Text Mining of Scouting Reports as a Novel Data Source for Improving NHL Entry Draft Analytics

Timo Seppa Analytics consultant

Ottawa Hockey Analytics conference 2017 – May 6, 2017

Scouting or stats?

Why not both?

But in a different way than you might think...

Scouting reports can provide a valuable textual data source for an analytics approach



Over a half million words

Seven years of scouting reports were used, for 2010-2016

Textual data comprised of over **583,000** words

More can be added easily

18-2016 BLACK BOOK

Abramov, Vitalii LW - Gatineau Olympiques (QMJHL) 5'10", 170 HockeyProspect.com Ranking: 29

Vitall entered his first full season in the QMUHL as he was selected in the CHL limport Draft in 2015 coming over from the Russian JL league. He did not look out of place at all where in his 1st game he scored 4 points and finished the year with 93 points ted for 5th in league scoring.

Vitall is an undersized forward with great speed and high hockey IQ that creates a lot of offinative chances for himself and hist teammates. He is a good skater and is really good on his edges and has good mobility to get out of tight areas, and has quick acceleration to separate himself from defenders. He has a quick and accurate release with his shot and is really good at coming off the half boards or off the rush and outting into the solt and taking shot on net. Vitall has very good stickhanding ability where he can handle the puck at top speed coming into the zone and can stickhanding in a phone booth. He has high hockey IQ, he is really good at anticipating how the play is going to develop in front of him, where his teammates are going to be on the ice, he is also very good at finding the open areas of the zone to get set up for an offensive chance. He hockey sense really shows on the power play where he can be very creative and seems to always find the open man, or taking the proper shot on net. Vitall ials can be very creative coming into the zone of the rush, whether its stopping and finding the trailer or making a move on the D, he always seems to show the game down. We would like to see him go hard to the net or to the dirty areas of the rink; he trands to do also to his work on the perimeted.

Vitalli defensive of game would be the thing he has to work on for the next level, he needs to get stronger along the boards, there were a lot of times in our viewings he was loaing the puck along the boards for the break out or if the defenseme pinches down. He also tends to run around too much in his own zone, instead of focusing on his own man and staying with them. Once he does get the puck along the boards he's good at these touch passes to his center curling or touch pass of the boards. Vitall has 1st round talent and hockey smarts, he just needs to focus on getting bigger and stronger and working on the little parts to the game away from the puck, be could be a very good hockey player at the next level and be very effective.

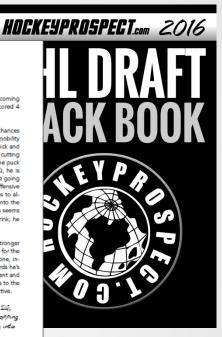
Quatable:" Vitalli is an undersized forward with good stating ability and High hockey IQ, In san create a lad of offensive characs, whether it's coming down the wing and stapping to wait for the play to develop in front of him, or coming down the wing and cutting into the slot wing his quick release and accurate shot - HP Scout, Justin Growie

Quotable:" He is a hard working forward that congretes every shift he is on the ice, seldom takes a night off $^{\prime}$ - NHL Scout (April 206)

Guestable" Vitall's vision on the ice is outstanding, he somehow finds the open non that nabody could find. I love his ability on the power play to control the puck and understand where and who the puck should go to '- MP Scout, Justin Sproule (Mpri) 206)

Quatable: One of our top interviews, he knows his game and gave great answers to our quastions." - NHL Societ (Combine week)

Qustable: "(Laughing) He was good (combine interview), I liked the kid. His english was The and he was group oriented, he spele to everyone in the room. He sold he knows he's s'i' but it doesn't deter him and that he uses different ways to get to the net. He had some inner drive in him. Good kid. He knows what he is'. MHL Scauld (Combine week)



Massaging the text: exclusion lists, lemmatization

List: User de	fined	•	Sort by: Case number						
Word: ABILITY	,	C	iontext delimiter: Paragraph						
	FREQ	CASENO		KEYWORD		CODE	CATEGORY	· •	
SHOW ALL ITEMS	=	2	ns and even coaches. Etem's game is built around speed. He uses his top-notch skating	ability	to gain space on opposing defensemen, protect the puck and create opportunities. Alth	c ANA	Teams		
1731 - ability	9	2	at times on the defensive side of the puck. Etem certainly has the speed and offensive	ability	to be a productive NHL player. Whether or not he is a 3rd line center or 1st line will dep	e ANA	Teams		
+ 349	8	2	pend on how well he develops his overall game. With his attitude and natural leadership	ability	, it is safe to say that Etem could slide into a 3rd line NHL Role, but his speed makes top	- ANA	Teams		
307 - ability	7	4	e forecheck, doesn't take very many shifts off and prides himself in his strong two-way	ability	. Rakell is particularly effective when on the penalty kill. Has good hands and controls	h ANA	Teams		
± 513		4	we project him to play if he cracks an NHL lineup some day. We don't see top 6 scoring	ability	. In fact we don't think he is close as far as scoring and NHL shooting ability. Rakell sho	ANA	Teams	_	
 1809 1785 - ability 	6	4	op 6 scoring ability. In fact we don't think he is close as far as scoring and NHL shooting	ability	. Rakell showed nicely at the World Juniors in Buffalo by banging players around and p		Teams		
1785 - adility + 66	6	8	Itors mirroring forward Victor Rask's production. Karlsson showed off his great skating	ability	, combined with his physical play. Karlsson plays at a very high energy, but is also ver		Teams	-	
± 125	5	8	hysical play. Karlsson plays at a very high energy, but is also very creative, and has the	ability	to make high difficulty passes under pressure.	ANA	Teams	-	
1863 - ability	5	11	p for it, he just didn't need to in Lincoln. Roy is easily a draftable player. His goal scoring	ability	was something unlike the USHL has seen in years. In terms of projecting a round, with		Teams	-	
1822 - ability	5	11	is patient enough to wait the extra half-second to make the better play. His goal scoring	-	is off the charts (scoring 54 goals in the USHL is unheard of and his goals were no flu		Teams	-	
1733 - ability	5	12	I's, 5'9.25" 170 Besse is a sleeper Minnesota hockey player who has great goal scoring		. He played this season for Wayzata of Minnesota high school hockey. You just cant di		Teams	-	
1716 - ability	5	15					Teams	-	
+ 2037	5		at shot and is a natural goal scorer. He's a good skater and you can see a ton of natural	-	. He had some shifts that looked kind of lazy. He has good instincts and knows how to				
1924 - ability	5	17	he could get them on net more often. The biggest concern of Theodore is his defensive		. He has a tough time getting into proper position without the puck, and often gets caug		Teams		
790 - ability	5	17	as skilled with the puck as Shea Theodore in the upcoming NHL Entry Draft. He has the	ability	to go coast to coast and score a goal to change the momentum of the game, and provi		Teams	-	
1204 - abilities	4	17	ng reach to keep pucks away from opponents as he quickly picks up speed. He has the	ability	to quickly identify open lanes, and changes his direction in a hurry to drive to the net an		Teams	- 1	
± 117	4	17	I the tools to become an NHL defenseman one day, but his play without the puck and his	ability	to think the game will determine just how successful he can be. Quotable: "I think his o		Teams	- 1	
± 1126	4	18	, aware, poised, and fundamentally sound, with pinches at the right times, puck-rushing	ability	, excellent passing, the ability to quarterback a power play and get pucks on net, and in	II: ANA	Teams		
143 - abilities	4	18	entally sound, with pinches at the right times, puck-rushing ability, excellent passing, the	ability	to quarterback a power play and get puc WordStat 7.1.4 -	_		_	_
141 - ability	4	19	puck well and can carry the puck end to end seemingly at will. It was this puck rushing	ability	along with his hockey sense and quick st	= Ere	equencies	🚯 Extra	ction
1358 - ability ± 1336	4	20	ur viewings winning some draws clean. Flashed a good wrister and some stickhandling	abilities		= 114	quencies	CAUG	cuor
1849 - ability	4	20	'0" 165 Kase has shown us flashes of having great skills and excellent speed. Liked his	ability	to change speeds quickly, delaying and d Dictionaries				
1849 - ability 1810 - ability	4	23	ind anticipates developing plays quickly. He doesn't get enough credit for his playmaking	ability	, probably because he has had so little ta			Zedit	8
± 178	4	23	e is one of the most NHL ready prospects in this draft. He blends size, strength, scoring	ability	and a mean streak into his game. We exp			a cuit	Ø
1768 - ability	4	25	arsson is a promising defenseman blessed with projectable size and impressive skating	ability	. At 17 he already received some icetime Stemming: English (Porter)		-		
1748 - ability	4	26	ongest player in corners and you don't see him dishing out any big hits. Nattinen has the	ability	to make his linemates better, he sees the	tion		M main	
1743 - ability	4	26	eup to be physical. I think his strength is defensive hockey. While He has shown me the	ability	to distribute the puck, I think he really exc	luon		Curr	-
± 1721	4	27	all three forward positions. Sideroff is a particularly good fit on the wing because of his	ability	to drive wide with his speed. Indeed, Side Exclusion: English.exc		- 6	Gopen	9
+ 17	4 🔻	29	and the original position of a particularly good in or the wing boards or in a any accurate shot that he is able to keep low for deflections or rebounds. His shot and		to make the right play is especially helpful Categorization:			Copen	

from the point. His slap shot could have more velocity on it, while he could get them on net more often. The biggest concern of Theodore is his defensive abilities. He has a tough time getting into proper position without the puck, and

Lemmatization: grouping together different inflected forms of a word for analysis

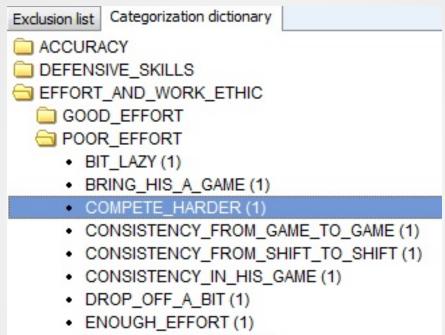
Exploring the data via topic extraction

ŀ	Dictionaries	11 Options Frequencies Straction	S.Cooccur	rences 🖁	Crosstab	III Key	word-In-Context	-Classification
-	Topics D Phras	es 💈 🎭 Named Entities 🛛 🍄 Misspellings & Unkn	owns					
iegt	mentation: by Parage	aph 🔹 No. topics: 20 🚡 Loading: 0.30 👘	Run	11 *	rej 🐒	0 0		S 💼 😤 🖬 🖉
NO	NAME	KEYWORDS	EIGENVAL	% VAR	FREQ	CASES	% CASES ^	ut. 🔄 🖾 Rate per 10,000 words 🔹 BY CODE 🔹
1	RUSHING DEFENSEMAN	DEFENSEMAN; BLUELINE; GAP; RUSH; MOBILITY; FOOTWORK	13.59	1.25	1779	999	47.41%	
2	HOCKEY PLAYER	HOCKEY; SENSE; SMART; IQ; QMJHL; PLAYER; GREAT	3.44	1.15	3841	1728	82.01%	24-
3	MINOR MIDGET	MINOR; MIDGET; OHL	3.27	0.92	979	685	32.51%	22-
4	GOALTENDER	GOALTENDER; SAVE; ATHLETIC; CREASE; REBOUND; GLOVE; POST	3.13	1.26	911	464	22.02%	20-
5	BETTER	BETTER; MORE; CONFIDENCE	2.84	0.98	1722	1297	61.56%	18
6	CYCLE; NET DRIVE	CYCLE; DRIVE; LOW; NET; PROTECT; DOWN; CREATE; PRESSURE; SHOT; FORECHECK; DEEP	2.52	1.14	3701	1565	74.28%	16-
7	WORLD CHAMPIONSHIP	WORLD; CHAMPIONSHIP; JUNIOR; TOURNAMENT; CHALLENGE; PERFORMANCE	2.34	0.93	990	594	28.19%	
8	COMBINE INTERVIEW	INTERVIEW; COMBINE	2.21	0.75	157	102	4.84%	
9	NEED; FRAME	NEED; FRAME; IMPROVE; SKATE; STRENGTH	2.20	1.05	3712	1783	84.62%	
0	PENALTY KILL; BLOCK SHOT	PENALTY; KILL; BLOCK; SHOT; SITUATION	2.19	0.95	1579	860	40.82%	
1	ABOVE AVERAGE	AVERAGE; ABOVE; WHL	2.06	0.78	947	675	32.04%	
2	AGILITY; QUICK	AGILITY; QUICK; ACCURATE; WRISTSHOT	2.02	1.03	1490	1054	50.02%	
3	BATTLE; WIN	BATTLE; WIN WORK; HARD; ETHIC	1.99	0.89	2745	1372	65.12%	
4	EXCELLENT; ABILITY	EXCELLENT; ABILITY; DEFENDER	1.90	1.07	1688	1255	59.56%	
5	PHYSICAL	PHYSICAL; HIT; OPPOSITION	1.84	0.86	1749	1204	57.14%	ANA ANA ANA ANA ANA ANA ANA ANA ANA ANA
6	ZONE; PUCK IN THE OFFENSIVE	ZONE; OFFENSIVE; PASS; STICK; PUCK; DEFENSIVE	1.80	1.29	6238	1924	91.31%	
7	CONSISTENT	CONSISTENT	1.79	0.80	224	224	10.63%	and and a second s
8	TAKE AWAY; QUICK	AWAY; QUICK; OFF; TAKE; TRY; TIMES	1.77	1.22	3850	1698	80.59%	La F E Rate per 10,000 words • BY CATEGORY •

Creating categories -> sentiment analysis

	WordStat 7.1.4 -	No.				-	
11/10	🛨 🤗 Dictionari	es 🔃 Opti	ons 📻 Fre	quencies	🚯 Extra	iction 义	Cooccurren
	Dictionaries						
	Preprocessing:	3-grams		•	🞽 Edit	滂 New	💼 Delete
	Stemming:	English (Porter))	•			
	Substitution:	English - Lemma	atization	•	🞽 Edit	滂 New	💼 Delete
	Exclusion:	English.exc		•	🔒 Open	🛞 New	💼 Delete
	Categorization:	Hocketdraft1.c	at	•	产 Open	😁 New	💼 Delete
		Level: Up to lev	/el 🔻 2		Categories	only 📃 U:	se full path as
		<u>_</u> .					
	Dictionary viewer						
	+ Add	to Options	Exclusion list	Categoriza	ation dictiona	ry	
	- Remove	Suggest					
			DEFENS			\ \	
	E <u>d</u> it	A Sort	N	_AND_W	ORK_ETHIC T	, 	
	🔄 Undo	និន្ត្រៃ Merge		R_EFFOR			
	Print	Comment					
		<u> </u>		_	IG_AND_IN RSATILITY		
	📃 Drag & drop edi	ting		_			
	Select items		DF_NET_DF	RIVE_AND	NETFRON	т	
	Allow overlaps		DASSING				
	View definition		PENALT	-			
	Show warnings		PHYSIC/	ALITY_ANI	D_SIZE		
	_ _		DOISE				
			DUCK_E	_			
			DUCK_F		ON_AND_C	YCLING	
			DUCK_S	_	_	SSION	
			CUICK_				
						RING	
			E SHOT_E		,		
				-			

What constitutes "Poor effort"?



LEFT_I_WANT_MORE (1)

Some of the 2-, 3-, 4-, and 5-word phrases categorized as POOR_EFFORT

Is "compete hard" = "compete harder"?

[Case #4]

He just needs to compete harder, pure and simple

[Case #10]

he ought to use that 6'1 frame to compete harder

[Case #49]

he'll have to compete harder to get the puck

[Case #52]

he is kid who could make us look like idiots...if he ever started to compete harder

[Case #60]

We'd like to see him block shots and compete harder in one-on-one battles

[Case #5]

I'm not sure he always competes hard enough

[Case #9]

He was able to win puck battles against men, compete hard along the boards

[Case #9]

He competes hard very consistently

[Case #12]

He competes hard and is a tough kid who plays a lunch pail blue collar game

[Case #13]

He shows good backcheck and competes hard defensively

Conclusion: COMPETE_HARDER is categorized as POOR_EFFORT, while COMPETE_HARD is not

What can we say about undrafted players?

Skill									
OOR_VISION	Undrafted vs. drafted							291%	
POOR_PUCK_SKILLS	Undrafted vs. drafted						246%		
POOR_PASSING	Undrafted vs. drafted					21	16%		
POOR_SHOOTING_AND_SCORING	Undrafted vs. drafted		182%						
POOR_PUCK_BATTLES	Undrafted vs. drafted		164%						
POOR_PHYSICALITY_AND_SIZE	Undrafted vs. drafted		142%						
GOOD_PUCK_BATTLES	Undrafted vs. drafted		136%						
POOR_ACCURACY	Undrafted vs. drafted		129%						
GOOD_EFFORT	Undrafted vs. drafted		128%						
GOOD_NET_DRIVE	Undrafted vs. drafted	122%							
POOR_DEFENSIVE_SKILLS	Undrafted vs. drafted	116%							
GOOD_PHYSICALITY_AND_SIZE	Undrafted vs. drafted	116%							
POOR_INTANGIBLES	Undrafted vs. drafted	116%							
GOOD_POWER_PLAY	Undrafted vs. drafted	115%							
GOOD_PENALTY_KILLING	Undrafted vs. drafted			1149	6				
POOR_DECISION_MAKING	Undrafted vs. drafted			1149	6				
GOOD_GOALTENDING	Undrafted vs. drafted			1139	6				
GOOD_NASTINESS	Undrafted vs. drafted			112%					
GOOD_DEFENSIVE_SKILLS	Undrafted vs. drafted			110%					
POOR_GOALTENDING	Undrafted vs. drafted			110%					
POOR_NET_DRIVE	Undrafted vs. drafted			108%					
POOR_SKATING	Undrafted vs. drafted								
GOOD_SHOT_BLOCKING	Undrafted vs. drafted			101%				1 null	
		0%	50%	100%	150% Value	200%	250%	300%	

Undrafted players are better at...

Undrafted - good skills		Ľ	Undrafted - poor skills			
Skill			Skill			
GOOD_PU Skill ATTLES	Undrafted vs. drafted	136%	POOR_VISION	Undrafted vs. drafted	291%	
GOOD_EFFORT	Undrafted vs. drafted	128%	POOR_PUCK_SKILLS	Undrafted vs. drafted	246%	
GOOD_NET_DRIVE	Undrafted vs. drafted	122%	POOR_PASSING	Undrafted vs. drafted	216%	
GOOD_PHYSICALITY_AND_SIZE	Undrafted vs. drafted	116%	POOR_SHOOTING_AND_SCORING	Undrafted vs. drafted	182%	
GOOD_POWER_PLAY	Undrafted vs. drafted	115%	POOR_PUCK_BATTLES	Undrafted vs. drafted	164%	
GOOD_PENALTY_KILLING	Undrafted vs. drafted	114%	POOR_PHYSICALITY_AND_SIZE	Undrafted vs. drafted	142%	
GOOD_GOALTENDING	Undrafted vs. drafted	113%	POOR_ACCURACY	Undrafted vs. drafted	129%	
GOOD_NASTINESS	Undrafted vs. drafted	112%	POOR_DEFENSIVE_SKILLS	Undrafted vs. drafted	116%	
GOOD_DEFENSIVE_SKILLS	Undrafted vs. drafted	110%	POOR_INTANGIBLES	Undrafted vs. drafted	116%	
GOOD_SHOT_BLOCKING	Undrafted vs. drafted	101%	POOR_DECISION_MAKING	Undrafted vs. drafted	114%	
GOOD_PUCK_PROTECTION	Undrafted vs. drafted	96%	POOR_GOALTENDING	Undrafted vs. drafted	110%	
GOOD_POISE	Undrafted vs. drafted	96%	POOR_NET_DRIVE	Undrafted vs. drafted	108%	
GOOD_TRANSITION_AND_RUSH	Undrafted vs. drafted	94%	POOR_SKATING	Undrafted vs. drafted	103%	
GOOD_PASSING	Undrafted vs. drafted	89%	POOR_PUCK_PROTECTION	Undrafted vs. drafted	86%	
GOOD_DECISION_MAKING	Undrafted vs. drafted	85%	POOR_EFFORT	Undrafted vs. drafted	72%	
GOOD_PUCK_SKILLS	Undrafted vs. drafted	82%	POOR_NASTINESS	Undrafted vs. drafted	69%	
GOOD_VISION	Undrafted vs. drafted	80%	POOR_TRANSITION_AND_RUSH	Undrafted vs. drafted	52%	
GOOD_SKATING	Undrafted vs. drafted	77%	POOR_POISE	Undrafted vs. drafted	1 null	
GOOD_SHOOTING_AND_SCORING	Undrafted vs. drafted	75%			0% 100% 200% 300%	
GOOD_RELEASE	Undrafted vs. drafted	65%			Value	
GOOD_ACCURACY	Undrafted vs. drafted	62%				
GOOD_INTANGIBLES	Undrafted vs. drafted	57%				
		0% 50% 100% 150% Value				

Topic extraction (on categories)-> common skillsets

👫 Topics 📴 Phrases 😺 Named Entities 😻 Misspellings & Unknowns								
Segr	mentation: by Paragraph 🔹 N	Io. topics: 8 🚔 Loading: 0.30 🛓	Run	🖹 🏅 reľ	R 🗊			S 🧰 💺 🖬 🖓
NO	NAME	KEYWORD5	EIGENVAL	% VAR	FREQ	CASES	% CASES	Lul. 🔄 🚾 Rate per 10,000 words 🔻 BY CODE
1	TOP-SIX FORWARD or TOP PAIRING DEFENSEMAN	GOOD_PUCK_SKILLS; GOOD_RELEASE; GOOD_SKATING; GOOD_VISION; POOR_PHYSICALITY_AND_SIZE; GOOD_ACCURACY	2.53	4.22	2416	1355	64.31%	
2	POWER FORWARD OR GOOD DEFENSEMAN	GOOD_NET_DRIVE; GOOD_PUCK_PROTECTION; GOOD_SHOOTING_AND_SCORING; GOOD_PUCK_BATTLES	1.41	3.56	1429	982	46.61%	
3	GOALTENDER	POOR_GOALTENDING; GOOD_GOALTENDING	1.29	3.43	123	103	4.89%	
4	GOOD ROLE PLAYER OR DEFENSIVE SPECIALIST	GOOD_PENALTY_KILLING; GOOD_DEFENSIVE_SKILLS; GOOD_INTANGIBLES; GOOD_PUCK_BATTLES; GOOD_EFFORT; GOOD_SHOT_BLOCKING; GOOD_DECISION_MAKING	1.24	3.73	2183	1322	62.74%	
5	FOURTH-LINE BRUISER OR PHYSICAL DEFENSEMAN	GOOD_PHYSICALITY_AND_SIZE; GOOD_NASTINESS; POOR_SKATING; POOR_SHOOTING_AND_SCORING; POOR_INTANGIBLES	1.21	3.36	1537	1088	51.64%	* 15
6	ROSTER FILLER	GOOD_TRANSITION_AND_RUSH; GOOD_PASSING; GOOD_POISE; POOR_DECISION_MAKING	1.17	3.44	1198	884	41.96%	
7	POWER PLAY SPECIALIST OR SECONDARY SCORER	GOOD_POWER_PLAY; POOR_EFFORT; POOR_TRANSITION_AND_RUSH; GOOD_SHOOTING_AND_SCORING	1.16	3.05	209	195	9.25%	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
8	NOT MUCH TO LOOK AT	POOR_PUCK_BATTLES; POOR_PASSING; POOR_PUCK_SKILLS; POOR_PHYSICALITY_AND_SIZE; POOR_ACCURACY	1.15	3.10	119	110	5.22%	4 / ₩₩0.00,00000000000000000000000000000000

Topic extraction yields familiar hockey archetypes

Predicting ESG/60

Set of predictors	Adjusted r ²	Average cross-validation mean squared error		
Actual NHL team picks	2.9%	0.0573		
Performance Only	12.1%	0.0521		
Scouting Only	23.7%	0.0495		
Performance & Scouting	25.0%	0.0445		

Results for predicting AHL even-strength goals per 60 minutes This model used CHL ESG/60, GOOD_RELEASE, GOOD_EFFORT, POWER_FORWARD (not all positive coefficients)

Predicting ESA1/60

Set of predictors	Adjusted r ²	Average cross-validation mean squared error
Actual NHL team picks	0.4%	0.0498
Performance Only	5.1%	0.0485
Scouting Only	14.8%	0.0450
Performance & Scouting	18.8%	0.0425

Results for predicting AHL even-strength primary assists per 60 minutes This model used CHL ESA1/60, GOOD_ACCURACY, POOR_INTANGIBLES, GOOD_RELEASE, GOOD_VISION, GOOD_DEFENSIVE_SKILLS, POOR_PUCK_SKILLS (not all positive coefficients)

Example – Alex DeBrincat

- Erie Otters, 5'7" forward
- 51-50-101 in 60 GP in 2015-16
- 65-62-127 in 63 GP in 2016-17
- Selected 2nd round, 39th overall by Chicago

Our rank	Actual draft rank (of 40 CHL forwards)	Player (overall selection)	Predicted AHL ES primary pts/60
1	14	Alex DeBrincat (39)	1.53
2	40	Ty Ronning (201)	1.51
3	39	Dmitry Sokolov (196)	1.50
4	22	Brandon Gignac (80)	1.50
5	3	Alexander Nylander (8)	1.50
6	1	Pierre-Luc Dubois (3)	1.40

• Ottawa Hockey Analytics conference 2017 – May 6, 2017

Conclusions

- 1. Draft analytics have outperformed team picks.
- 2. Surprisingly (?) scouting-based analytics performed better than stat-based analytics at predicting future (pro) scoring rates.
- 3. Using both together provide the best predictions.
- 4. Scouting remains vital for assessing skills BUT
- 5. Decision making on Draft Day needs to evolve to include draft analytics in a prominent role, including text-mined scouting data.

Ongoing work - version 2.0

- Projections for both forwards and defensemen
- Performance analytics for all four major junior leagues (OHL, QMJHL, WHL, USHL)
- Scouting-based projections for all prospects (European, all junior leagues, high school)
- Use both NHL and AHL target variables
- Add 2017 scouting reports to corpus
- Project future performance of 2017 draft class

Thanks!



This presentation was based on a white paper by Timo Seppa, Michael Schuckers, and Mike Rovito

http://statsportsconsulting.com/2017/02/22/text-mining-and-improving-nhl-draft-analytics/

• Ottawa Hockey Analytics conference 2017 – May 6, 2017