

# Web Sports Data Extraction and Visualization

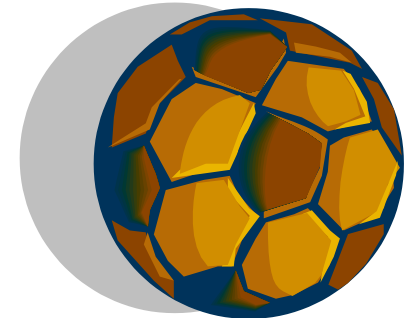


TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

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

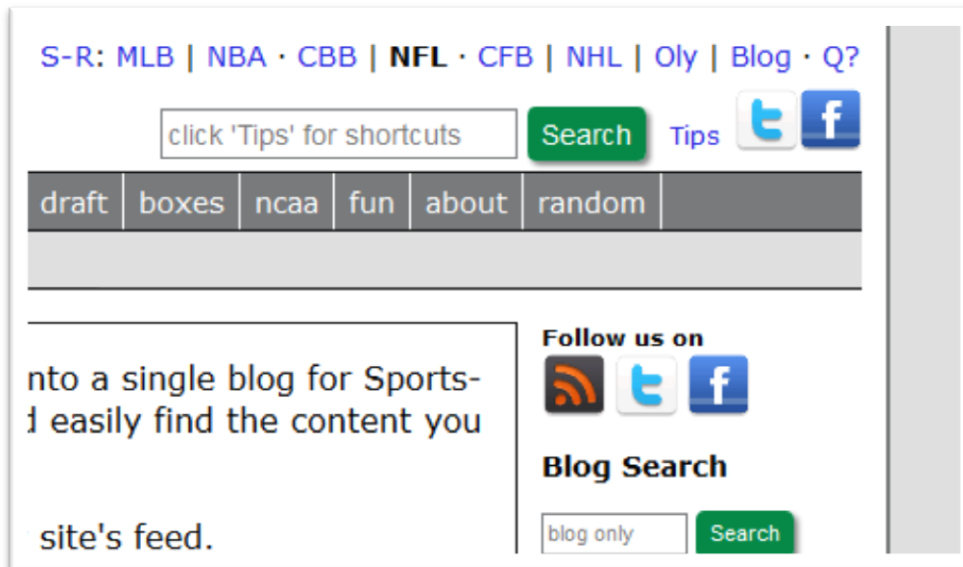
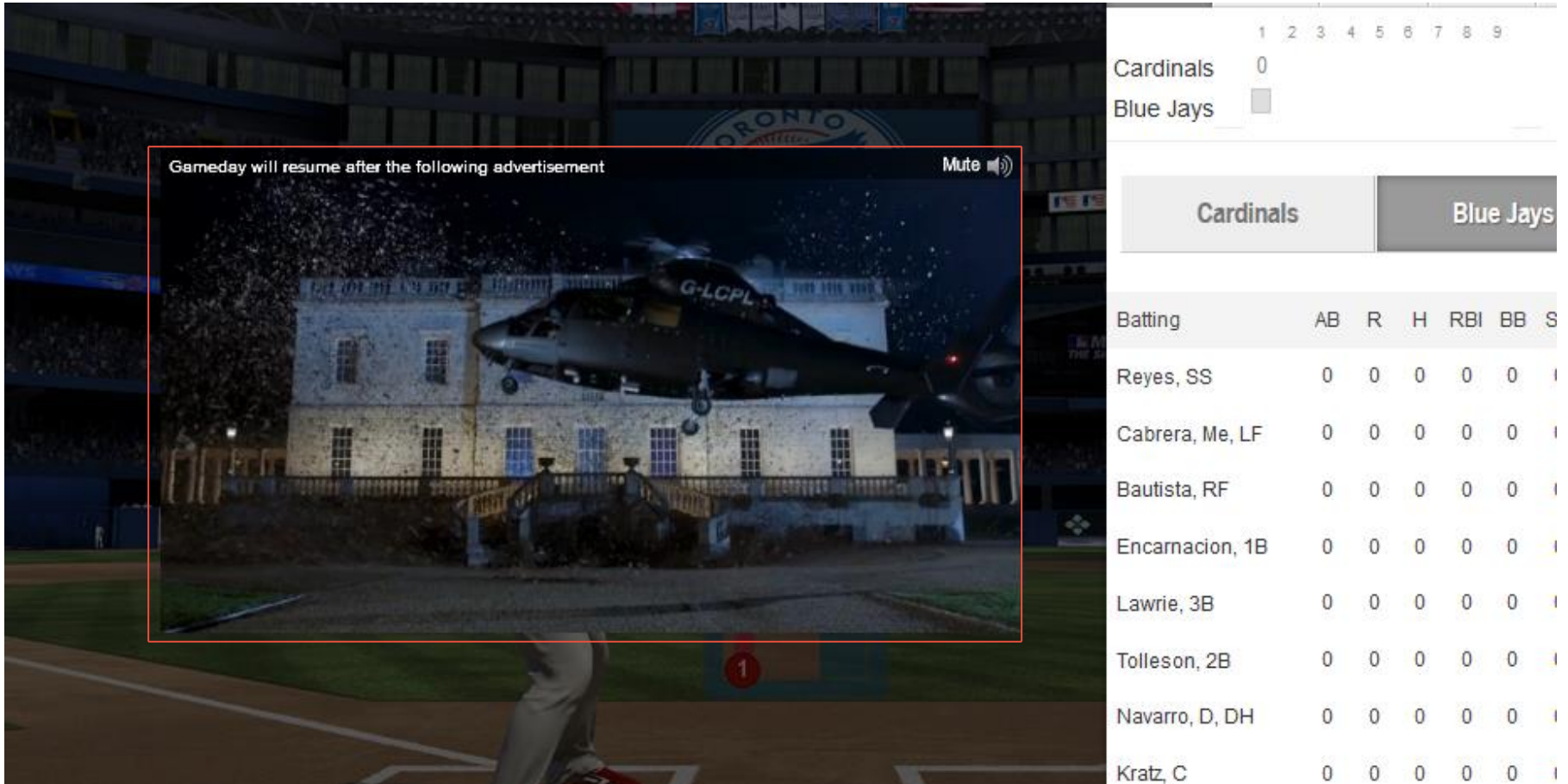


Abb.2



Abb.1

## 2 Web-Datenquellen Werbung



The screenshot shows a live broadcast of a baseball game. An advertisement for a helicopter is overlaid on the field. The helicopter is a Sikorsky UH-60 Black Hawk, registration G-LCPL, flying over a large, illuminated building at night. The text "Gameday will resume after the following advertisement" is visible in the top left of the ad, and "Mute" is in the top right. A red circle with the number "1" is at the bottom center of the ad.

Scoreboard:

	1	2	3	4	5	6	7	8	9
Cardinals	0								
Blue Jays									

Team Selection:

Cardinals | Blue Jays

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

Abb.7

## 2 Web-Datenquellen Abbildungen

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

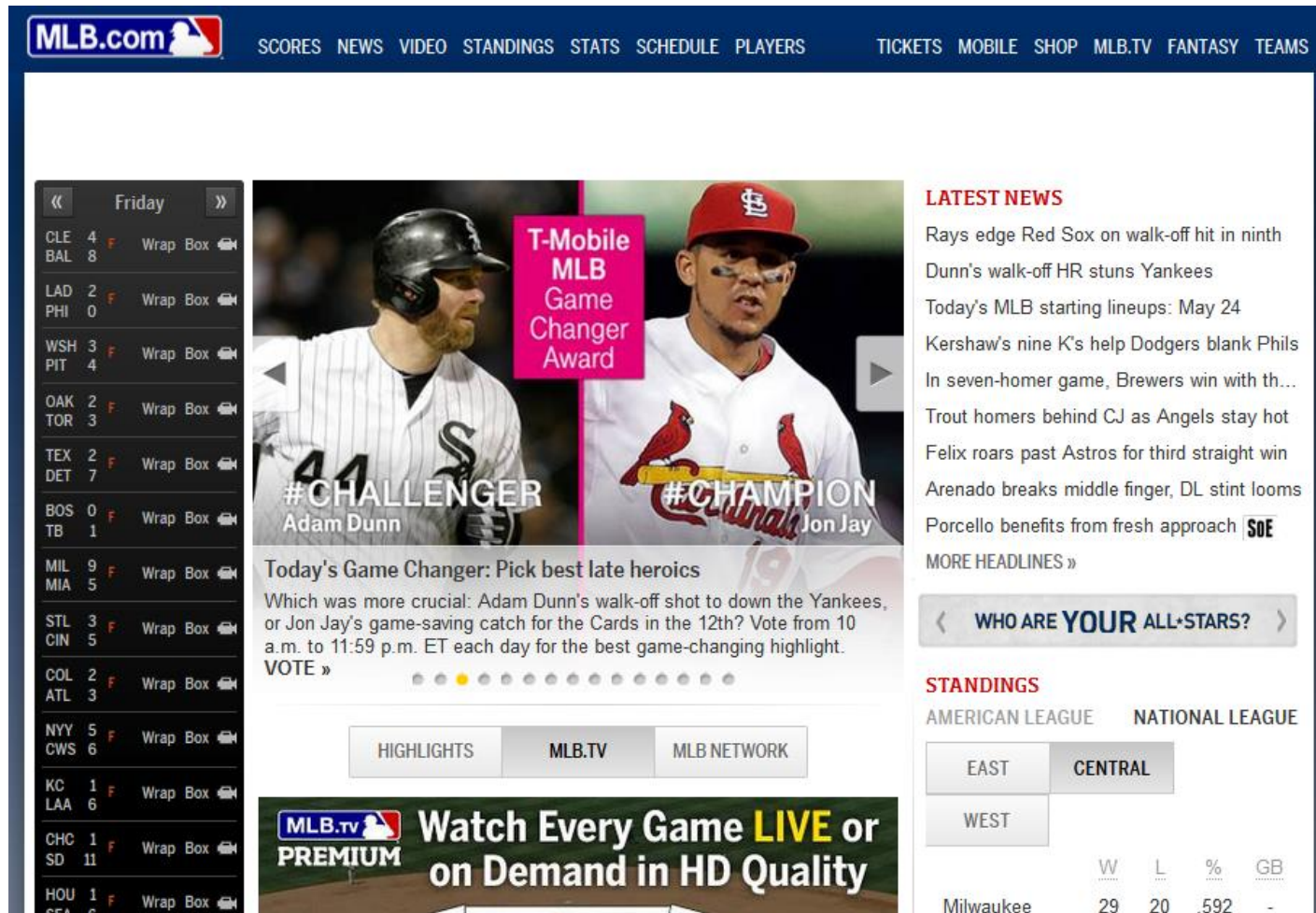


Abb.3



# 2.1 Baseball MLB Major League Baseball

MLB.com



The screenshot shows the MLB.com website interface. At the top, there is a navigation bar with links for SCORES, NEWS, VIDEO, STANDINGS, STATS, SCHEDULE, PLAYERS, TICKETS, MOBILE, SHOP, MLB.TV, FANTASY, and TEAMS. The main content area is divided into several sections:

- Game Results:** A list of games from Friday, including scores and links to wrap-ups for various teams like CLE, BAL, LAD, PHI, WSH, PIT, OAK, TOR, TEX, DET, BOS, TB, MIL, MIA, STL, CIN, COL, ATL, NYY, CWS, KC, LAA, CHC, SD, HOU, and SEA.
- Game Changer Award:** A central feature highlighting Adam Dunn (#CHALLENGER) and Jon Jay (#CHAMPION) for the T-Mobile MLB Game Changer Award. It includes a poll: "Today's Game Changer: Pick best late heroics" with options for Adam Dunn's walk-off shot or Jon Jay's game-saving catch. A "VOTE" button is present.
- Latest News:** A list of recent articles such as "Rays edge Red Sox on walk-off hit in ninth", "Dunn's walk-off HR stuns Yankees", and "Today's MLB starting lineups: May 24".
- Standings:** A section for "STANDINGS" with tabs for "AMERICAN LEAGUE" and "NATIONAL LEAGUE". Under "NATIONAL LEAGUE", there are tabs for "EAST", "CENTRAL", and "WEST". A partial table shows Milwaukee's record: W 29, L 20, % .592, GB -.
- Promotional Banners:** A "WHO ARE YOUR ALL-STARS?" poll banner and an "MLB.TV PREMIUM" banner for watching games live or on demand in HD quality.

Abb.4


## 2.1 Baseball MLB.com


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

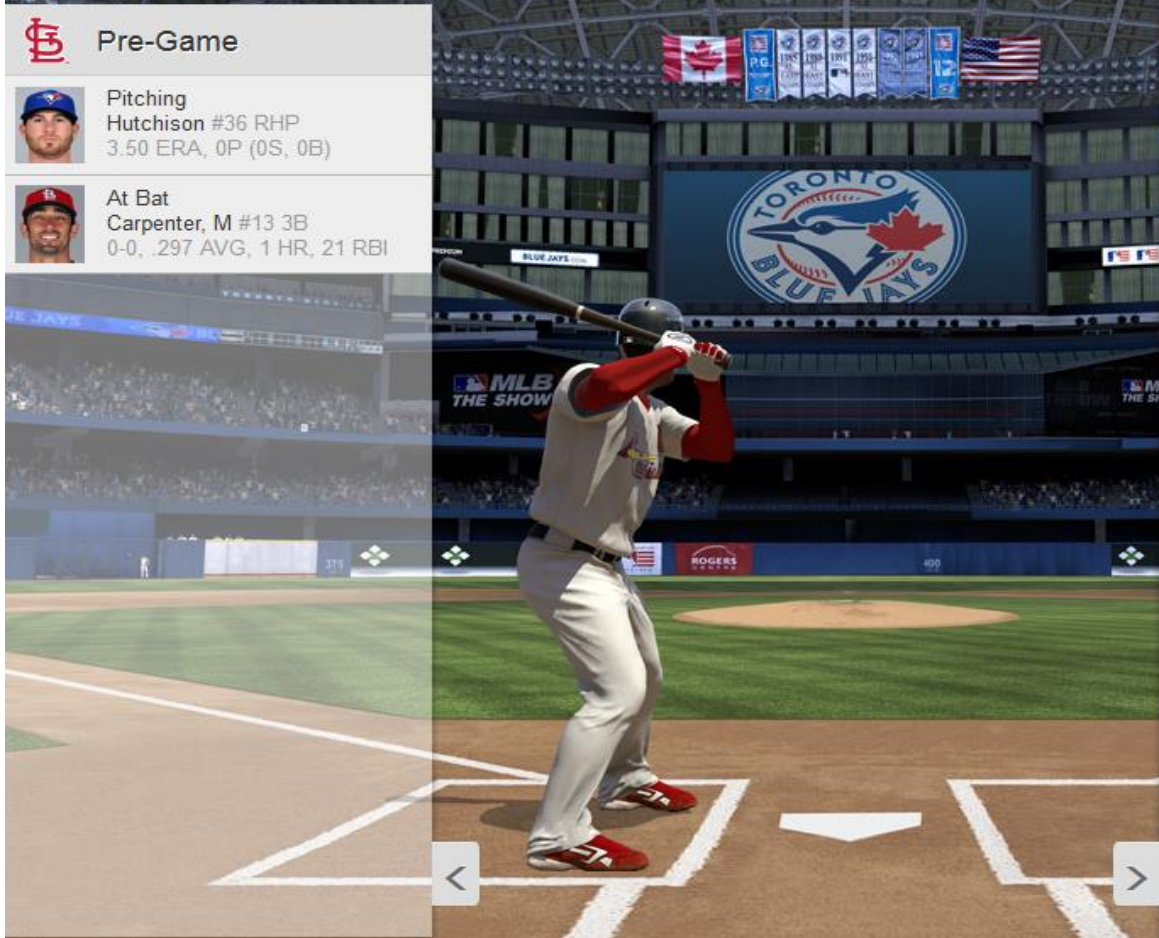
- Bsp. aus der Saison 2009: Cleveland's Shin-Soo Choo hat eine 0.405 Trefferquote gegen Kansas City im Kaufmann Stadion
- Statistiken von Schläger-Werfer-Paaren/Kombinationen
- [Bsp. PDF RED SOX Notes](#)


# 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

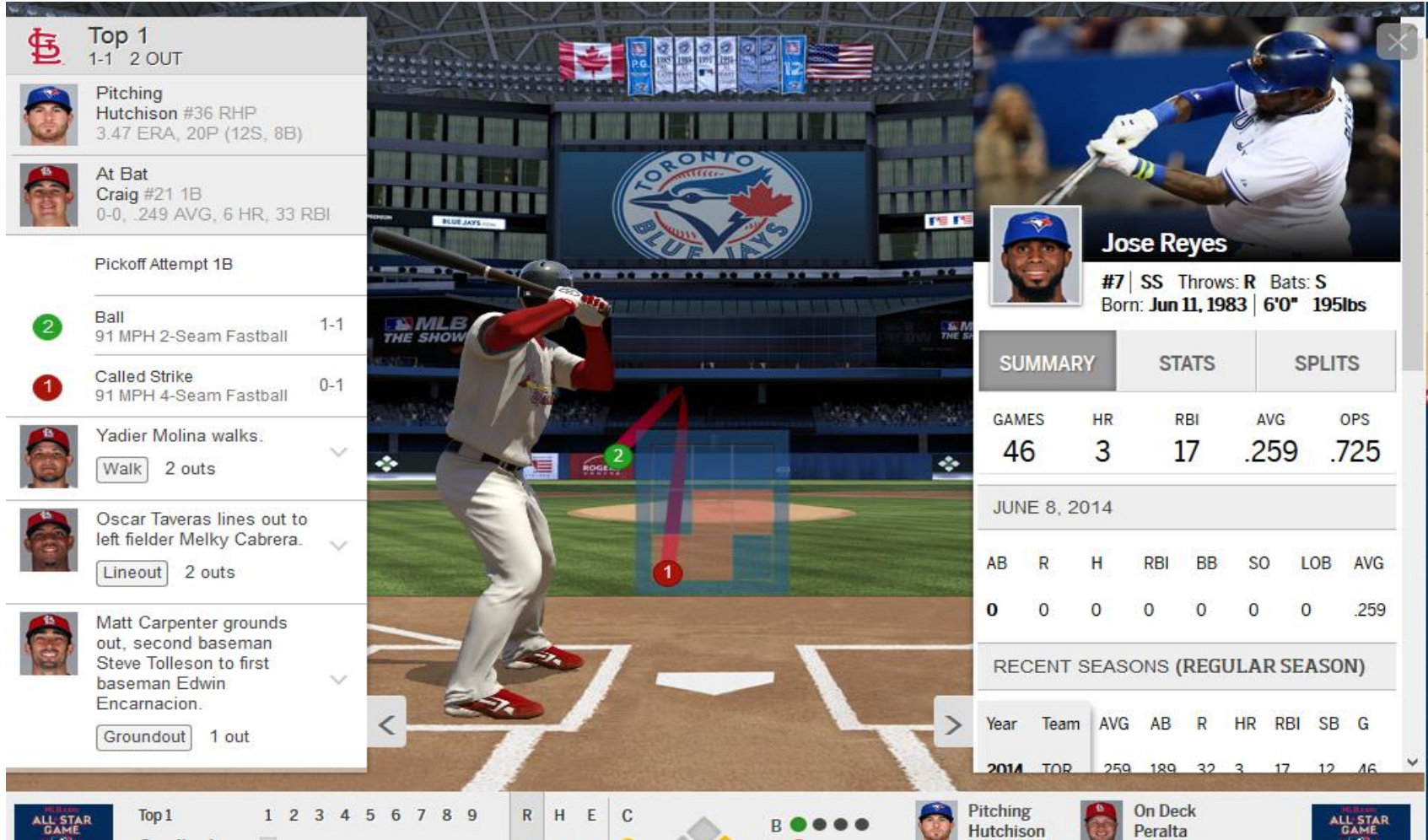
AB	R	H	RBI	BB	SO	LOB	AVG
0	0	0	0	0	0	0	.259

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	Final	R	H	E	Final	R	H	E
<b>STL</b> (32-31)	5	8	0	<b>HOU</b> (27-36)	0	5	2	<b>CLE</b> (31-31)	8	12	0
<b>TOR</b> (38-25)	0	3	1	<b>MIN</b> (29-31)	8	7	0	<b>TEX</b> (31-31)	3	8	1
W: <b>Miller, S</b> (7-5, 3.59) L: <b>Buehrle</b> (10-2, 2.04) HR: STL: <b>Grichuk</b> (1) .				W: <b>Gibson</b> (5-5, 3.91) L: <b>Feldman</b> (3-4, 4.12) HR: None.				W: <b>Tomlin</b> (4-2, 3.12) L: <b>Tepesch</b> (2-2, 4.91) HR: CLE: <b>Gomes, Y</b> (7) , <b>Santana, C</b> (7) .			
<span>📺</span> <span>WRAP</span> <span>BOX</span> <span>📺</span>				<span>📺</span> <span>WRAP</span> <span>BOX</span> <span>📺</span>				<span>📺</span> <span>WRAP</span> <span>BOX</span> <span>📺</span>			

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)					
<a href="#">PREVIEW</a>	<a href="#">NOTES</a>			<a href="#">PREVIEW</a>	<a href="#">NOTES</a>			<a href="#">PREVIEW</a>	<a href="#">NOTES</a>		

Abb.7

## 2.1 Baseball MLB Premium

Rot = hohe Trefferquote

Blau = niedrige Trefferquote

Nummerierung = Anzahl der pitches

Spieltages Tool

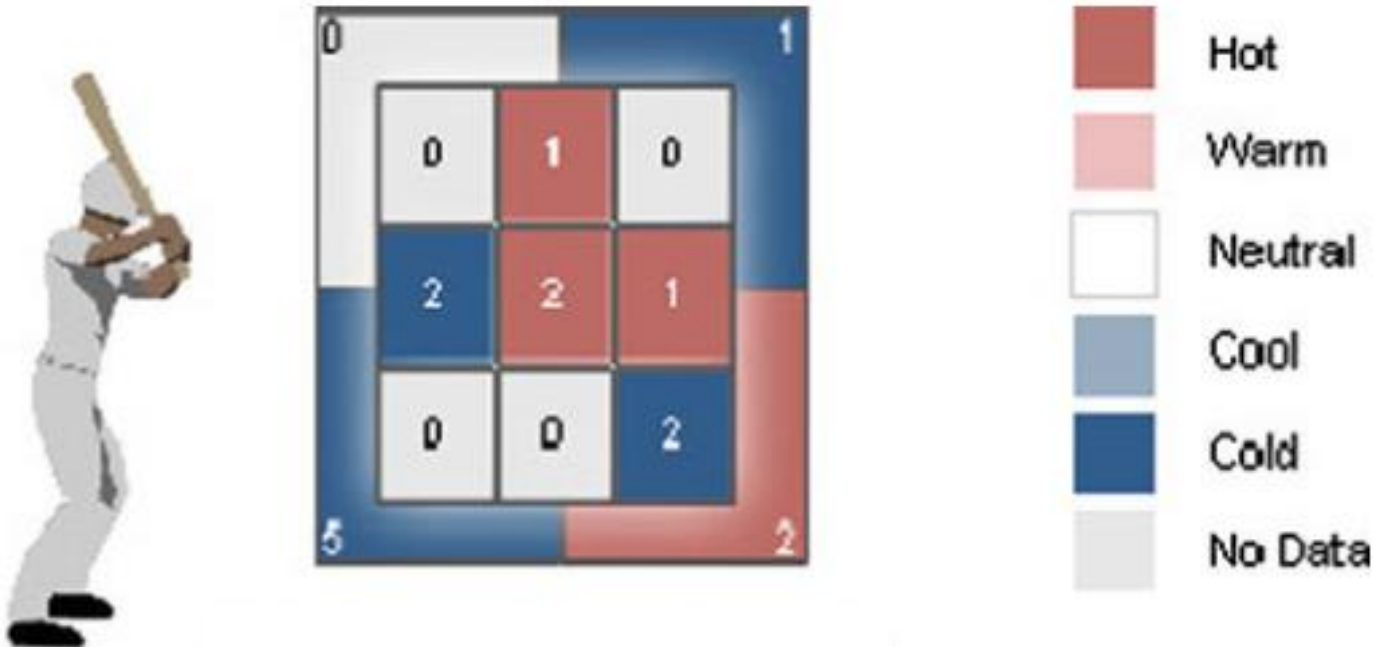


Abb.8

## 2.1 Baseball MLB Premium

### Pitcher Tendencies Tool

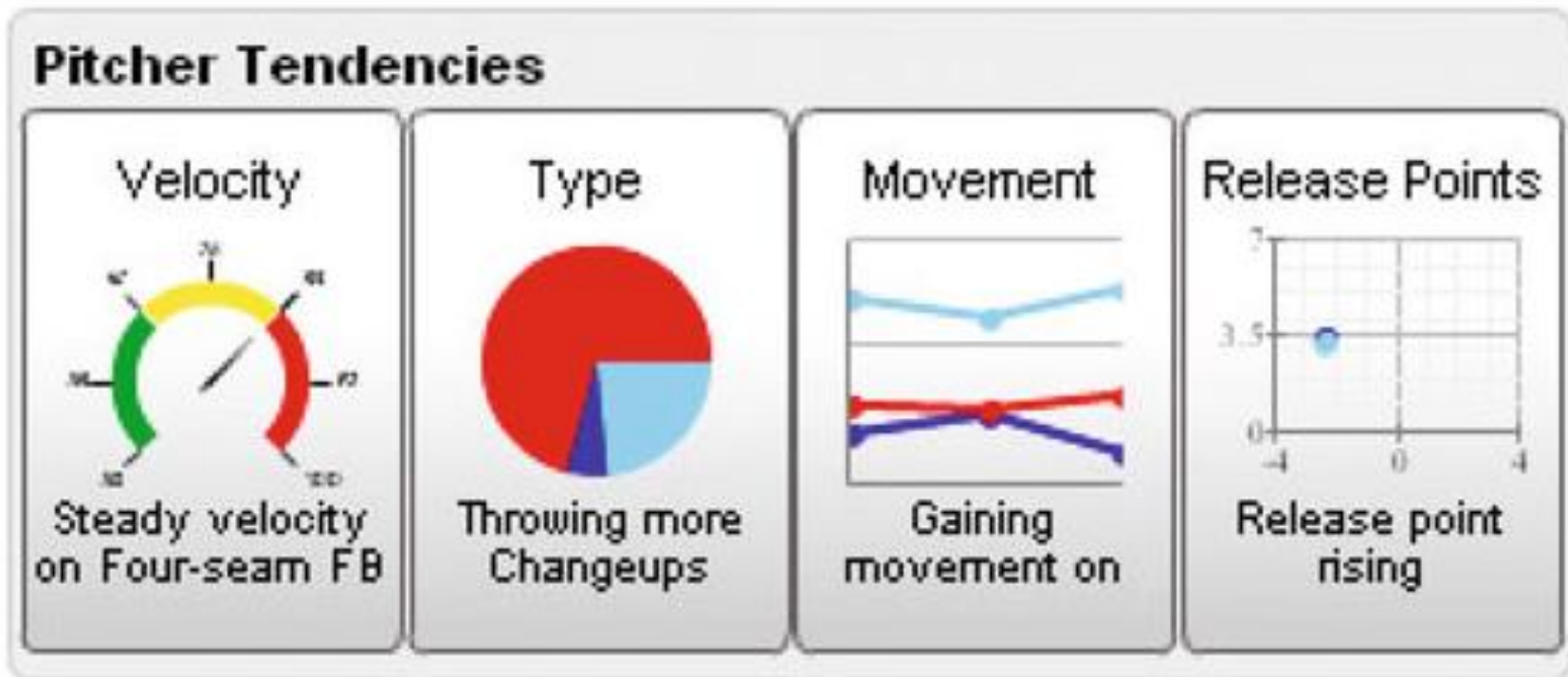


Abb.9



## 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](http://Baseball-reference.com)
  - Mit historischen und aktuellen Daten zusätzlich Minor-League-Teams (unterklassige Ligen)
- [Baseball1.com](http://Baseball1.com) (Sean Lehman-Baseball-Archiv)

## 2.1 Basketball NBA.com

NBA.com

The screenshot displays the NBA.com website interface. At the top, there is a navigation bar with the SPOX.com logo and the NBA DEUTSCHLAND logo. Below this, a menu lists various sports categories: HOME, FUSSBALL, NBA, FORMEL 1, BASKETBALL, BOXEN, TENNIS, GOLF, EISHOCKEY, HANDBALL, US-SPORT, MEHR, VIDEO, and MYSPOX. A secondary menu includes NBA LEAGUE PASS, Videos, Playoffs, Ergebnisse, Tabellen, Stats, Spieler & Teams, News, Fotos, LIVE-STREAM, and D-League. The main content area features a schedule for the NBA LEAGUE PASS, showing games for the days from Saturday, May 24 to Friday, May 30. The games listed include IND vs MIA, SAS vs OKC, and MIA vs IND. A live-stream video player is prominently displayed, showing a game between the Thunder and the Spurs. The video title is "Ein ungleiches Duell" and the subtitle is "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 video player is a sidebar titled "NBA-NEWS" with a list of recent news items, including "Spiel 3: George darf spielen" and "Blitzheilung bei OKC's Ibaka".

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

**NBA LEAGUE PASS**  
SEHEN SIE JEDES SPIEL DER PLAYOFFS  
**PLAYOFFS**

Abb. 11

## 2.1 Basketball NBA.com

NBA.com

The screenshot shows the NBA.com website interface in German. At the top, there is a 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 banner for 'NBA DEUTSCHLAND' and a navigation menu with 'R1', 'R2', 'CF', and 'F'. Below the banner, there is a section titled 'Tonight on NBA TV: 'The84Draft' (9 ET) | Film review | 84 facts on the '84 Draft | 2014 Finals'. The main content area is divided into three columns. The left column features a large image of three Miami Heat players in white jerseys with 'HEAT' written on them. Below the image is the text 'Morning Tip: Sticks And Stones ...' and a short paragraph: 'Scoffers and haters, take note. The Miami Heat barely hear your words in their three-peat chase. Hang Time: Defensive change helps Heat | Where has Kawhi gone? | 20 key stats |'. The middle column contains a video player with a play button icon and the title 'Mini-Movie: Nothing Easy'. The right column contains another video player with a play button icon and the title '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

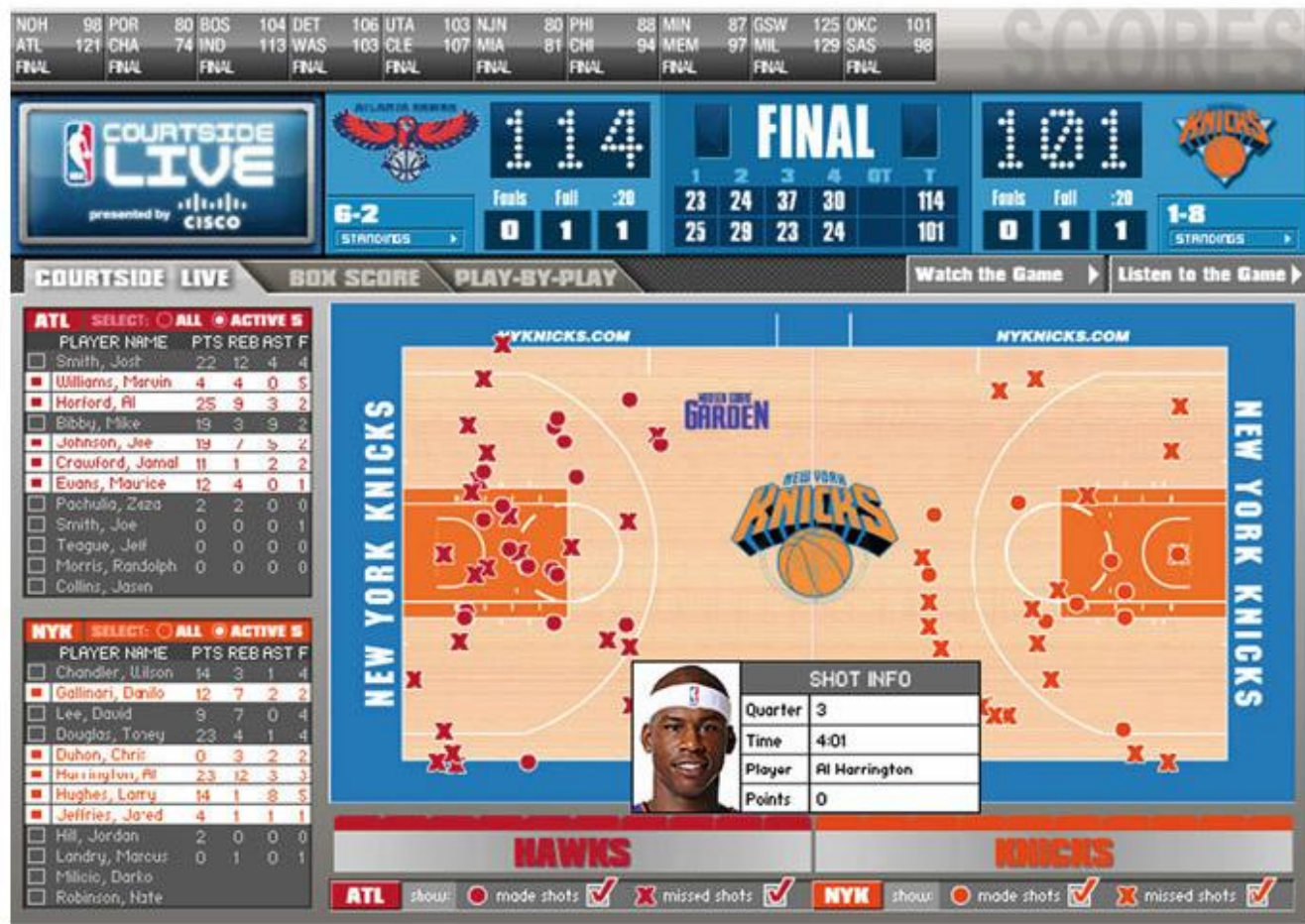
### Top Five-Player Combinations

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



**SCORES**

NOH	98	POR	80	BOS	104	DET	106	UTA	103	NJN	80	PHI	88	MIN	87	GSW	125	OKC	101
ATL	121	CHA	74	IND	113	WAS	103	CLE	107	MIA	81	CHI	94	MEM	97	MIL	129	SAS	98
FINAL	FINAL	FINAL	FINAL	FINAL	FINAL	FINAL	FINAL	FINAL	FINAL	FINAL	FINAL	FINAL	FINAL	FINAL	FINAL	FINAL	FINAL	FINAL	FINAL

**COURTSIDE LIVE** presented by CISCO

**ATLANTA HAWKS** 114 **FINAL** **NEW YORK KNICKS** 101

Feats Fall :20

1	2	3	4	OT	T
23	24	37	30		114
25	29	23	24		101

Feats Fall :20

**6-2** STANDINGS **1-8** STANDINGS

**COURTSIDE LIVE** **BOX SCORE** **PLAY-BY-PLAY** Watch the Game Listen to the Game

**ATL** SELECT: ALL ACTIVE 5

PLAYER NAME	PTS	REB	AST	F
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
Evans, Maurice	12	4	0	1
Pachulia, Zaza	2	2	0	0
Smith, Joe	0	0	0	1
Teague, Jeff	0	0	0	0
Morris, Randolph	0	0	0	0
Collins, Jason				

**NYK** SELECT: ALL ACTIVE 5

PLAYER NAME	PTS	REB	AST	F
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
Milicic, Darko				
Robinson, Nate				

**SHOT INFO**

Quarter	3
Time	4:01
Player	Al Harrington
Points	0

**HAWKS** **KNICKS**

ATL show:  made shots  missed shots NYK show:  made shots  missed shots

Abb. 13

## 2.1 Basketball NBA.com



“Stats at a Glance“

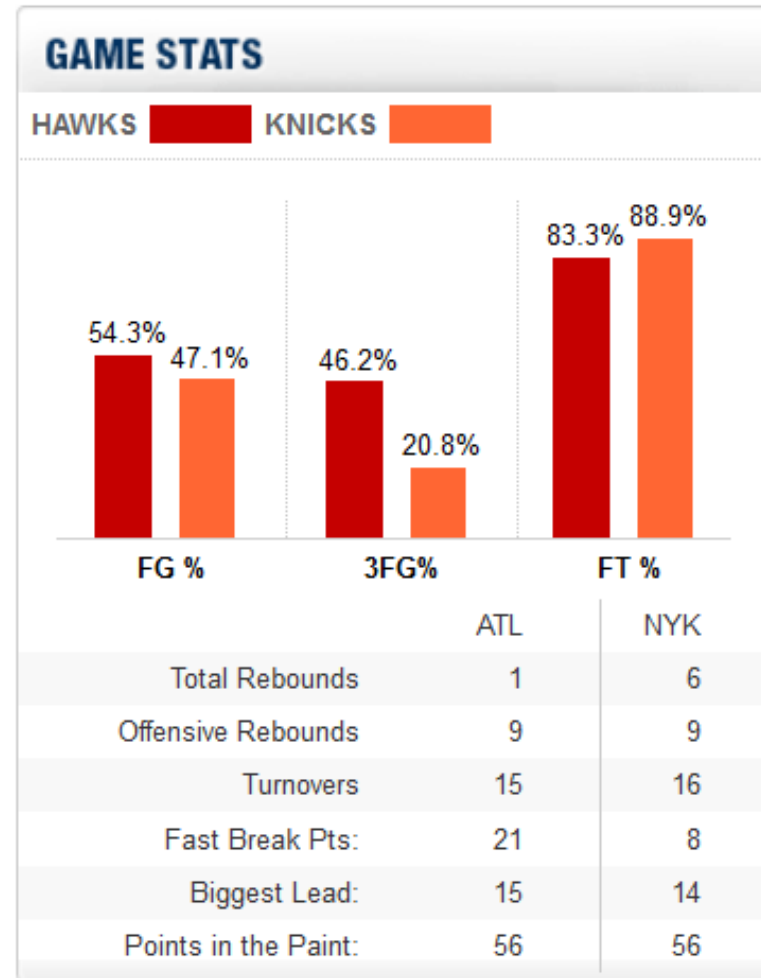


Abb. 14

---

## 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 left is the logo with the text "BASKETBALL-REFERENCE.COM". To the right are navigation links for various sports: S-R: MLB | NBA · CBB | NFL · CFB | NHL | Oly | Blog · Q?, a search bar, and social media icons for Twitter and Facebook. Below the navigation is a menu with categories like "play index", "box scores", "players", "teams", "seasons", "coaches", "leaders", "awards", "playoffs", "draft", "olympics", and "more [+]", with "playoffs" selected. A breadcrumb trail reads "Mobile Site You Are Here > Basketball-Reference.com > Basketball Statistics and History". A news snippet is visible: "News: s-r blog:DNP/Inactive Records For Player Gamelogs". The main content area features a search bar and a "Search" button. Below it is a "Basketball-Reference.com News" section with several blog links. The central focus is the "2013-14 NBA Playoffs" section, which includes a "Finals" subsection for "San Antonio (1) vs. Miami (2)". It provides details such as "San Antonio leads the series 1-0", "Next game: Sunday Jun 8th @ 8:00 PM EST", "Broadcast: ABC", and win probabilities for the series and finals. A "Yesterday's Top Performers" box on the right indicates "No games played." and offers links to "See all of yesterday's players", "2013-14 NBA Debuts", and "2013-14 NBA All-Stars".

Abb. 15

# 2.1 Cricket Cricinfo.com



The screenshot shows the ESPNcricinfo website. 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 banner: "Now you can set your own homepage. Choose from the dropdown below." The main header includes "CHOOSE EDITION UK", the ESPNcricinfo logo, a Facebook Like button (3.5m), a search bar, and a "Welcome Guest" message with "Sign In" and "Log In" options. A secondary navigation bar contains: Series, Countries, Live Scores, Fixtures, Results, News, Features, Photos, Video & Audio, Blogs, Statistics, Archive, Fantasy, Mobile. Below this is a "Quick Links" section with items like "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 several sections:

- VIDEO NEWS AND ANALYSIS:** A large video player showing a NatWest T20 match with the text "The best night of NatWest T20 so far".
- WI v NZ, 1st Test, Kingston, 1st day:** A headline "Latham and Williamson build slowly" with a "Live scorecard" link. Sub-headlines include "Lunch report - Taylor strikes on return", "How do you police the doosra?", and "Cordon - The return of Benn, Bravo and Taylor".
- Chris Gayle's 100th Test:** A headline "Gayle targets 'huge performance'" with sub-headlines: "Want to keep scoring and to represent WI as long as I can", "Cozier: The Jamaican man mountain", and "Much more than a short-format batsman".
- Northants v Sri Lankans, Northampton, 4th day:** A headline "SL buoyant despite Keogh ton" with sub-headlines: "Scorecard | Report - Prasad states Test case" and "Battle with England seamers pivotal - Adams".
- Sri Lanka tour of England 2014:** A headline "Senanayake to be tested in Cardiff" with sub-headline: "Offspinner to undergo biomechanical analysis at new centre".

At the bottom of the main content area, there is a "BET NOW" section: "Jordan to blast Sri Lanka away 100/30 >>".

On the right side, there is a "Scores" section with tabs for "Scores", "Results", and "Fixtures". It lists "International" and "Domestic" matches:

- International:** New Zealand 94/1 v West Indies, Sri Lankans 558/8d & 99/4d v Northants 345 (Match drawn).
- Domestic:** 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, there are links for "Live scores pages", "Scores on Mobile", and a "Live" indicator.

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 "TEAM SITES". A search bar is located on the right. Below the navigation bar, there are tabs for "FANTASY", "NEWS", "VIDEO", "SCORES", "SCHEDULES", "STANDINGS", "STATS", "WATCH GAMES", "TICKETS", and "SHOP". A banner for "WATCH NFL NETWORK NOW: AMERICA'S GAME 11:00 AM ET" is visible. 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, there 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 a featured article titled "MILLER TIME" about Von Miller, showing a photo of him in an orange jersey with the number 52 and a star rating of 76. 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, SLOVENČINA, DEUTSCH) and social media icons. Below this is a schedule for Round 3 of the Stanley Cup Playoffs, listing games from Saturday 5/24 to Friday 5/30. The main content area features a navigation menu (SCORES, STANDINGS, SCHEDULE, STATS, PLAYERS, TEAMS, NEWS, VIDEO, GAMECENTER LIVE, NHL NETWORK, PLAYOFFS, FANTASY, TICKETS, SHOP) and a central game preview for Game 3 between the Chicago Blackhawks (CHI) and the Los Angeles Kings (LAK), which is tied 1-1. To the right, there are sections for 'TOP HEADLINES' and 'NHL GREEN' with various news items. A large video player in the center shows a player in a blue jersey, with the headline 'STEPAN SIDELINED' and a sub-headline 'Rangers center sustains broken jaw, out indefinitely following surgery'. A sidebar on the left promotes 'NHL GAMECENTER LIVE' and 'WATCH THE STANLEY CUP™ PLAYOFFS LIVE'. The bottom of the page includes a copyright notice: '© NHL 2014. All Rights Reserved.'

Abb. 18

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


[MLS SOCCER.COM](#)
[Scores](#)
[World Cup](#)
[News](#)
[Videos](#)
[Schedule](#)
[Standings](#)
[Fantasy](#)
[Shop](#)
[More ▾](#)

[Login](#) | [Sign Up](#)




**VS**


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

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

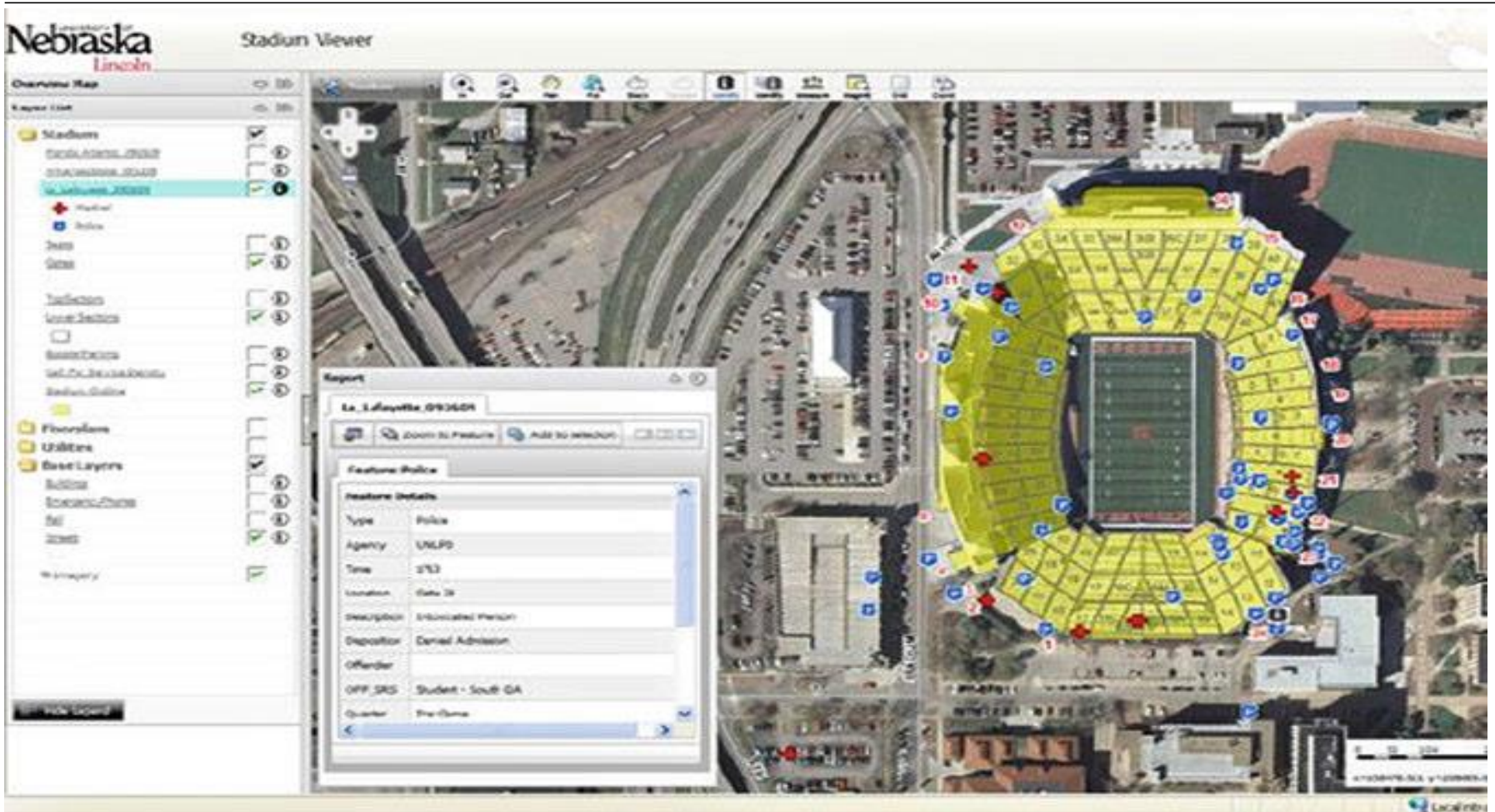


Abb.20

## 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:

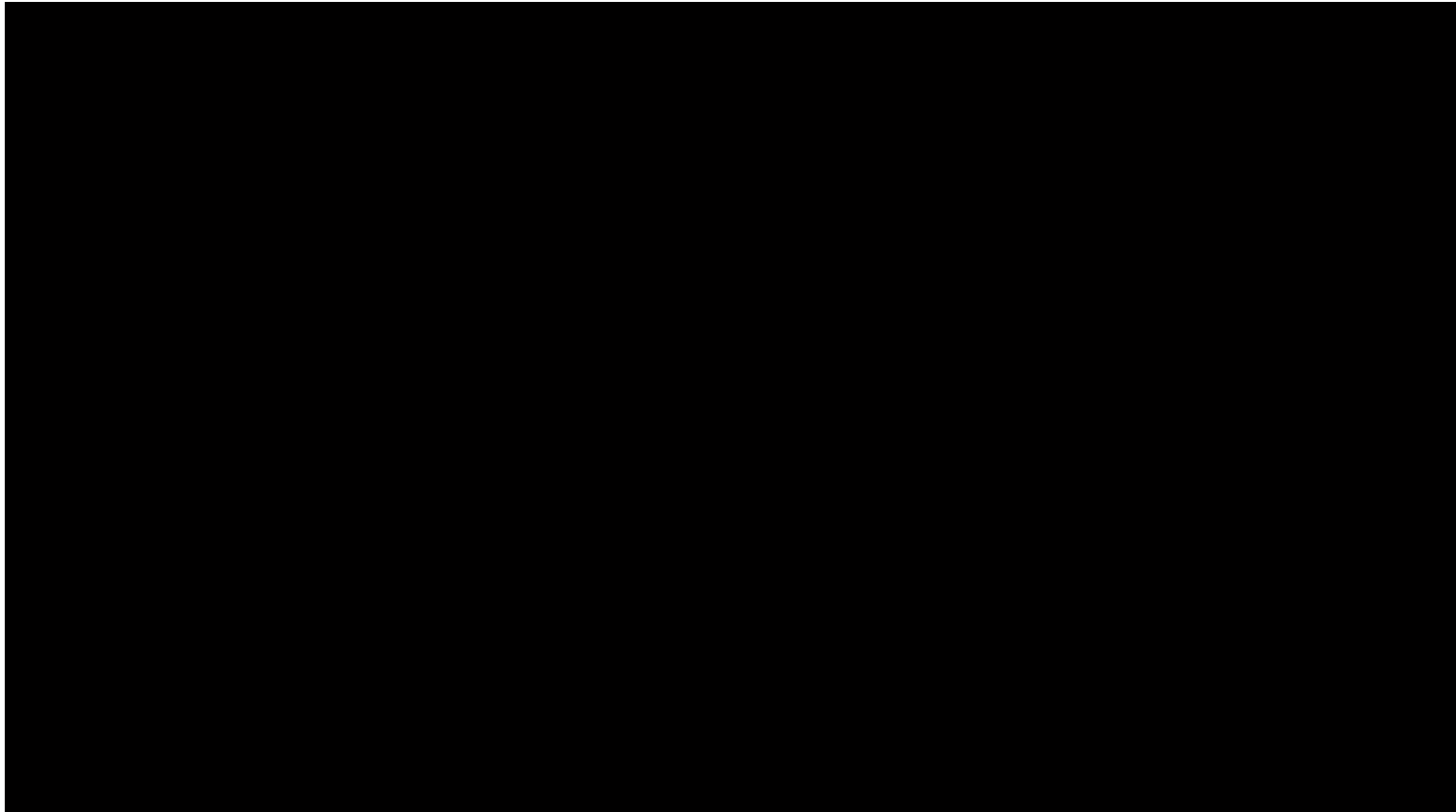
- Anwendung von Transformationsmatrizen auf jedes Einzelbild
- Nutzung von Punkten von stationären Objekt
- Kalman Filter

(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](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](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](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](http://upload.wikimedia.org/wikipedia/commons/thumb/2/22/Centre_Court.jpg/300px-Centre_Court.jpg)



**Vielen Dank  
für Ihre Aufmerksamkeit!**