

Web Sports Data Extraction and Visualization

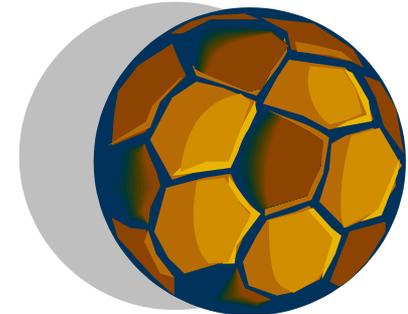
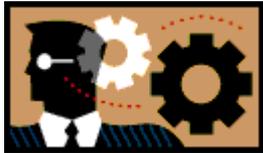


TECHNISCHE
UNIVERSITÄT
DARMSTADT

Seminar: Sports and Data Mining

Referent: Markus Schelmbauer

SS 2014



Agenda

- 1 Einleitung**
- 2 Web-Datenquellen
- 3 Datenextraktion
- 4 Zusammenfassung/ Ausblick

1 Einleitung

Wofür sind Web-Datenquellen nützlich?

In welcher Form liegen die Daten vor und haben sie sich geändert?

Was für Programme gibt es und wofür sind diese nützlich?

Agenda

1 Einleitung

2 Web-Datenquellen

3 Datenextraktion

4 Zusammenfassung/ Ausblick

2 Web-Datenquellen

Daten, die mit Sport in Zusammenhang stehen, sind leichter zugänglich und weiter verbreitet als noch vor ein paar Jahren

Anbieter von Web-Datenquellen sind:

- Verschiedene offizielle Dachverbände
- Weitere nicht offizielle Quellen

2 Web-Datenquellen

Finanzierung der Daten:

- Durch Abonnementen-Modelle
- Bei kostenlosen Anbietern: durch Werbung oder andere Einnahmequellen

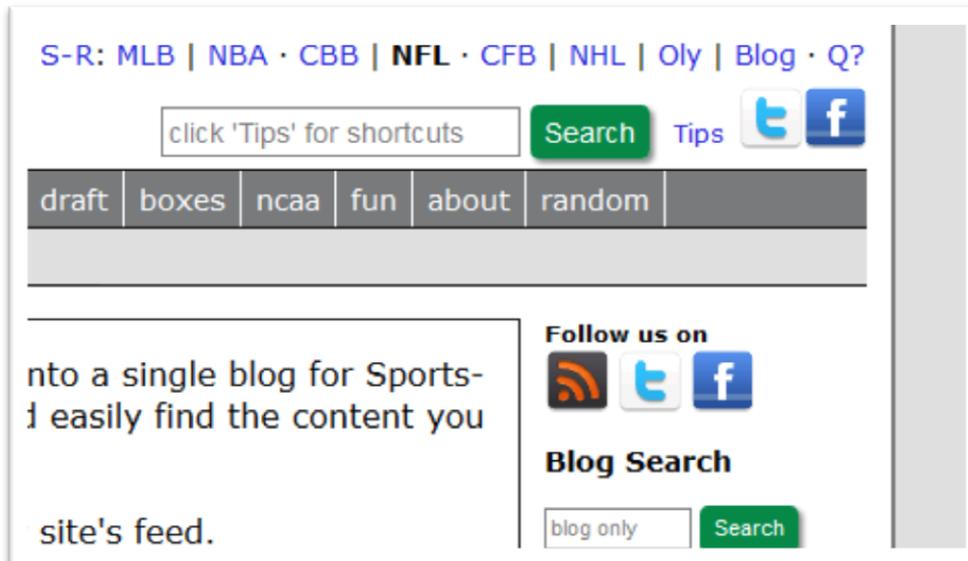
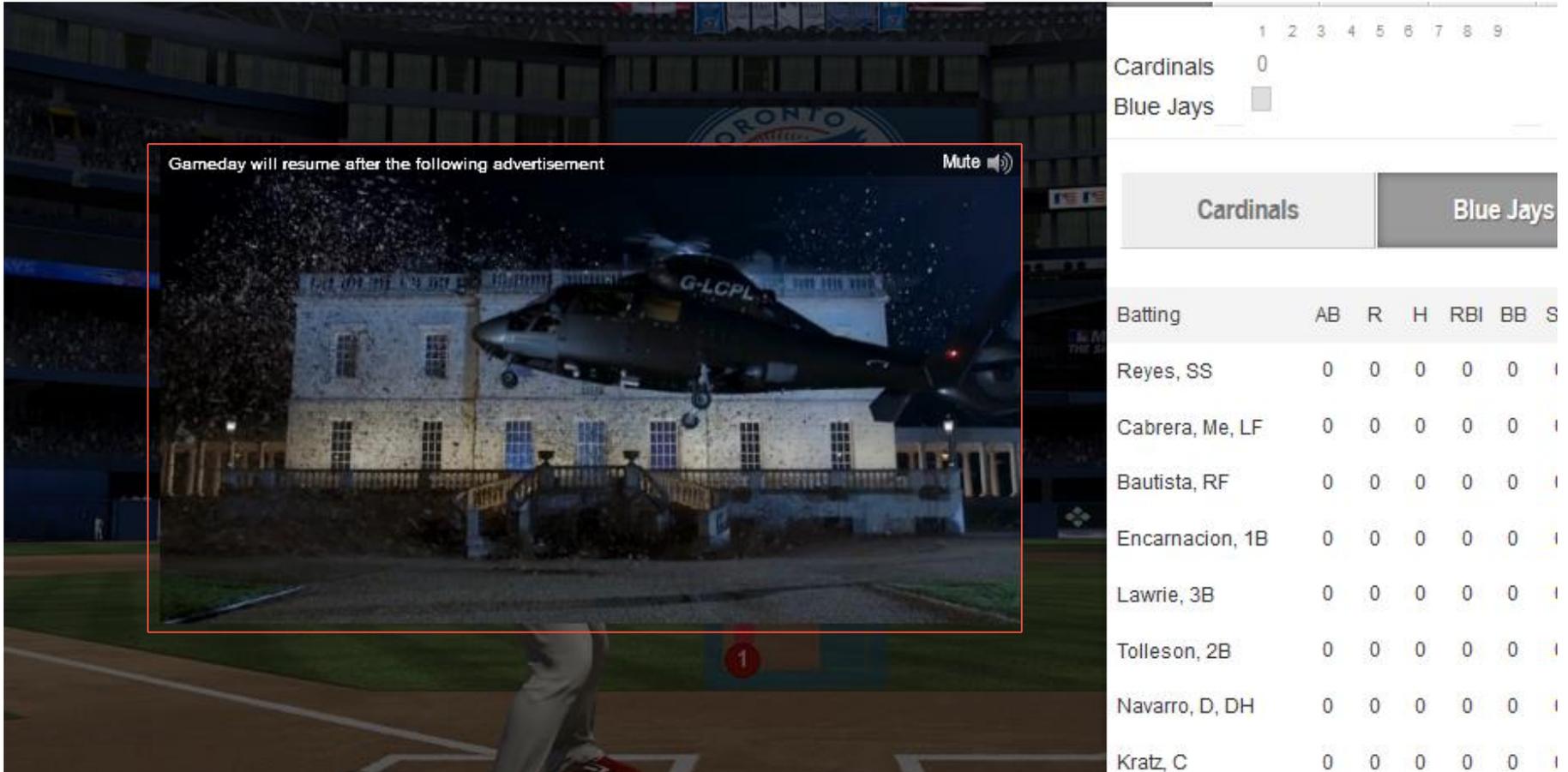


Abb.2



Abb.1

2 Web-Datenquellen Werbung



The image shows a screenshot of a live broadcast from a baseball stadium at night. In the foreground, a large advertisement for a helicopter is displayed. The helicopter is dark with the registration number 'G-LCPL' on its side. It is positioned in front of a large, illuminated building. The text 'Gameday will resume after the following advertisement' is overlaid on the top left of the ad, and a 'Mute' button is on the top right. A red circle with the number '1' is visible at the bottom center of the ad.

To the right of the ad, a scoreboard is visible. It shows the following statistics:

Batting	AB	R	H	RBI	BB	S
Reyes, SS	0	0	0	0	0	I
Cabrera, Me, LF	0	0	0	0	0	I
Bautista, RF	0	0	0	0	0	I
Encarnacion, 1B	0	0	0	0	0	I
Lawrie, 3B	0	0	0	0	0	I
Tolleson, 2B	0	0	0	0	0	I
Navarro, D, DH	0	0	0	0	0	I
Kratz, C	0	0	0	0	0	I

At the top of the scoreboard, the score is shown as Cardinals 0 and Blue Jays 0. The top of the scoreboard also has a row of numbers 1 through 9, likely representing innings.

Abb.7

2 Web-Datenquellen Abbildungen

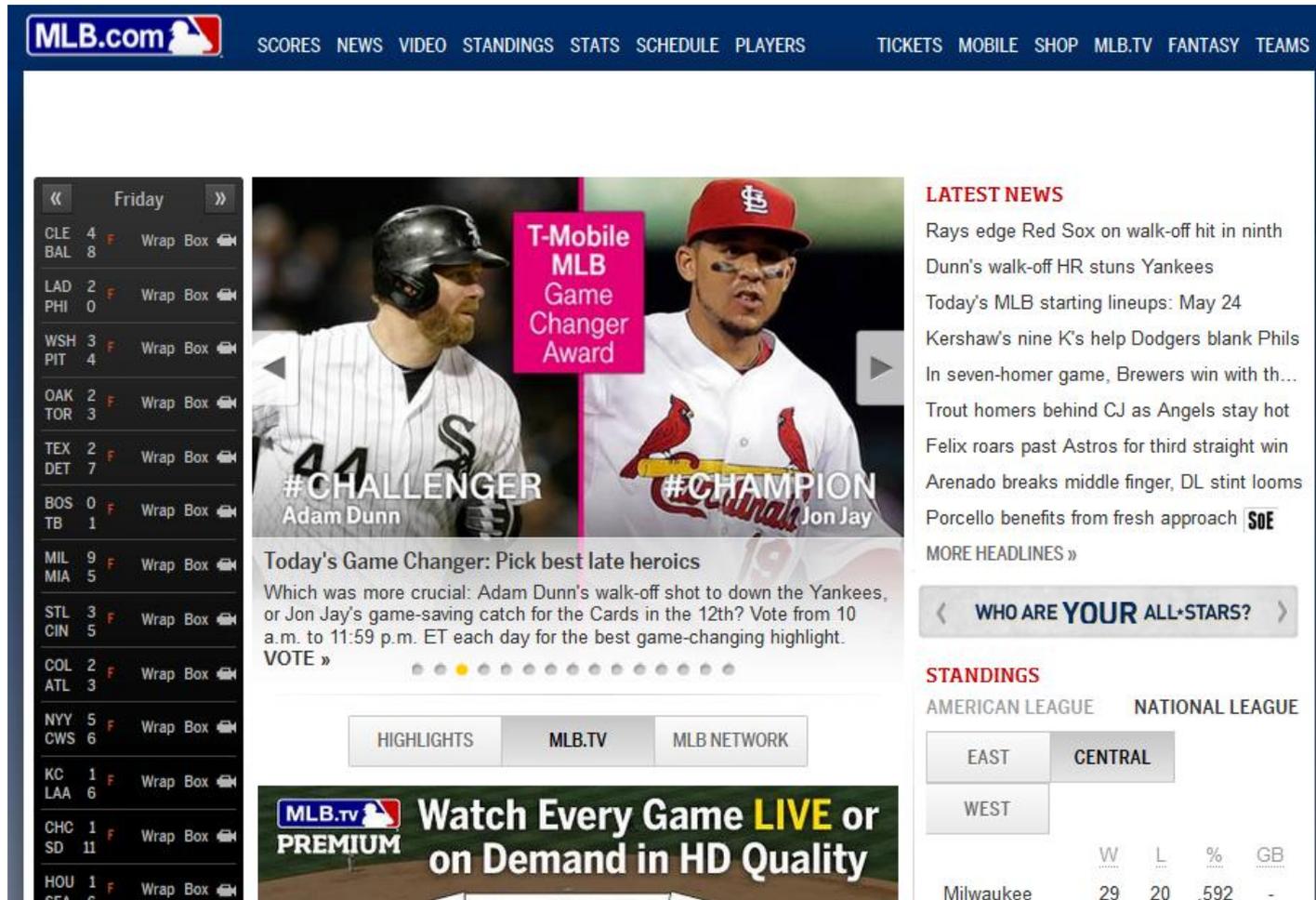
Die Daten werden auf verschiedenen Internetseiten in leicht verständlichen Diagrammen, Grafiken und Projektionen angeboten.



Abb.3

2.1 Baseball MLB Major League Baseball

MLB.com



MLB.com SCORES NEWS VIDEO STANDINGS STATS SCHEDULE PLAYERS TICKETS MOBILE SHOP MLB.TV FANTASY TEAMS

Friday

CLE	4	F	Wrap Box
BAL	8		
LAD	2	F	Wrap Box
PHI	0		
WSH	3	F	Wrap Box
PIT	4		
OAK	2	F	Wrap Box
TOR	3		
TEX	2	F	Wrap Box
DET	7		
BOS	0	F	Wrap Box
TB	1		
MIL	9	F	Wrap Box
MIA	5		
STL	3	F	Wrap Box
CIN	5		
COL	2	F	Wrap Box
ATL	3		
NYY	5	F	Wrap Box
CWS	6		
KC	1	F	Wrap Box
LAA	6		
CHC	1	F	Wrap Box
SD	11		
HOU	1	F	Wrap Box
SEA	6		

T-Mobile MLB Game Changer Award

#CHALLENGER Adam Dunn

#CHAMPION Jon Jay

Today's Game Changer: Pick best late heroics
Which was more crucial: Adam Dunn's walk-off shot to down the Yankees, or Jon Jay's game-saving catch for the Cards in the 12th? Vote from 10 a.m. to 11:59 p.m. ET each day for the best game-changing highlight. **VOTE »**

HIGHLIGHTS MLB.TV MLB NETWORK

MLB.TV PREMIUM Watch Every Game **LIVE** or on Demand in HD Quality

LATEST NEWS

- Rays edge Red Sox on walk-off hit in ninth
- Dunn's walk-off HR stuns Yankees
- Today's MLB starting lineups: May 24
- Kershaw's nine K's help Dodgers blank Phils
- In seven-homer game, Brewers win with th...
- Trout homers behind CJ as Angels stay hot
- Felix roars past Astros for third straight win
- Arenado breaks middle finger, DL stint looms
- Porcello benefits from fresh approach **SOE**

MORE HEADLINES »

WHO ARE YOUR ALL-STARS? >

STANDINGS

AMERICAN LEAGUE NATIONAL LEAGUE

	EAST	CENTRAL		
			WEST	
				W L % GB
Milwaukee				29 20 .592 -

Abb.4

2.1 Baseball MLB.com

Es gibt viele Daten, Grafiken und Darstellungen von Spielerleistungen sowie Grundlagenstatistiken, die kostenlos zur Verfügung gestellt werden.

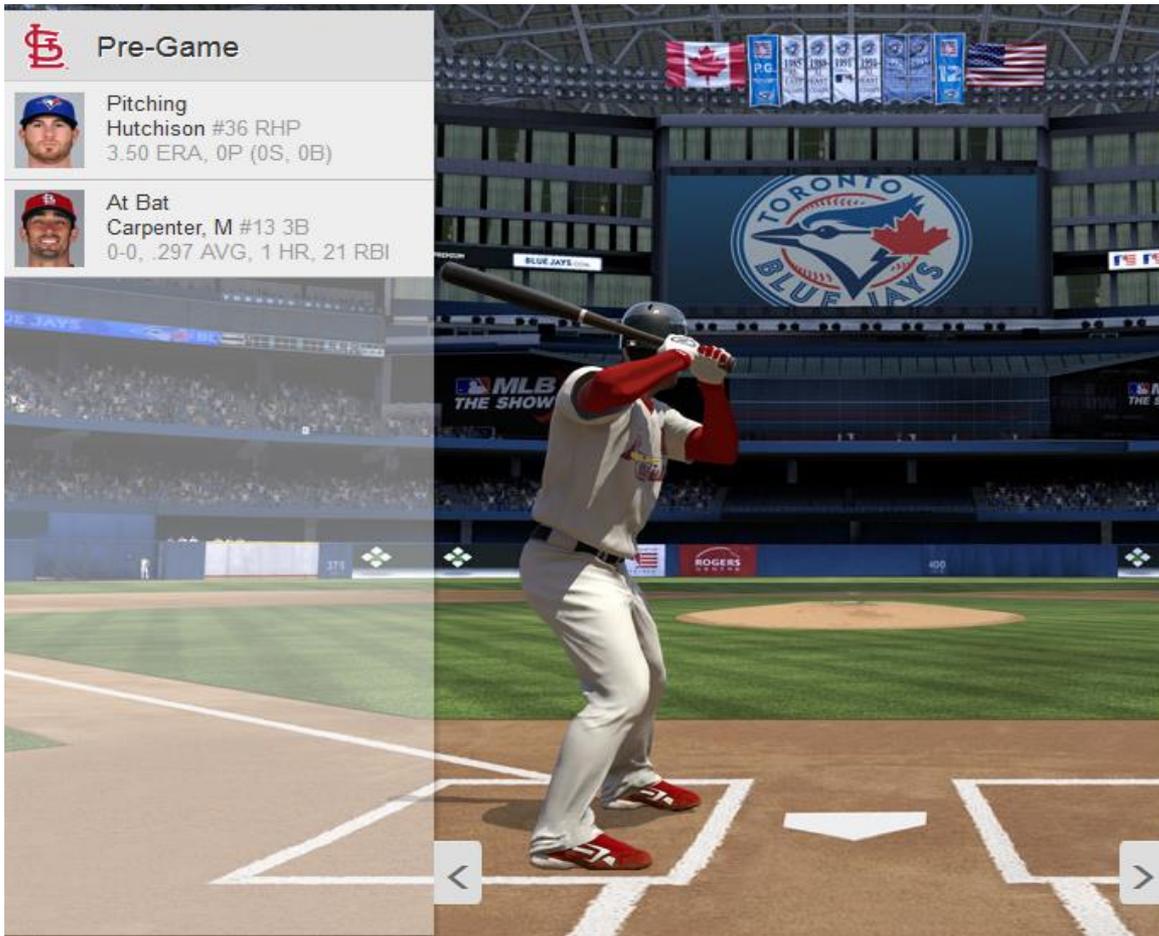
- Bsp. aus der Saison 2009: Cleveland's Shin-Soo Choo hat eine 0.405 Trefferquote gegen Kansas City im Kaufmann Stadion
- Statistiken von Schläger-Werfer-Paaren/Kombinationen
- [Bsp. PDF RED SOX Notes](#)

2.1 Baseball MLB.com

Pre-Game

Pitching
Hutchison #36 RHP
3.50 ERA, 0P (0S, 0B)

At Bat
Carpenter, M #13 3B
0-0, .297 AVG, 1 HR, 21 RBI





Jose Reyes

#7 | SS Throws: R Bats: S
Born: Jun 11, 1983 | 6'0" 195lbs

SUMMARY		STATS		SPLITS	
GAMES	HR	RBI	AVG	OPS	
46	3	17	.259	.725	

JUNE 8, 2014

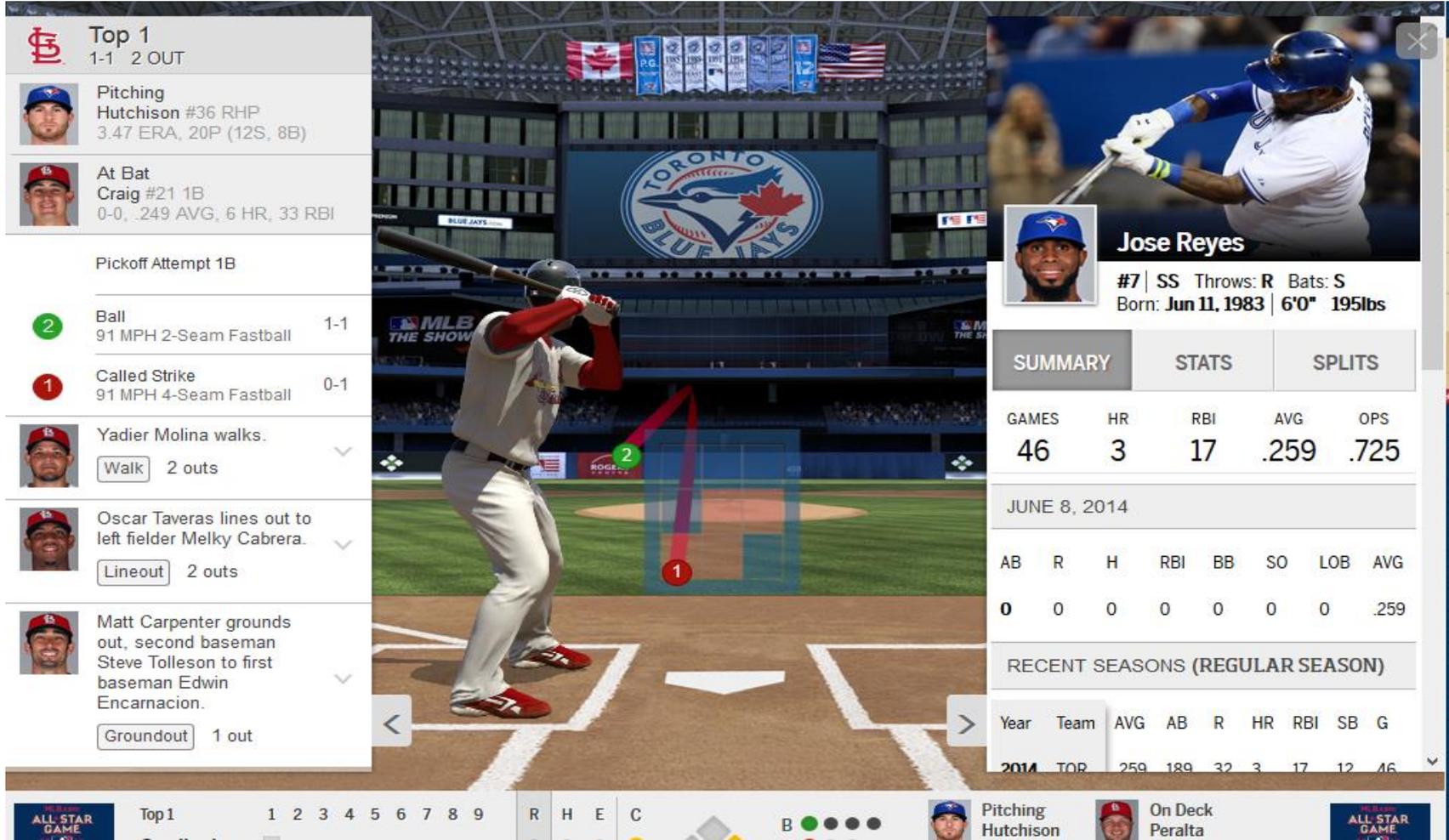
AB	R	H	RBI	BB	SO	LOB	AVG
0	0	0	0	0	0	0	.259

RECENT SEASONS (REGULAR SEASON)

Year	Team	AVG	AB	R	HR	RBI	SB	G
2014	TOR	.259	189	32	3	17	12	46

Abb.5

2.1 Baseball MLB.com



Top 1
1-1 2 OUT

Pitching
Hutchison #36 RHP
3.47 ERA, 20P (12S, 8B)

At Bat
Craig #21 1B
0-0, .249 AVG, 6 HR, 33 RBI

Pickoff Attempt 1B

2 Ball
91 MPH 2-Seam Fastball 1-1

1 Called Strike
91 MPH 4-Seam Fastball 0-1

Yadier Molina walks.
Walk 2 outs

Oscar Taveras lines out to left fielder Melky Cabrera.
Lineout 2 outs

Matt Carpenter grounds out, second baseman Steve Tolleson to first baseman Edwin Encarnacion.
Groundout 1 out

Jose Reyes
#7 | SS Throws: R Bats: S
Born: Jun 11, 1983 | 6'0" 195lbs

SUMMARY		STATS		SPLITS	
GAMES	HR	RBI	AVG	OPS	
46	3	17	.259	.725	

JUNE 8, 2014

AB	R	H	RBI	BB	SO	LOB	AVG
0	0	0	0	0	0	0	.259

RECENT SEASONS (REGULAR SEASON)

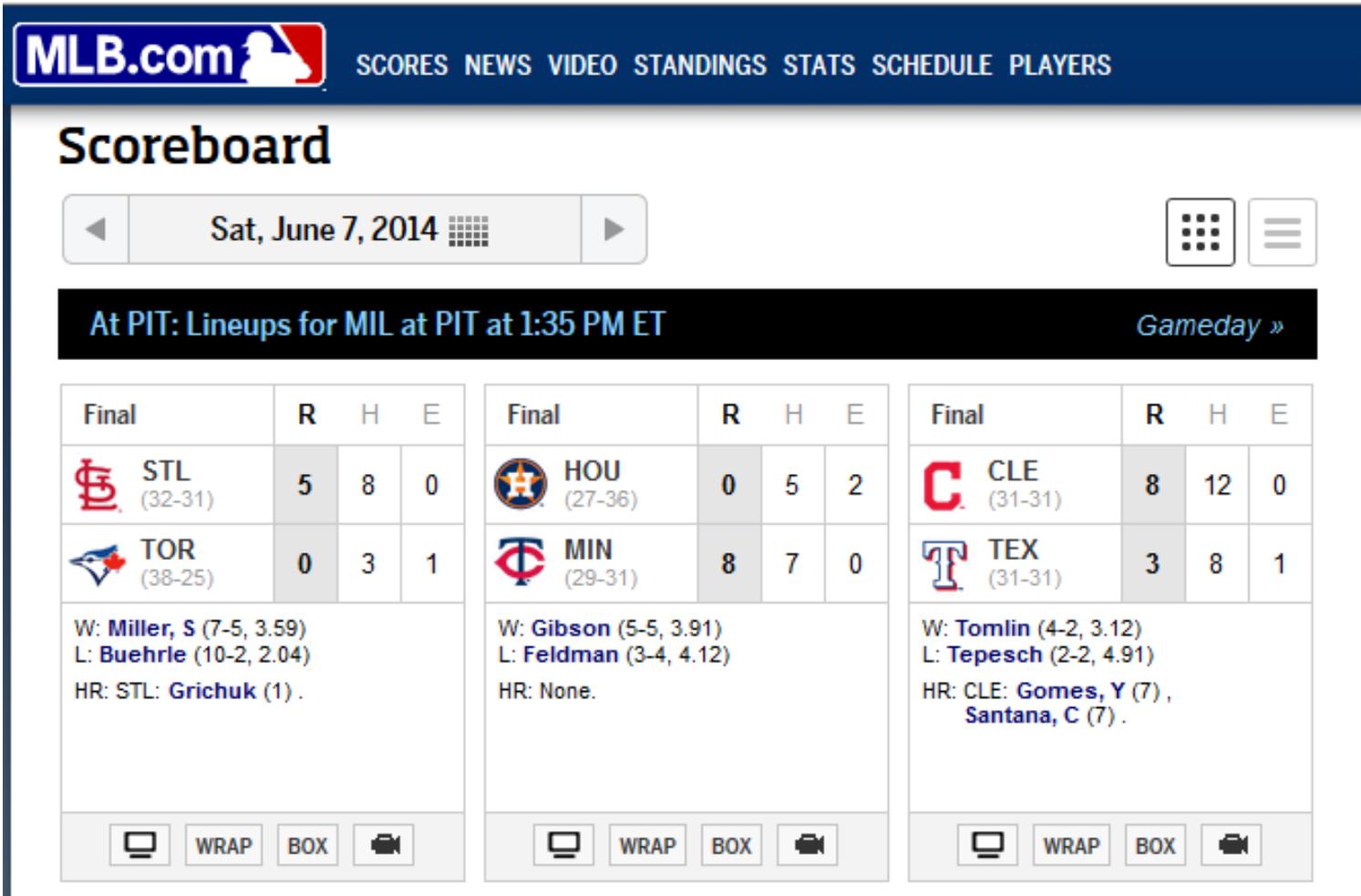
Year	Team	AVG	AB	R	HR	RBI	SB	G
2014	TOR	.259	189	32	3	17	12	46

Top 1 1 2 3 4 5 6 7 8 9 R H E C B ●●●●●

Pitching Hutchison On Deck Peralta

Abb.5

2.1 Baseball MLB Premium



MLB.com SCORES NEWS VIDEO STANDINGS STATS SCHEDULE PLAYERS

Scoreboard

Sat, June 7, 2014

At PIT: Lineups for MIL at PIT at 1:35 PM ET *Gameday »*

Final	R	H	E
 STL (32-31)	5	8	0
 TOR (38-25)	0	3	1
W: Miller, S (7-5, 3.59) L: Buehrle (10-2, 2.04) HR: STL: Grichuk (1) .			

Final	R	H	E
 HOU (27-36)	0	5	2
 MIN (29-31)	8	7	0
W: Gibson (5-5, 3.91) L: Feldman (3-4, 4.12) HR: None.			

Final	R	H	E
 CLE (31-31)	8	12	0
 TEX (31-31)	3	8	1
W: Tomlin (4-2, 3.12) L: Tepesch (2-2, 4.91) HR: CLE: Gomes, Y (7) , Santana, C (7) .			

Abb.6

2.1 Baseball MLB Premium



SCORES NEWS VIDEO STANDINGS STATS SCHEDULE PLAYERS

Scoreboard

Mon, June 9, 2014

At PIT: Middle 7th - MIL 1, PIT 0 Gameday »

1:10 PM ET	R	H	E	7:05 PM ET	R	H	E	7:05 PM ET	R	H	E
 SEA (32-29)				 BOS (27-34)				 CHC (25-34)			
 TB (24-39)				 BAL (31-29)				 PIT (29-32)			
Probables: SEA: Ramirez, Er (1-4, 6.82) TB: Price (4-5, 4.03)			Probables: BOS: Peavy (1-3, 4.72) BAL: Norris, B (4-5, 4.41)			Probables: CHC: Jackson, E (4-5, 4.59) PIT: Morton (2-7, 3.31)					
PREVIEW	NOTES			PREVIEW	NOTES			PREVIEW	NOTES		

Abb.7

2.1 Baseball MLB Premium

Rot = hohe Trefferquote

Blau = niedrige Trefferquote

Nummerierung = Anzahl der pitches

Spieltages Tool

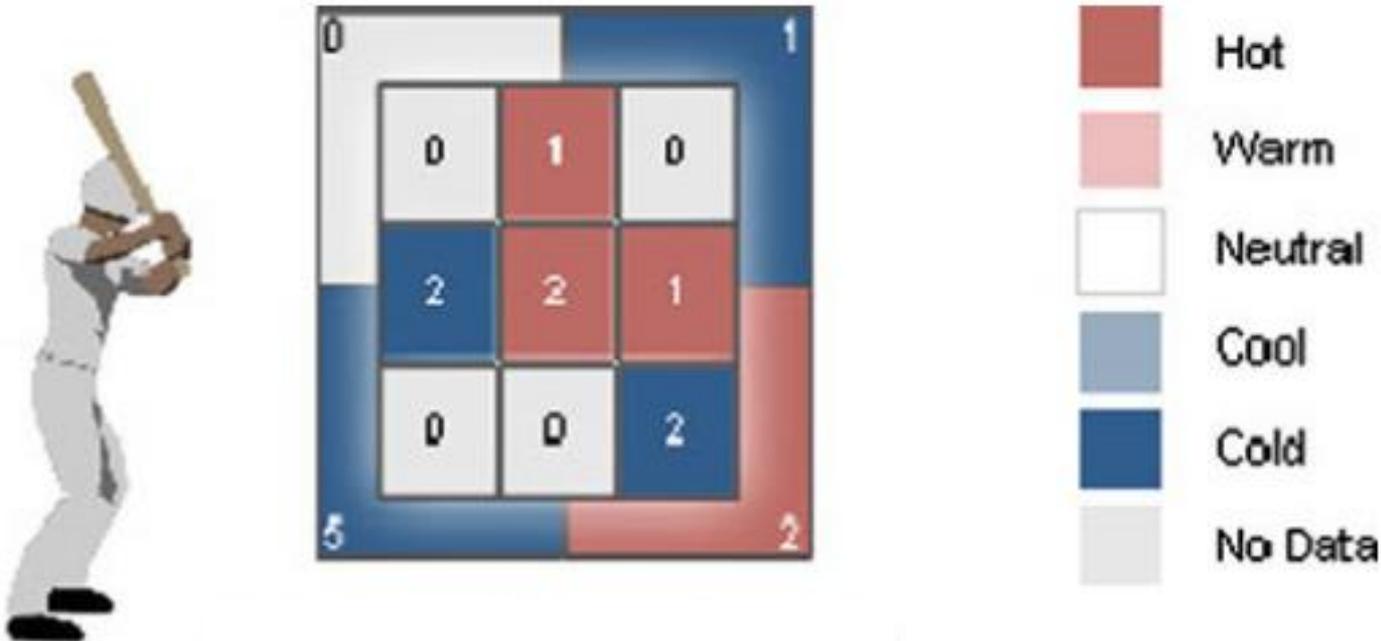


Abb.8

2.1 Baseball MLB Premium

Pitcher Tendencies Tool

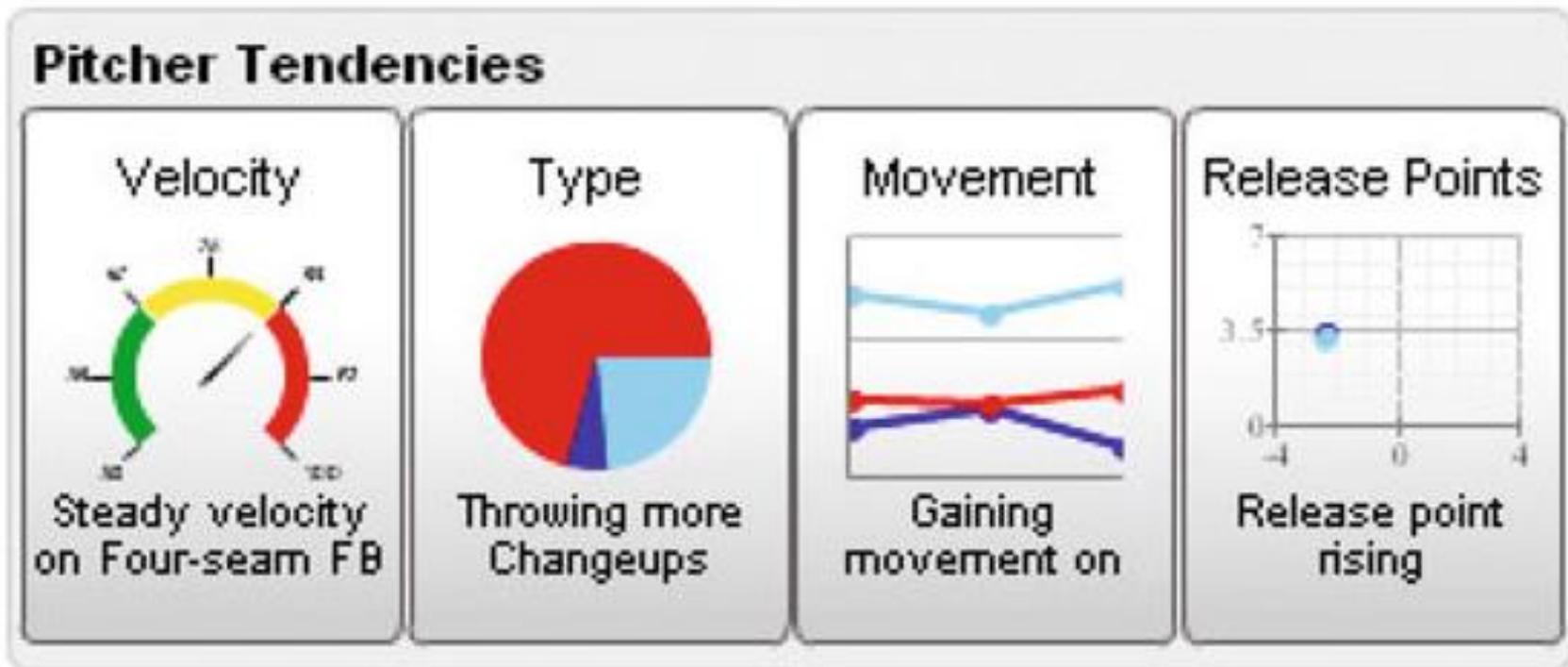


Abb.9

2.1 Baseball Retrosheet.org



Historische und kontinuierlichen Daten seit 1952

RETROSHEET

Home	About ↓	Games/People/Parks ↓	Data downloads ↓	Features ↓	Organization ↓	Archives ↓
------	---------	----------------------	------------------	------------	----------------	------------

Keeping Score
Retrosheet was founded in 1989 for the purpose of computerizing play-by-play accounts of as many pre-1984 major league games as possible. [Learn more about the organization.](#)

[Site Map](#)

[Discussion of discrepancies with official records](#)

Retrosheet's Most Wanted
Not criminals, but games. Lists by year of all games we are missing from 1920-73 [Can you help?](#)

What's New

- 6/2/2014: Homer replay in Boston. See [\[Noteworthy events/Homer Replays\]](#)
- 5/25/2014: Josh Beckett crafts the first no hitter of the season. See [\[No Hitters and Cycles\]](#)

Abb. 10

2.1 Baseball Weitere Internetseiten

- Baseball-reference.com
 - Mit historischen und aktuellen Daten zusätzlich Minor-League-Teams (unterklassige Ligen)
- Baseball1.com (Sean Lehman-Baseball-Archiv)

2.1 Basketball NBA.com

NBA.com

The screenshot shows the NBA.com website interface. At the top, there is a blue header with the SPOX.com logo and the NBA DEUTSCHLAND logo. To the right of the logos, there are social media links for Facebook (Gefällt mir 161.647) and Twitter (@NBA_de folgen 5.593 Follower). Below the header is a navigation menu with categories: HOME, FUSSBALL, NBA, FORMEL 1, BASKETBALL, BOXEN, TENNIS, GOLF, EISHOCKEY, HANDBALL, US-SPORT, MEHR, VIDEO, and MYSPOX. Below the navigation menu is a sub-menu for NBA LEAGUE PASS with options: Videos, Playoffs, Ergebnisse, Tabellen, Stats, Spieler & Teams, News, Fotos, LIVE-STREAM, and D-League. The main content area features a table of NBA League Pass games for the week of May 24th to 30th. The table has columns for the days and rows for the games. The games listed are: Sa, Mai 24 (IND vs MIA, 2:30 AM, NBA LEAGUE PASS Sport 1 US), So, Mai 25 (SAS vs OKC, 2:30 AM, NBA LEAGUE PASS Spox.com), Mo, Mai 26 (IND vs MIA, 2:30 AM, NBA LEAGUE PASS Sport 1 US), Di, Mai 27 (SAS vs OKC, 3:00 AM, NBA LEAGUE PASS Sport 1 US), Mi, Mai 28 (MIA vs IND, 2:30 AM, NBA LEAGUE PASS Sport 1 US), Do, Mai 29 (OKC vs SAS, 3:00 AM, NBA LEAGUE PASS), and Fr, Mai 30 (IND vs MIA, 2:30 AM, NBA LEAGUE PASS). Below the table is a large image of a basketball game between the Thunder and the Spurs. The image shows a player from the Spurs (Tony Parker) dribbling the ball while being defended by a player from the Thunder (Russell Westbrook). Below the image is a live-stream article titled "Ein ungleiches Duell" with a sub-headline "LIVE-STREAM Tony Parker vs. Russell Westbrook. Das Point-Guard-Duell fiel bisher recht einseitig zu Gunsten des Franzosen aus. Kann Russ in dieser Serie noch auftrumpfen? »". To the right of the image is a section titled "NBA-NEWS" with a list of news items: 08:15 Spiel 3: George darf spielen, 23.05. Blitzheilung bei OKC's Ibaka, 23.05. George im Training - Einsatz offen, 22.05. Carter-Williams führt All-Rookie First Team an, 22.05. Nach dem Blowout: Ein Nachbericht in 35 Punkten, 22.05. Barkley: LeBron sollte nach Cleveland zurückkehren, 22.05. Überragende Spurs zerlegen OKC, and 21.05. George erleidet Gehirnerschütterung. Below the news section is a banner for NBA LEAGUE PASS with the text "SEHEN SIE JEDES SPIEL DER PLAYOFFS" and the PLAYOFFS logo.

NBA LEAGUE PASS	Sa, Mai 24	So, Mai 25	Mo, Mai 26	Di, Mai 27	Mi, Mai 28	Do, Mai 29	Fr, Mai 30
	IND vs MIA	SAS vs OKC	IND vs MIA	SAS vs OKC	MIA vs IND	OKC vs SAS	IND vs MIA
JETZT ANSEHEN	2:30 AM NBA LEAGUE PASS Sport 1 US	2:30 AM NBA LEAGUE PASS Spox.com	2:30 AM NBA LEAGUE PASS Sport 1 US	3:00 AM NBA LEAGUE PASS Sport 1 US	2:30 AM NBA LEAGUE PASS Sport 1 US	3:00 AM NBA LEAGUE PASS	2:30 AM NBA LEAGUE PASS

LIVE-STREAM: Thunder vs. Spurs (Mo., 2.30 Uhr)

Ein ungleiches Duell

LIVE-STREAM Tony Parker vs. Russell Westbrook. Das Point-Guard-Duell fiel bisher recht einseitig zu Gunsten des Franzosen aus. Kann Russ in dieser Serie noch auftrumpfen? »

NBA-NEWS

- 08:15 Spiel 3: George darf spielen
- 23.05. Blitzheilung bei OKC's Ibaka
- 23.05. George im Training - Einsatz offen
- 22.05. Carter-Williams führt All-Rookie First Team an
- 22.05. Nach dem Blowout: Ein Nachbericht in 35 Punkten
- 22.05. Barkley: LeBron sollte nach Cleveland zurückkehren
- 22.05. Überragende Spurs zerlegen OKC
- 21.05. George erleidet Gehirnerschütterung

NBA LEAGUE PASS
SEHEN SIE JEDES SPIEL DER PLAYOFFS

Abb. 11

2.1 Basketball NBA.com

NBA.com

The screenshot displays the NBA.com website interface in German. At the top, there is a blue navigation bar with the NBA logo, a search bar, and various utility links like 'Login', 'Gefällt mir', 'Follow @NBA', and 'NBA STORE'. Below this is a secondary navigation bar with links for 'Tickets', 'Teams', 'Scores & Schedules', 'News', 'Blogs', 'Video', 'Players', 'Standings', 'Stats', 'Mobile', 'TV', 'League Pass', and 'More Store'. A search bar is also present with the text 'Search NBA.com'. The main content area features a grey banner for 'NBA DEUTSCHLAND' with a German flag icon. Below this, there is a blue banner with the text 'Tonight on NBA TV: 'The84Draft' (9 ET) | Film review | 84 facts on the '84 Draft | 2014 Finals'. The main content area is divided into several sections: a large image of three Miami Heat players in white jerseys with 'HEAT' written on them, titled 'Morning Tip: Sticks And Stones ...'; a video thumbnail titled 'Mini-Movie: Nothing Easy'; and another video thumbnail titled 'GameTime: Oh Yeah, That Guy'. At the bottom left, there is a link to 'Create a custom NBA video channel' and a 'Connect with Facebook' button.

Abb. 11

2.1 Basketball NBA.com



Welcher Nutzen kann aus diesen Seiten gezogen werden?

- Informationen über die Leistung der Spieler
- Identifiziert erfolgreichste fünf Kombinationen

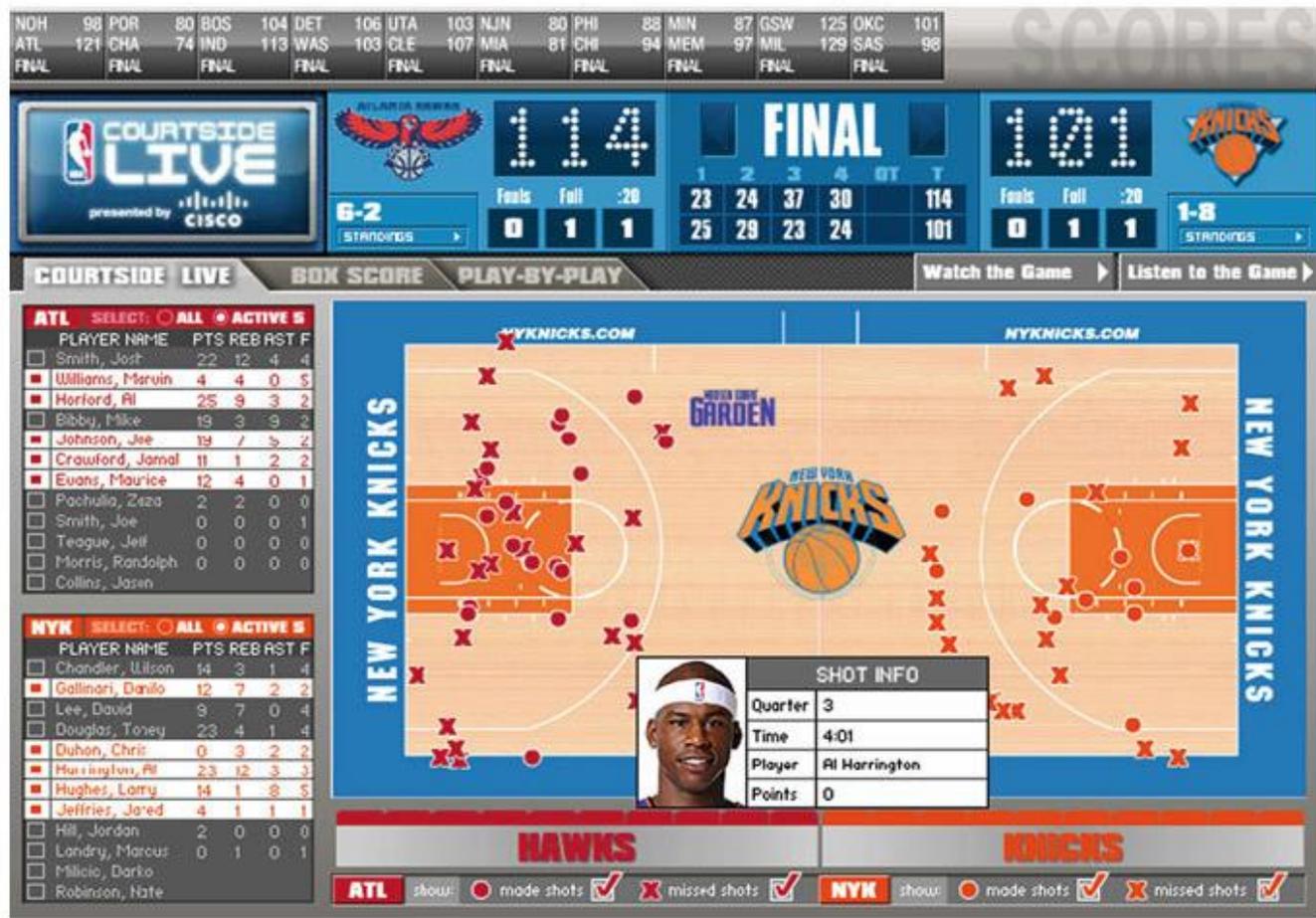
Top Five-Player Combinations

PLAYER 1	PLAYER 2	PLAYER 3	PLAYER 4	PLAYER 5	TEAM	+	-	+/-
J. Kidd	E. Dampier	D. Nowitzki	S. Marion	J. Terry	Mavericks	172	-105	67
M. Bibby	J. Crawford	J. Johnson	Jo. Smith	A. Horford	Hawks	155	-100	55
J. O'Neal	Q. Richardson	D. Wade	M. Beasley	M. Chalmers	Heat	293	-245	48
K. Garnett	R. Allen	P. Pierce	K. Perkins	R. Rondo	Celtics	490	-443	47
M. Camby	B. Davis	R. Butler	C. Kaman	E. Gordon	Clippers	178	-137	41

Abb. 12

2.1 Basketball NBA.com

Courtside Live



The screenshot displays the NBA.com Courtside Live interface for a game between the Atlanta Hawks (ATL) and the New York Knicks (NYK). The interface is divided into several sections:

- Top Scoreboard:** Shows the final score of 114 for the Hawks and 101 for the Knicks. The game is in the 3rd quarter. The Hawks' record is 6-2, and the Knicks' record is 1-8.
- Quarterly Scores:**

Q	1	2	3	4	OT	T
Hawks	23	24	37	30		114
Knicks	25	29	23	24		101
- Player Statistics:**

Team	Player Name	PTS	REB	AST	F
ATL	Smith, Jost	22	12	4	4
	Williams, Marvin	4	4	0	5
	Hortford, Al	25	9	3	2
	Bibby, Mike	19	3	9	2
	Johnson, Jee	19	7	5	2
	Crawford, Jamal	11	1	2	2
NYK	Chandler, Wilson	14	3	1	4
	Gallinari, Danilo	12	7	2	2
	Lee, David	9	7	0	4
	Douglas, Tony	23	4	1	4
	Duhon, Chris	0	3	2	2
	Harrington, Al	23	12	3	3
	Hughes, Lorry	14	1	8	5
	Jeffries, Jared	4	1	1	1
	Hill, Jordan	2	0	0	0
	Londry, Marcus	0	1	0	1
- Court Diagram:** A live court diagram showing player positions. A shot info window for Al Harrington is open, showing he is in the 3rd quarter with 4:01 remaining and has 0 points.
- Legend:** ATL shows made shots (red circle) and missed shots (red X). NYK shows made shots (orange circle) and missed shots (orange X).

Abb. 13

2.1 Basketball NBA.com



“Stats at a Glance“

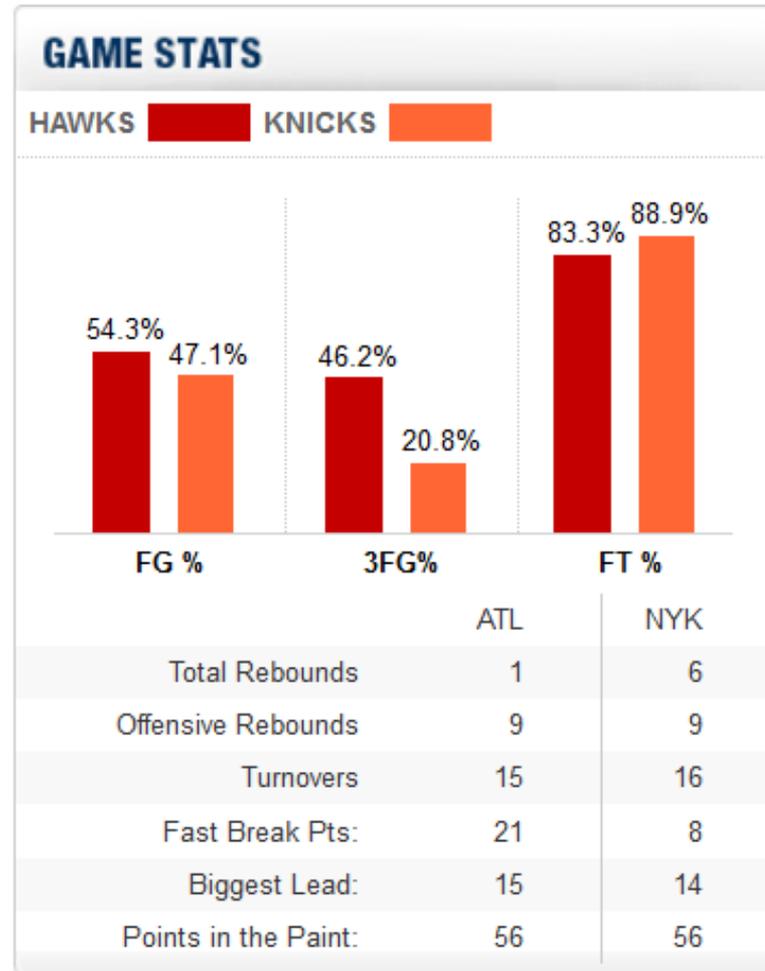


Abb. 14

2.1 Basketball

Verfügt, ebenso wie Baseball, über viele Online-Datenquellen

2.1 Basketball Basketball-reference.com

- Vergleichbar mit Baseball-reference.com
- Die Seite gibt es seit 2003



The screenshot shows the Basketball-Reference.com website. At the top, there is a navigation bar with links for various sports: S-R: [MLB](#) | [NBA](#) · [CBB](#) | [NFL](#) · [CFB](#) | [NHL](#) | [Oly](#) | [Blog](#) · [Q?](#). There is a search bar and social media icons for Twitter and Facebook. Below the navigation bar, there is a menu with links for [play index](#), [box scores](#), [players](#), [teams](#), [seasons](#), [coaches](#), [leaders](#), [awards](#), [playoffs](#), [draft](#), [olympics](#), and [more \[+\]](#). The main content area is titled "2013-14 NBA Playoffs" and features a "Finals" section for the game between San Antonio (1) and Miami (2). The text indicates that San Antonio leads the series 1-0, with the next game on Sunday, June 8th at 8:00 PM EST, broadcasted on ABC. The series win probability is SAS (65.6%) / MIA (34.4%), and the Finals win probability is SAS (65.6%) / MIA (34.4%). The season series was tied 1-1, with a link to [Playoff history SAS vs. MIA](#). On the right side, there is a "Like" button with 2,976 likes and a "g+1" button with 209 recommendations. Below this is a "Yesterday's Top Performers" section, which currently shows "No games played" and a link to [See all of yesterday's players](#). At the bottom, there are links for [2013-14 NBA Debuts](#) and [2013-14 NBA All-Stars](#). On the left side, there is a "Basketball-Reference.com News" section with several news items, including "DNP/Inactive Records For Player Gamelogs", "New Team and Opponent Statistics Tables", "Shooting Hand Added to Player Season Finder", and "Basketball-Reference.com is Ten Years Old".

Abb. 15

2.1 Cricket Cricinfo.com



The screenshot shows the ESPNcricinfo website interface. At the top, there is a navigation bar with categories: FOOTBALL, RUGBY, F1, BOXING, GOLF, TENNIS, MORE SPORTS, VIDEO, GAMES, TV. Below this is a message: "Now you can set your own homepage. Choose from the dropdown below." with a close button (X). The main header includes "CHOOSE EDITION UK", the ESPNcricinfo logo, a Facebook Like button (3.5m), a "Welcome Guest" message, and a search bar. A secondary navigation bar contains links for Series, Countries, Live Scores, Fixtures, Results, News, Features, Photos, Video & Audio, Blogs, Statistics, Archive, Fantasy, and Mobile. Below this is a "Quick Links" section with links for Eng v SL, County cricket 2014, WI v NZ, Ban v Ind, Cricket on Twitter, Betting Blog, ICC Rankings, and ESPNcricinfo apps.

The main content area is divided into two columns. The left column features a "VIDEO NEWS AND ANALYSIS" section with a large image for "The best night of NatWest T20 so far" and two smaller video thumbnails: "#politeenquiries: Will Broad investigate Moeen's Doosra?" and "County Cricket Show: Lyth and Lees flail Northants". The right column displays cricket news articles. The top article is "WI v NZ, 1st Test, Kingston, 1st day" with the headline "Latham and Williamson build slowly" and sub-headline "Gayle targets 'huge performance'". Below it is "Chris Gayle's 100th Test" with the headline "Gayle targets 'huge performance'". The next article is "Northants v Sri Lankans, Northampton, 4th day" with the headline "SL buoyant despite Keogh ton". The bottom article is "Sri Lanka tour of England 2014" with the headline "Senanayake to be tested in Cardiff".

On the right side of the page, there is a "Scores" section with tabs for "Scores", "Results", and "Fixtures". It lists "International" matches: "New Zealand 94/1 v West Indies", "Sri Lankans 558/8d & 99/4d v Northants 345" (Match drawn). Below that, it lists "Domestic" matches: "England Domestic", "Kent 387 v Essex 253/1", "Leics 311 v Derbyshire 12/0", "Sussex 210/9 v Somerset", "Lancashire 188/6 v Warwickshire", "Worcs 175/4 v Hampshire", "Notts 205 v Yorkshire 52/2", "Gloucs 207/3 (20/20 ov) v Glamorgan" (No result). At the bottom of the scores section are links for "Live scores pages", "Scores on Mobile", and "Live".

Abb. 16

2.1 Football NFL.com

The screenshot shows the NFL.com website interface. At the top, there is a navigation bar with the NFL logo, 'Sign In | Register', and a search bar. Below this is a menu with categories: FANTASY, NEWS, VIDEO, SCORES, SCHEDULES, STANDINGS, STATS, WATCH GAMES, TICKETS, and SHOP. A secondary bar includes 'WATCH NFL NETWORK NOW: AMERICA'S GAME 11:00 AM ET', 'NFL NETWORK', 'MOBILE', 'EN ESPAÑOL', and 'HEALTH & SAFETY'. The main content area features 'DRAFT RESULTS ROUND 1' with a list of picks: Jadeveon CLOWNEY (DE, Pick No. 1), Greg ROBINSON (OT, Pick No. 2), Blake BORTLES (QB, Pick No. 3), and Sammy WATKINS (WR, Pick No. 4). Below the draft results is a section titled 'TOP 10 BEST OFFSEASON TEAMS' with logos for the Chicago Bears, Jacksonville Jaguars, Tampa Bay Buccaneers, Minnesota Vikings, New York Jets, and New Orleans Saints. To the right, there is an article titled 'The TOP 100 Players of 2014' featuring a photo of Von Miller with a star rating of 76 and the sub-heading 'MILLER TIME'. The article text states: 'Von Miller dropped 67 spots from last year, but he's still right at home on the "Top 100 Players of 2014" list. Check out numbers 80-71 here.' Below the text is a list of bullet points: 'Watch: Talib says Miller should be Top 10', 'Welker's tough act has him at No. 73', 'The debate is on: Is Romo a top 10 QB?', and 'Are Tony Romo's best days behind him?'.

Abb. 17

2.1 Hockey NHL.com

The screenshot shows the NHL.com website interface. At the top, there are navigation links for various languages (ENGLISH, FRANÇAIS, РУССКИЙ, SUOMI, SVENSKA, CESTINA, SLOVENCINA, DEUTSCH) and social media icons. Below this is a schedule for the week of May 24-30, 2014, with games between Chicago and Los Angeles, and Montreal and NY Rangers. The main navigation bar includes links for SCORES, STANDINGS, SCHEDULE, STATS, PLAYERS, TEAMS, NEWS, VIDEO, GAMECENTER LIVE, NHL NETWORK, PLAYOFFS, FANTASY, TICKETS, and SHOP. The central content area features the Stanley Cup Playoffs 2014 logo and a game preview for Game 3 between the Chicago Blackhawks (CHI) and the Los Angeles Kings (LAK), which is tied 1-1. The preview includes the game date and time (SAT MAY 24, 8:00 PM EDT). To the right, there are sections for TOP HEADLINES and NHL GREEN, with a list of news items such as 'CHI-LAK: Teams know each other very well' and 'Hawks' Shaw ready to go | Game 3 preview'. A large video player in the center shows a hockey player in a blue jersey, with the headline 'STEPAN SIDELINED' and a sub-headline 'Rangers center sustains broken jaw, out indefinitely following surgery'. A 'READ MORE' link is provided below the headline. On the left side, there is a vertical banner for 'NHL GAMECENTER LIVE' with the text 'WATCH THE STANLEY CUP™ PLAYOFFS LIVE' and the URL 'NHL.COM/GCL'. The bottom left corner contains the copyright notice '© NHL 2014. All Rights Reserved.'

Abb. 18

2.1 Fußball MLSnet.com zu <http://www.mlssoccer.com/>



USA VS Red Bull NEW YORK

SUNDAY | 5 PM ET
GILLETTE STADIUM
MLS LIVE
#NEvNY

You can't stop the Revs. Can NY contain them?



LA's Arena: Of course we expect to win the SuperClasico, as do Chivas



USA-Nigeria HIGHLIGHTS: All the action & analysis from MLSsoccer.com



Larentowicz rails against himself & Fire teammates after another loss

6/7 FINAL	6/7 FINAL	6/7 FINAL	6/7 FINAL	6/7 FINAL	6/7 FINAL	6/7 FINAL	6/7 FINAL	6/7 FINAL	6/7 FINAL	6/7 FINAL	6/7 FINAL	6/7 FINAL	6/8 5:00pm	6/8 8:00pm					
TOR 1	SLV 0	USA 2	DC 0	PHI 3	CHI 2	DAL 3	RSL 1	NE 1	SJ 0	ESP 2	NGA 1	CLB 0	VAN 3	SEA 3	COL 2	POR 3	NY 3	LA 1	CHV 1
MLS	Road to Brazil	Friendly	MLS	MLS	MLS	MLS	MLS	MLS	MLS										
RECAP	RECAP	RECAP	RECAP	RECAP	RECAP	RECAP	RECAP	RECAP	RECAP	RECAP	RECAP	RECAP	MLS LIVE	UDN					

6/12 4:00pm	6/13 12:00pm	6/13 3:00pm	6/13 6:00pm	6/14 12:00pm	6/14 3:00pm	6/14 6:00pm	6/14 9:00pm	6/15 12:00pm	6/15 3:00pm
BRA 1	MEX 0	ESP 1	CHI 0	COL 0	URU 0	ENG 0	CIV 0	SUI 0	FRA 0
CRO 0	CMR 0	NED 0	AUS 0	GRE 0	CRC 0	ITA 0	JPN 0	ECU 0	HON 0
ESPN	ESPN2	ESPN	ESPN2	ABC	ABC	ESPN	ESPN	ABC	ABC

Abb. 19

2 Web-Datenquellen

Was ist euch bei den einzelnen Webseiten aufgefallen?

- ...

Agenda

- 1 Einleitung
- 2 Web-Datenquellen
- 3 Datenextraktion**
- 4 Zusammenfassung/ Ausblick

3 Datenextraktion

- Extraktion der Daten, hat eine große Bedeutung
- Offene und geschlossene Systemprogramme
- GIS Mapping von Ereignissen „Geographic Information Systems“

3.1 Programme

Nicht viele kommerzielle Programme vorhanden um Daten von Web-Datenquellen zu verarbeiten.

Was glaubt ihr warum das so ist?

3.1 Programme

Mangel an Homogenität der Daten über die verschiedenen Websites

- Generalisierte Programme

Auf dem Server, Limitierung der Bandbreiten

- Benutzer, die zu exzessiv nutzen, können gebannt oder geblockt werden.
- Einige Datamining Programme können als exzessive Nutzer angesehen werden.
- Es werden Programme entwickelt, die möglichst keine Aufmerksamkeit erzeugen.

3.1.1 Crowd Control Programme

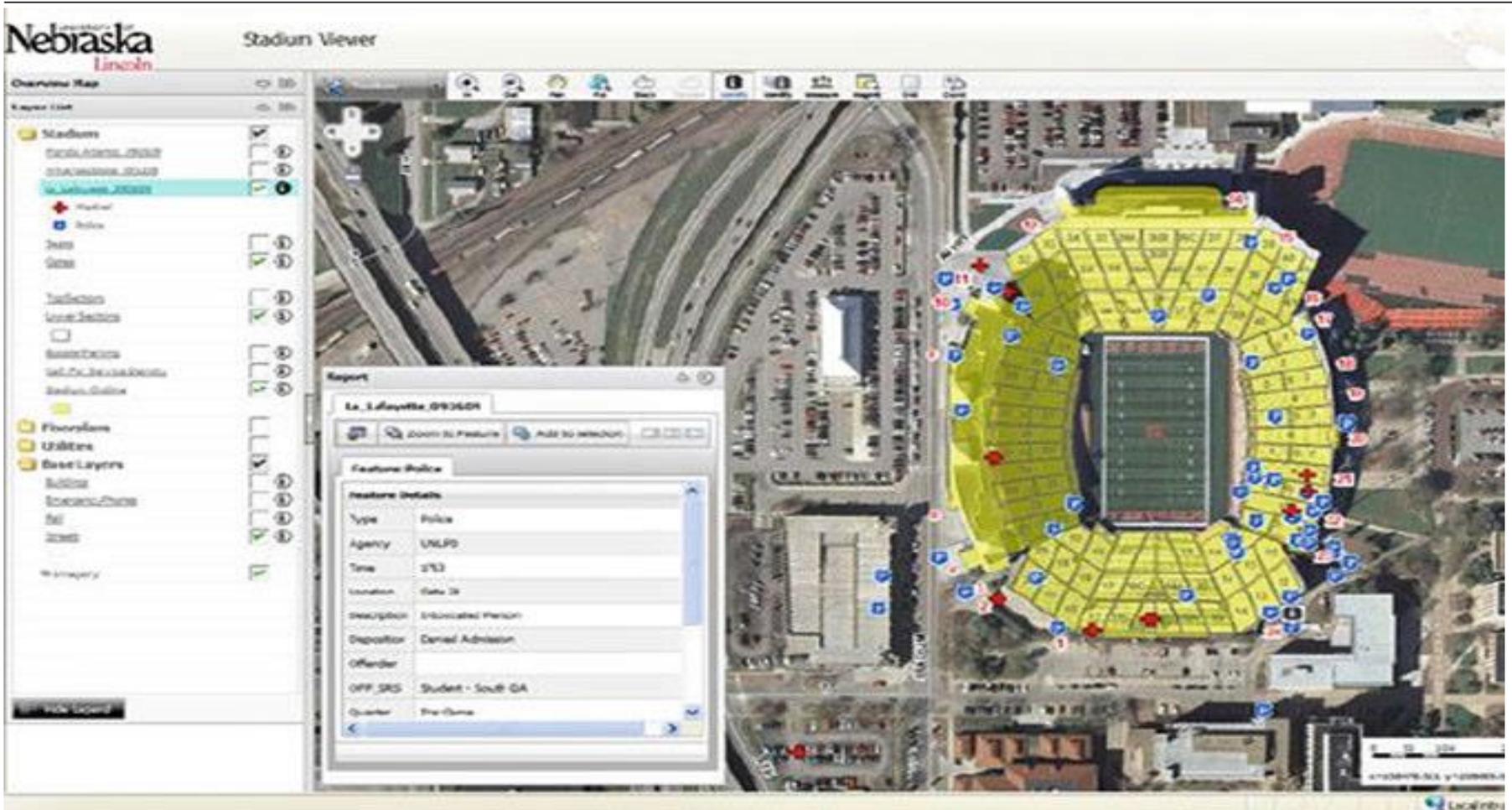


Abb.20

3.1.1 Crowd Control Programme

Früher:

- Sicherheitspersonal wurde aufmerksam gemacht.

Heute:

- Zusätzlich SMS Service

Probleme:

- Scherzanrufe

Tabellarische Erfassung von Problemabschnitten

3.1.2 Hochgeschwindigkeitssportarten

Aus dem Bereich der Multimedia-Sportforschung

- Eislaufen, Bobfahren, Skifahren, Skaten



Abb.21

3.1.2 Hochgeschwindigkeitssportarten

Nachteil von Kameras:

- Weniger Details, wenn die Kamera weiter entfernt ist



Abb.22

3.1.2 Hochgeschwindigkeitssportarten

Das Problem entsteht, wenn EDV-Tracking-Tools das Videomaterial überwachen.

→ Überforderung von Trackingsystemen

→ Lösung:

- Anwendung von Transformationsmatrizen auf jedes Einzelbild
- Nutzung von Punkten von stationären Objekt
- Kalman Filter

(vgl. Liu et al. 2009)

3.1.2 Hochgeschwindigkeitssportarten



TECHNISCHE
UNIVERSITÄT
DARMSTADT



<http://www.youtube.com/watch?v=qn5YQVvW-hQ>

Agenda

- 1 Einleitung
- 2 Web-Datenquellen
- 3 Datenextraktion

4 Zusammenfassung/ Ausblick

4 Zusammenfassung und Ausblick

- Technologien werden immer wichtiger
- Wir haben ein Angebot von Daten in den verschiedensten formen.
- Wahrscheinlich wird sich der Informatikanteil im Sport noch vergrößern je besser die Programme und Technologien werden.
- Dies hat zur Folge, dass es einfacher wird, Spiele zu kontrollieren und Schlussfolgerungen zu ziehen.
- Sportbezogene Daten haben sich in den letzten Jahren stark gewandelt

Literaturverzeichnis

- Stamm, C. (2013). *Hochgeschwindigkeitsvideoverarbeitung im Sport*. Schweiz: Springer
- Liu, G. & X. Tang, et al. 2009. *A Novel Approach for Tracking High Speed Skaters using a Panning Camera*. *Pattern Recognition* 42(11): 2922–2935.
- Red Sox.(2014). *Red Sox Game Notes Saturday, June 7, 2014*. Abgerufen unter <http://pressroom.redsox.com/Directory.aspx?&Location=GAME%20DAY%20INFO/Red%20Sox%20Game%20Notes/Regular%20Season&dir=June>.
- Robert P. Schumaker, Osama K. Solieman, Hsinchun Chen. (2010). *Sports Data Mining (Integrated Series in Information Systems)* New York Heidelberg Dordrecht London: Springer.



Abbildungsverzeichnis



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Abb. 1: Werbung auf Offiziellen Internet Seiten. Zugriff Juni 2014 unter <http://www.nhl.com/>

Abb. 2: Werbung auf Third-Party Internet Seiten. Zugriff Juni 2014 unter <http://www.nhl.com/>

Abb. 3: Playoffs. Zugriff Juni 2014 unter <http://de.global.nba.com/stats/league/playoffs/index.xhtml>

Abb. 4: MLB.com. Zugriff unter <http://mlb.mlb.com/home>

Abb. 5: Matchup. Zugriff Juni 2014 unter http://mlb.mlb.com/mlb/gameday/index.jsp?gid=2014_06_08_slmlb_tormlb_1&mode=gameday

Abb. 6: Vortag. Zugriff Juni 2014 unter http://mlb.mlb.com/mlb/scoreboard/index.jsp?tcid=gameday_about#date=6/7/2014

Abb. 7: Zukunft. Zugriff Juni 2014 unter http://mlb.mlb.com/mlb/scoreboard/index.jsp?tcid=gameday_about#date=6/8/2014

Abb. 8: *Premium. Robert P. Schumaker / Osama K. Solieman & Hsinchun Chen, 2010, S. 73. (Fig. 8.1 Hot/cold zones, courtesy of <http://www.mlb.com/mlb/gameday>)*



Abbildungsverzeichnis



Abb. 9: *Premium*. Robert P. Schumaker / Osama K. Solieman & Hsinchun Chen, 2010, S. 73. (Fig. 8.2 Pitching tendencies, courtesy of <http://www.mlb.com/mlb/gameday>)

Abb. 10: Internet Seite Retrosheet. Zugriff Juni 2014 unter <http://retrosheet.org/>

Abb. 11: NBY.co,. Zugriff Juni 2014 unter <http://www.nba.com/?22>

Abb. 12: *Premium*. Robert P. Schumaker / Osama K. Solieman & Hsinchun Chen, 2010, S. 75. (Fig. 8.3 Top five plus/minus rankings, courtesy of <http://www.nba.com/statistics/plusminus/>)

Abb. 13: *Pusminus*. Robert P. Schumaker / Osama K. Solieman & Hsinchun Chen, 2010, S. 75. (Fig. 8.3 Top five plus/minus rankings, courtesy of <http://www.nba.com/statistics/plusminus/>)

Abb. 14: *Stats at glance*. Zugriff Juni 2014 unter <http://www.nba.com/games/20091111/ATLNYK/gameinfo.html>

Abb. 15: [basketball-reference.com](http://www.basketball-reference.com/). Zugriff Juni 2014 unter <http://www.basketball-reference.com/>



Abbildungsverzeichnis



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Abb. 16: Cricket. Zugriff Juni 2014 unter <http://www.espnccricinfo.com/>

Abb. 17: Football. Zugriff Juni 2014 unter <http://www.nba.com/?22>

Abb. 18: Hockey. Zugriff Juni 2014 unter <http://www.nhl.com/>

Abb. 19: Fußball. Zugriff Juni 2014 unter <http://www.mlssoccer.com/>

Abb. 20: *Stadion GSI. Robert P. Schumaker / Osama K. Solieman & Hsinchun Chen, 2010, S. 86. (Fig. 8.14 Nebraska's crowd control program using GIS mapping, courtesy of UNL police)*

Abb. 21: Tennis Blur. Zugriff Juni 2014 unter <http://static.guim.co.uk/sys-images/Guardian/Pix/pictures/2013/11/7/1383840386630/Rafael-Nadal--014.jpg>

Abb. 22: Tennis Stadion. Zugriff Juni 2014 unter http://upload.wikimedia.org/wikipedia/commons/thumb/2/22/Centre_Court.jpg/300px-Centre_Court.jpg



**Vielen Dank
für Ihre Aufmerksamkeit!**