

A Community-Driven Data-to-Text Platform for Football Match Summaries

Pedro Fernandes, Luís António Santos, Sérgio Nunes

INESC TEC · FEUP, University of Porto
ICS · University of Minho

This work describes the design and deployment of Prosebot, a community-driven data-to-text platform tailored for generating textual summaries of football matches derived from match statistics.

Prosebot uses a template-based Natural Language Generation module to produce initial match reports, which are subsequently refined by the reading community.

zerozero.pt is a prominent Portuguese sports portal, offering comprehensive access to match-related data and statistics. It also provides news coverage for premier competitions, curated by a team of journalists.

This coverage encompasses merely 1.4% of the matches cataloged in zerozero.pt’s database.

Monday 12 June 2000 19h45 Philips Stadion (NED) (Enthoven) 31000 Attendance Anders Frisk (SWE) Euro 2000 - Group Stage Group A	
Portugal	3-2 England
22 Luis Figo 37 Joao Vieira Pinto 59 Nuno Gomes	Paul Scholes 3 Steve McManaman 18
Half Time 2-2	
MATCH REPORT LIVE COMMENTARY PERFORMANCE STADIUM REFEREE VIDEOS PHOTOS COMMENTS MATCH REPORT	
PORTUGAL	
1 Vitor Bala (C)	22
2 Jorge Costa	23
5 Fernando Couto	24
13 Dimas	25
14 Abel Xavier	26
4 Luis Vidigal	27
7 Luis Figo	28
10 Rui Costa	29
8 Joao Vieira Pinto	30
17 Paulo Bento	31
21 Nuno Gomes	32
ENGLAND	
1 David Seaman	11
3 Gary Neville	12
4 Phil Neville	13
5 Sol Campbell	14
6 Tony Adams	15
7 Paul Ince	16
8 David Beckham	17
11 Steve McManaman	18
9 Paul Scholes	19
10 Alan Shearer (C)	20
12 Michael Owen	21
SUBS	
12 Pedro Espinha	13 Nigel Martyn
22 Quim	23 Richard Wright
20 Carlos Secretário	6 Martin Keown
16 Beto	12 Gareth Southgate
3 Rui Jorge	15 Gareth Barry
6 Paulo Sousa	17 Dennis Wise
15 Costinha	16 Steven Gerrard
19 Capucho	18 Nick Barmby
9 Sá Pinto	19 Emile Heskey
11 Sérgio Conceição	20 Kevin Phillips
18 Pedro Pauleta	21 Robbie Fowler
MANAGERS	
Humberto Coelho	Kevin Keegan

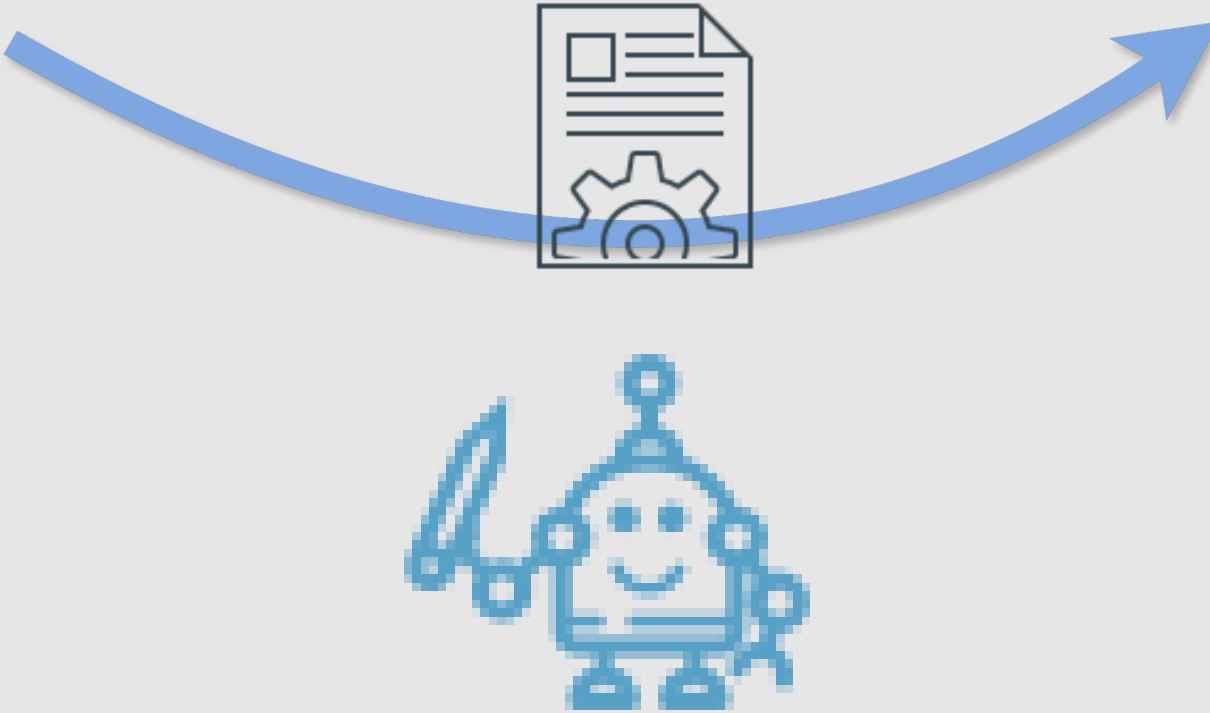
Match Data Tables



Platform Users

1. Identify texts lacking textual reports
2. Start new match reports
3. Generate Prosebot version
4. Validate or change text
5. Publish

Prosebot Draft



Prosebot

Monday 12 June 2000 19h45 Philips Stadion (NED) (Enthoven) 31000 Attendance Anders Frisk (SWE) Euro 2000 - Group Stage Group A	
Portugal	3-2 England
22 Luis Figo 37 Joao Vieira Pinto 59 Nuno Gomes	Paul Scholes 3 Steve McManaman 18
Half Time 2-2	
MATCH REPORT LIVE COMMENTARY PERFORMANCE STADIUM REFEREE VIDEOS PHOTOS COMMENTS MATCH REPORT	
PORTUGAL GRIND OUT VICTORY FOR PORTUGAL FULL OF GOALMOUTH ACTION	
Portugal overcome England	
Portugal triumphed over England, 3-2, on Monday, in a match relating to the Grp A of the European Championship. Both Portugal's national team and Kevin Keegan's team held their first game in the competition.	
Early in the game, Paul Scholes opened the scoring for England, through a header inside the penalty area. Kevin Keegan's team scored their second goal of the game through Steve McManaman, with 18 minutes on the clock, with a right-foot strike. After 22 minutes, Luis Figo fired home for Portugal, with a right-foot shot from outside the box. After 37 minutes, Joao Vieira Pinto struck the second goal for Portugal's national team, through a header inside the penalty box.	
Nuno Gomes scored the winning goal in the 59th minute with a right-foot strike inside the box. Nuno Gomes reaching the landmark of twenty goals for the season.	
Thanks to this result, Portugal are through to the next round of the competition. In relation to forthcoming league matches, the portuguese will play against Romania. Meanwhile, England will play against Germany.	
This article was automatically generated from an algorithm programmed by playmakerstats.com	
Relevant? ***** Quality? *****	

Match Report

Match Data (JSON)

"data": { "info": { "id": "8613676", "senior": 1, "gender": "1", "date": "2022-04-26 20:00:00", "name": "", "full_time": "1", "neutral_field": "0", "id_home_team": "86", "name_home_team": "Manchester City", "type_home_team": "0", "id_away_team": "50", "type_away_team": "0", "name_away_team": "Real Madrid", "home_goals": "4", "away_goals": "3", "home_pen_goals": "0", "away_pen_goals": "0", "referee": "Istvan Kovács", "id_stage": "124977", "name_stage": "MHP", "duration_int": "45", "phase_type": "elim2", "phase_num_hands": "2", "phase_winner": "86", "stadium": "Etihad Stadium", "local": "Manchester", "stadium_capacity": "55097", "winner": "1", "hand": "1", "competition": "Liga dos Campeões",

Template Code

"goal": { "text": "{template.time}, {scorer.name} struck for {team.name}, {template.goal_type}, {template.assisted}.", "condition": "{match_goals=1}" }, "text": "{template.time}, {scorer.name} scored {team.name}'s only goal of the game, {template.goal_type}, {template.assisted}.", "condition": "{match_goals=1}" }, "text": "{team.name} scored their second goal of the game through {scorer.name}, {template.time}, {template.goal_type}, {template.assisted}, {template.commentary_scorer_score}.", "condition": "{team_goal=2}" }, "text": "{template.time}, {scorer.name} opened the scoring for {team.name}, {template.goal_type}, {template.assisted}.", "condition": "{match_goals=1 && match_goals=2}" }, "text": "{template.time}, {scorer.name} fired home {team.name} second goal, {template.goal_type}, {template.assisted}.", "condition": "{team_goal=2}" }, "text": "{template.time}, {scorer.name} netted the final goal of the game, {template.goal_type}, {template.assisted}.", "condition": "{match_goal=match_goals && match_goals=3}" }, }
--

Section	Category	Variants
Title		All-purpose, 2+ goal difference, 4+ goal difference, home team won, away team won, game ended in draw, match ended with penalties
Subtitle		All-purpose, favourite team lost, decisive goal in the last minutes, no goals scored, many goals scored
Small text		Final score, starter and benched relevant players
Introduction		Final score, previous results, best player
Events	Goal	All-purpose, first goal, only goal, own goal, hat-trick, poker, second goal, last goal, goal drew the match, goal increased/decreased the goal difference
	Substitution	Gamechanger player is subbed in, relevant player is subbed in/out
	Missed penalty	Goalkeeper saved, penalty taker missed
	Red card	Direct red card, accumulation of yellow cards
Debriefing		Post-match classification, next games, match stats
Trivia	Stats	Best/worst result of the season for the team, best/worst overall result of the season
	Streaks	Increased or broke a sequence of matches

Table 1: Available templates.

Changes to Initial Drafts

To assess the similarity between the initial drafts generated by Prosebot and the final published texts one month post-launch (174 texts), we used the Dice text similarity metric.

In a large majority of the published texts (70%), the Dice coefficient exceeds 90%, indicating a high degree of similarity between the drafts and the final texts.

Survey to Participants

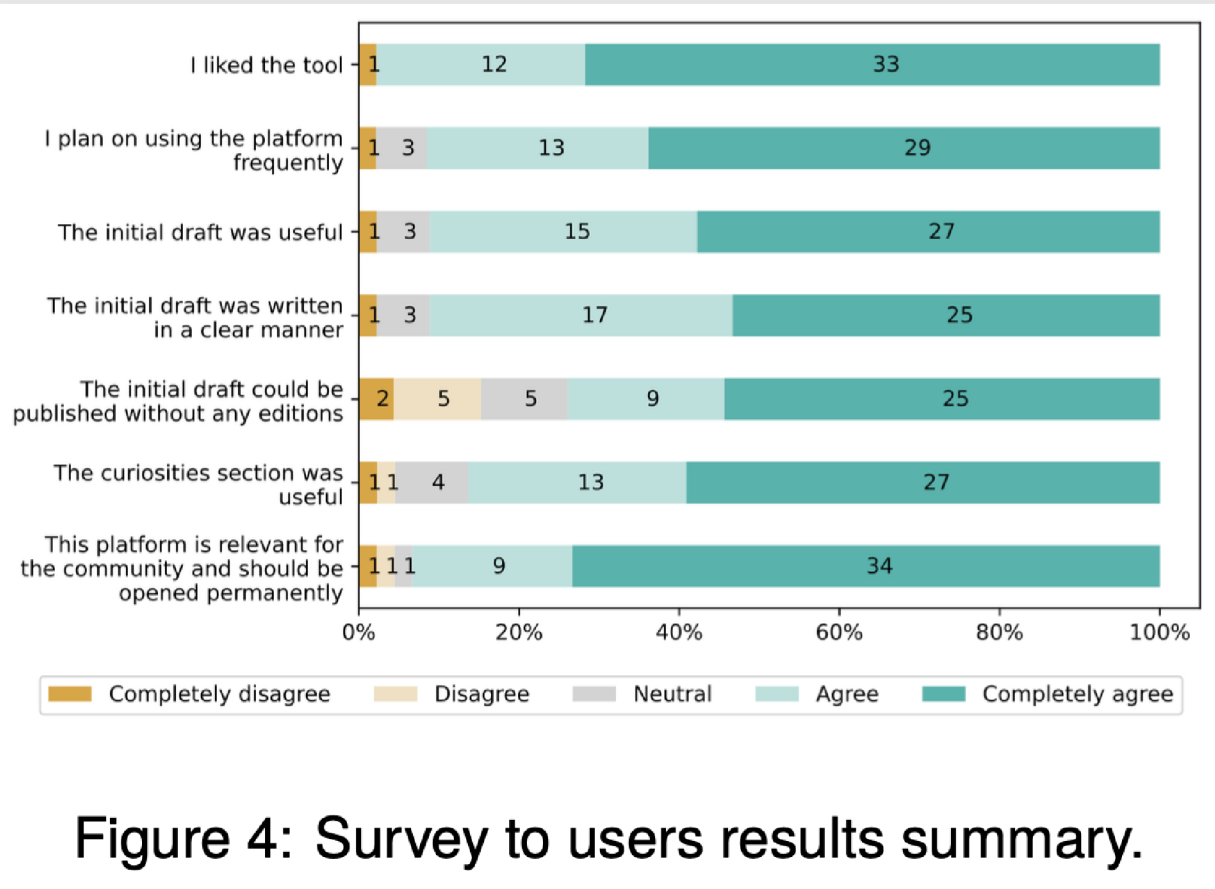


Figure 4: Survey to users results summary.

Platform Engagement

Indicator	May 2021	March 2024
Number of summaries	174	8,079
Number of authors	43	587
Number of teams covered	197	3,116
Number of competitions covered	56	295
Mean summaries per author	4.0	13.8
Mean summaries per team	1.8	5.1
Mean summaries per competition	3.1	27.4
Mean visits per summary	53.0	363.8

Table 2: Prosebot platform key statistics.