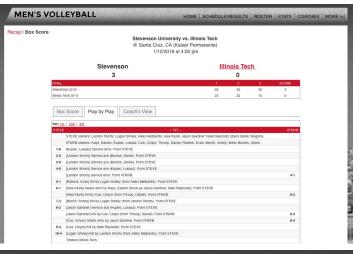# Article Generation
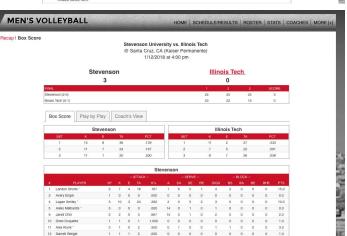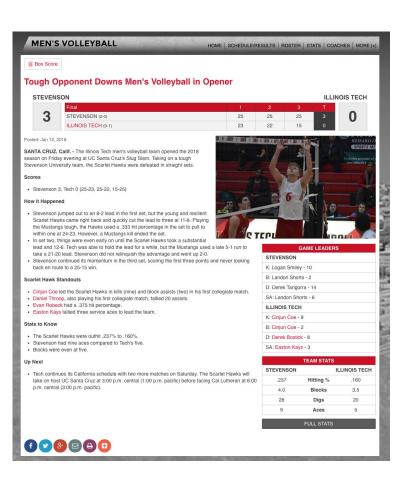## Volleyball Match Summarization

Michael Drews

Terminology

- 3 to 5 sets
  - Score at least 25 points, win by two
- Stats
  - Kill
  - Block
  - Dig
  - Hitting Percentage

Goal

Natural Language Generation

1. Content determination
2. Text structuring
3. Sentence aggregation
4. Lexicalisation
5. Referring expression generation
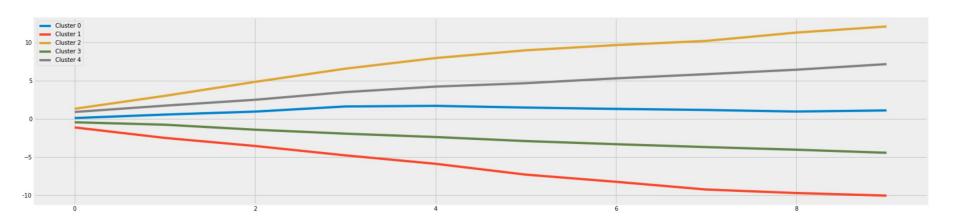6. Linguistic realisation

Approach

- Web crawl 121 historical articles
  - 571 sets
  - 2862 sentences
- Vectorize the play-by-play data for each set and then cluster them
- Compute stat leaders and find notable performances
- Vectorize the play-by-play data for each set and classify them
- Generate sentences using templates that are based off of the language used in those clusters

Clustering

- Creates list of the point differential after each play
- Creates a list of length 10, where each element is a sublist from the original list
  - [1, 2, 3, 4, 5, 6, 7, 8, 9]
  - [[1, 2, 3], [4, 5, 6], [7, 8, 9]]
- Final vector is the list of the average value in each sublist
  - [2, 5, 8]

Clusters

Sections

- Title
- Introduction
- How It Happened
- Scarlet Hawk Standouts
- Stats To Know
- Up Next

Title Templates

```python
def generate_title(info_dict, result, opponent, gender, num_sets):
    corpus = {
        # One Class for W
        "W": [
            "<GENDER>'s Volleyball <good-verb> <OPP>",
            "Scarlet Hawks <good-verb> <OPP>",
            "<GENDER>'s Volleyball <good-verb> <OPP> <SET-SCORE>",
            "Scarlet Hawks <good-verb> <OPP> <SET-SCORE>",
            "<GENDER>'s Volleyball Impresses In Win Over <OPP>",
            "<GENDER>'s Volleyball <good-verb> <OPP> in <NUM-SETS>-Set Match",
            "Scarlet Hawks <good-verb> <OPP> in <NUM-SETS>-Set Match"
        ],

        # One class for L
        "L": [
            "<bad-adj> Opponent <lose-verb> <GENDER>'s Volleyball",
            "<bad-adj> <OPP> team <lose-verb> <GENDER>'s Volleyball",
            "<GENDER>'s Volleyball Battles <bad-adj> Opponent",
            "<OPP> <lose-verb> <GENDER>'s Volleyball",
            "<GENDER>'s Volleyball Defeated by <OPP>",
            "<GENDER>'s Volleyball Falls to <OPP>",
            "<GENDER>'s Volleyball Drops <NUM-SETS>-Set Match at <OPP>"
        ],
        "bad-adj": ["Tough", "Strong", "Solid"],
        "lose-verb": ["Beats", "Tops", "Defeats"],
        "good-verb": ["Sweeps", "Shuts Out", "Cruises Past"]
    }
```

Introduction Templates

```python
def generate_intro(info_dict, result, opponent, gender, num_sets):
    corpus = {
        "first": {
            "W": [
                "<LOCATION> - The Illinois Tech <GENDER>'s volleyball team <played-verb> <OPP>, <win-verb> in <NUM-SETS> sets.",
                "<LOCATION> - The Scarlet Hawks played <good-adj> <NUM-SETS>-set match against <OPP>.",
                "<LOCATION> - The Scarlet Hawks played <good-adj> <NUM-SETS>-set versus <OPP>, <win-verb> <SET-SCORE>"
            ],
            "L": [
                "<LOCATION> - The Illinois Tech <GENDER>'s volleyball team <lost-verb> <OPP>, <losing-verb> in <NUM-SETS> sets."
            ]

        },
        "second": {
            "W": [
                "The <event-noun> <moved-verb> Scarlet Hawks to a <RECORD> overall record on the season. "

            ],
            "L": [
                "<despite-np> the <result> <moved-verb> Scarlet Hawks to a <RECORD> overall record."

            ]

        },
        "played-verb": ["defeated", "beat", "outplayed", "outworked"],
        "win-verb": ["overcoming them", "winning the match"],
        "lost-verb": ["lost to", "were beaten by", "were defeated by", "couldn't beat", "couldn't overcome", "fell to"],
        "losing-verb": ["losing the match", "falling short"],
        "good-adj": ["a fun", "an exciting", "a great"],
        "event-noun": ["day's match", "day's event", "match", "game"],
        "moved-verb": ["has moved the", "moved the", "brings the", "has brought the"],
        "despite-np": ["Despite a strong effot,", "Despite a good performance,", "Despite their best efforts,",
                       "While the Hawks battled hard,", "While the Hawks worked hard,"],
        "result": ["loss", "tough loss", "match", "result"]

    }
```

Results

MEN'S VOLLEYBALL

Box Score

Men's Volleyball Defeated by Loras

LORAS 3   **ILLINOIS TECH** 1

Final	1	2	3	4	T
LORAS (5-2)	44	25	24	25	3
ILLINOIS TECH (3-7)	42	19	26	20	1

Posted: 05-03-2018

Chicago, Ill.(Keating Sports Cente) - The Illinois Tech Men's volleyball team fell to Loras, falling short in 4 sets. Despite their best efforts, the tough loss has moved the Scarlet Hawks to a (3-7) overall record.

Scores

- Illinois Tech 1, Loras 3 (42-44, 19-25, 26-24, 20-25)

How it Happened

- Loras prevailed in set 1 to take the set over Illinois Tech. The Scarlet Hawks held the lead for 18 points in the set. The set ended with a final score of 42-44.
- Loras stood their ground in set 2. Illinois Tech held the lead for 17 points in the set. The set ended with a final score of 19-25.
- In set 3, IIT played hard. They held the lead for 24 points. IIT held the lead to take the set 26-24.
- Loras didn't look back in set 4 to take the set over IIT. The Scarlet Hawks held the lead for 0 points in the set. The Scarlet Hawks couldn't prevail, falling to Loras 20-25.

Scarlet Hawk Standouts

- Lukasz Kupiec had an impressive outing with 18 kills.
- Derek Bostick dominated with 21 digs.
- Ben Peschl led the Scarlet Hawks in assists with 31.

Stats to Know

- IIT earned 72 digs while Loras ended the match with 74.
- Illinois Tech ended up with a team 0.196 attack percentage to Loras's 0.22.
- Loras racked up 6 aces while The Hawks racked up 2.

Up Next

- Illinois Tech will take on Fontbonne at Chicago, Ill. - Keating Sports Center on Sat. February 10, 2018.

GAME LEADERS

LORAS

K: Greg Kinney - 21

B: Ian Walsh - 10

D: Greg Kinney - 18

SA: Mihajlo Gomez - 2

ILLINOIS TECH

K: Lukasz Kupiec - 18

B: Evan Robeck - 4

D: Derek Bostick - 21

SA: Easton Kays - 1

TEAM STATS

LORAS		ILLINOIS TECH
0.22	Hitting %	0.196
13.5	Blocks	8.0
74	Digs	72
6	Aces	2

BOX SCORE

MEN'S VOLLEYBALL

Box Score

Men's Volleyball Battles Strong Opponent

LORAS 3   **ILLINOIS TECH** 1

Final	1	2	3	4	T
LORAS (5-2)	44	25	24	25	3
ILLINOIS TECH (3-7)	42	19	26	20	1

Posted: 05-03-2018

Chicago, Ill.(Keating Sports Cente) - The Illinois Tech Men's volleyball team lost to Loras, losing the match in 4 sets. While the Hawks worked hard, the result has moved the Scarlet Hawks to a (3-7) overall record.

Scores

- Illinois Tech 1, Loras 3 (42-44, 19-25, 26-24, 20-25)

How it Happened

- Loras didn't look back in set 1 to take the set over Illinois Tech. Illinois Tech tried hard, going on a 8-2 run to bring the score to 22-20. IIT couldn't prevail, falling to Loras 42-44.
- Loras prevailed in set 2. IIT held the lead for 17 points in the set. IIT couldn't prevail, falling to Loras 19-25.
- In set 3, The Scarlet Hawks played well. The Hawks went on a 9-1 run to build a 8-9 lead. The set ended with a final score of 26-24.
- Loras held the lead in set 4 to take the set over IIT. IIT played hard, but it wasn't enough to beat Loras. Illinois Tech couldn't prevail, falling to Loras 20-25.

Scarlet Hawk Standouts

- Lukasz Kupiec dominated with 18 kills.
- Derek Bostick tallied 21 digs.
- Ben Peschl (31) led the Scarlet Hawks in assists.

Stats to Know

- Loras had a team 0.22 attack percentage to The Hawks's 0.196.
- Loras had more digs than The Scarlet Hawks with 74 digs while The Scarlet Hawks ended up with 72.
- Illinois Tech had 0 kills to Loras's 0.

Up Next

- The Hawks will challenge Fontbonne on Sat. February 10, 2018 at Chicago, Ill. - Keating Sports Center.

GAME LEADERS

LORAS

K: Greg Kinney - 21

B: Ian Walsh - 10

D: Greg Kinney - 18

SA: Mihajlo Gomez - 2

ILLINOIS TECH

K: Lukasz Kupiec - 18

B: Evan Robeck - 4

D: Derek Bostick - 21

SA: Easton Kays - 1

TEAM STATS

LORAS		ILLINOIS TECH
0.22	Hitting %	0.196
13.5	Blocks	8.0
74	Digs	72
6	Aces	2

BOX SCORE

Results Cont.

MEN'S VOLLEYBALL

Box Score

Scarlet Hawks Sweeps Fontbonne

FONTBONNE | **1** | | **ILLINOIS TECH** | **3**

Final	1	2	3	4	T
FONTBONNE (4-2)	19	19	25	22	1
ILLINOIS TECH (4-7)	25	25	23	25	3

Posted: 05-03-2018

Chicago, Ill.(Keating Sports Centr) - The Scarlet Hawks played a great 4-set match versus Fontbonne, overcoming them 3-1 The game moved the Scarlet Hawks to a (4-7) overall record on the season.

Scores

- Illinois Tech 3, Fontbonne 1 (25-19, 25-19, 23-25, 25-22)

How it Happened

- Illinois Tech took set 1. They went on a 6-5 run to build a 24-19 lead. The Scarlet Hawks held the lead to take the set 25-19.
- In set 2, The Scarlet Hawks fought hard. The Hawks held the lead for 27 points. The Scarlet Hawks held the lead to take the set 25-19.
- Fontbonne didn't look back in set 3 to take the set over Illinois Tech. Illinois Tech fought hard, but it wasn't enough to beat Fontbonne. The set ended with a final score of 23-25.
- In set 4, Illinois Tech took the win. They held the lead for 40 points on their way to victory. Illinois Tech held the lead to take the set 25-22.

Scarlet Hawk Standouts

- Cinjun Coe had an impressive outing with 16 kills.
- Derek Bostick dominated with 13 digs.
- Daniel Throop tallied more assists than anyone else on the Scarlet Hawks with 43 assists.

Stats to Know

- IIT hit 0.213 compared to IIT's 0.084.
- Illinois Tech had more blocks than Fontbonne with 12.0 blocks while Fontbonne earned 8.0.
- The Hawks ended up with 3 aces while Fontbonne earned 3.

Up Next

- The Scarlet Hawks will face Maranatha Baptist on Sat. February 10, 2018 at Chicago, Ill. - Keating Sports Center.

GAME LEADERS

FONTBONNE

K: TJ Vorva - 7

B: Emir Hadzisalihovic - 4

D: Zach Rothstein - 20

SA: Jacob Lendy - 1

ILLINOIS TECH

K: Cinjun Coe - 16

B: Evan Robeck - 10

D: Derek Bostick - 13

SA: Easton Kays - 1

TEAM STATS

FONTBONNE		ILLINOIS TECH
0.084	Hitting %	0.213
8.0	Blocks	12.0
47	Digs	43
3	Aces	3

BOX SCORE

Looking Ahead

- Updating templates
  - Adding new stats
  - Increasing possible noun phrases and adjectives
  - Limiting redundancy
  - Combining simple sentences
- Different sports?