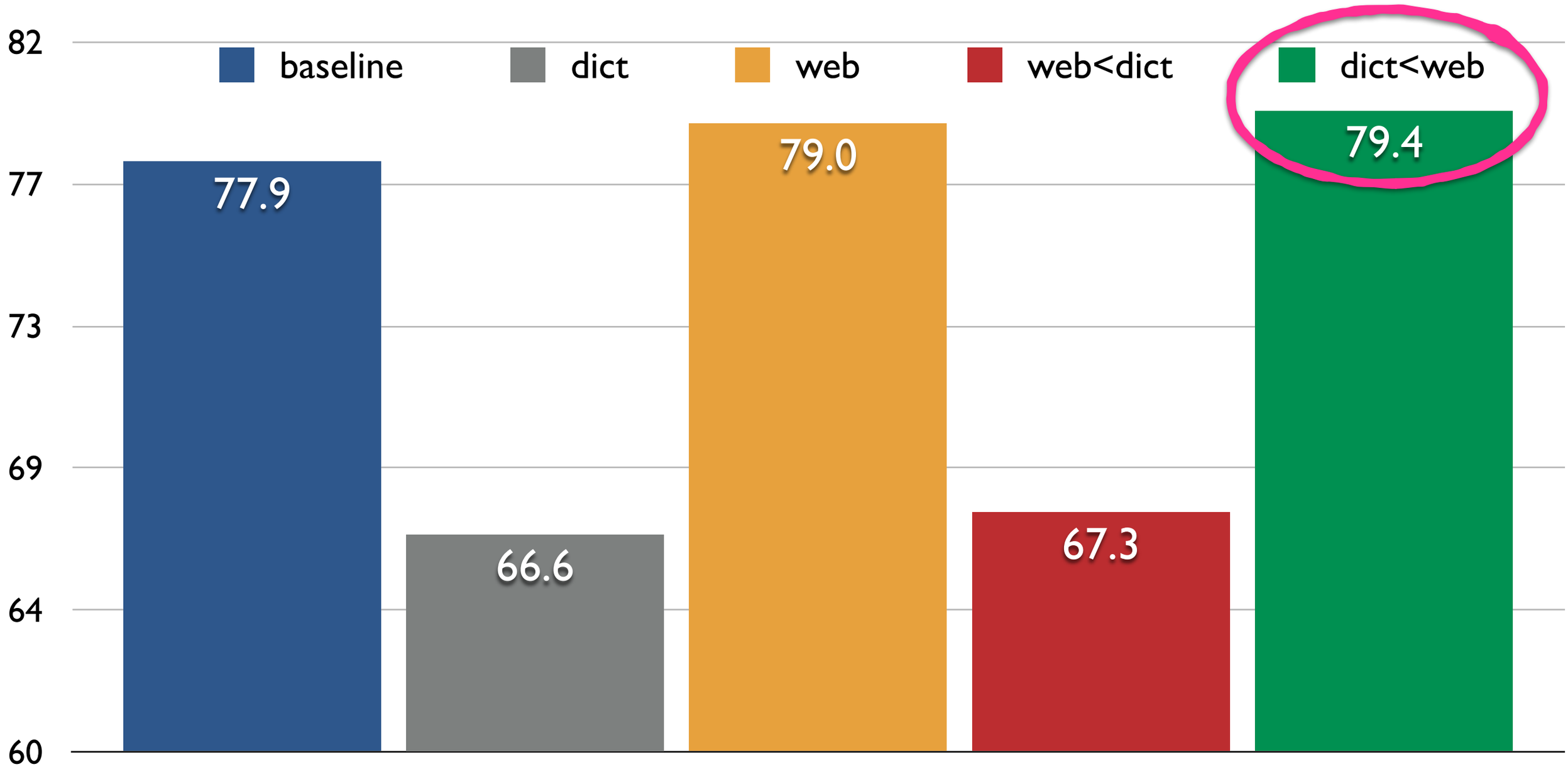
NER Results



Average over 3 test sets



NER Results



Average over 3 test sets