InsightChaser: Enhancing Visual Reasoning of Sports Tactical Visualization with Visual-Text Linking

Ziao Liu, Wenshuo Zhao, Xiao Xie, Anqi Cao, Yihong Wu, Hui Zhang, and Yingcai Wu

Abstract—In sports analytics, tactical visualization is widely used to convey valuable insights. However, due to the complex domain knowledge and contextual information involved in tactical visualizations, it is challenging for users to connect high-level tactical insights to corresponding visual patterns. This requires users to engage in a reasoning process to interpret insights within game contexts, which remains insufficiently supported in existing visual-text linking studies. In this work, we propose InsightChaser, a novel approach to bridge tactical insights and soccer visualizations through visual-text linking and visual reasoning enhancement. InsightChaser constructs knowledge graphs to represent both visual elements and contextual game information. Integrating large language models (LLMs), our approach retrieves relevant visual elements and establishes explicit links with insights. Moreover, InsightChaser utilizes LLMs to enhance these visual-text links by providing reasoning explanations and visual effects. We further develop an interactive visualization system that supports navigation and explanation of enhanced visual-text links. Users can explore linked tactical insights interactively and reason through enhanced visual explanations. We conduct two case studies using real-world soccer data and a user study to demonstrate the effectiveness of our approach.

Index Terms—Sports visualization, tactical analysis, visual-text linking

1 Introduction

Sports visualizations can effectively convey tactical insights from complex sports data [39] and are widely used in sports journalism to make such insights interpretable to a broad audience, leading to a more engaging experience. Despite the use of charts to improve readability, most fans still struggle to understand tactical articles. The challenge lies not only in the specialized knowledge required but also in connecting the textual insights to the visual patterns in the charts. Without clear guidance, readers often find it difficult to interpret which elements of the visualization correspond to the tactical explanations.

While recent studies have made significant progress in exploring visual-text interplay for basic chart types [22, 42], their application to sports tactical analysis presents unique challenges. Tactical analysis requires the interpretation of highly specialized domain knowledge that integrates multifaceted elements, including player dynamics, team strategies, and evolving spatial relationships. The complexity is further compounded by the hybrid nature of sports visualizations, which combine conventional graphical elements with domain-specific contextual layers. For example, a soccer passing network merges standard node-link diagrams with the spatial context of a pitch, resulting in composite representations that demand novel approaches for linking tactical insights with heterogeneous visual components.

In this work, we propose InsightChaser, a novel method to establish deeper connections between sports visualizations and tactical insights. We select soccer as our scenario for its complex tactical nature. InsightChaser allows users to interactively explore the insights with corresponding visual elements. Users can further investigate detailed explanations of tactical insights and corresponding visual effects, transparently revealing the underlying analytical reasoning processes. During development, we faced two key challenges.

Bridging tactical insights with visual evidence. Soccer visualizations encode complex spatial and contextual relationships (e.g., player positioning, tactical formations), yet clearly decoding this information based on the textual insight is difficult due to the abstractness of text. For example, tracing how abstract conclusions (e.g., "exploited the left flank") map to specific visual patterns (e.g., features of passing networks or heatmaps) requires joint interpretation of team perspective (i.e., which team's viewpoint is being described), team strategies, and player roles. However, this contextual information is often omitted (treated as prior knowledge) or fragmented in the text. To bridge this gap, structured representations of insights and visual elements are thus needed to unify domain knowledge with visual-textual correspondence.

Elaborating tactical insights with visual enhancement. Due to limited knowledge, users often struggle to comprehend how tactical insights are derived from visual-text links. It requires clear analytical paths to follow. However, these insights often lack explicit reasoning paths, making it challenging to explain the complex analytical process. Moreover, users need intuitive presentations to navigate the linked elements and understand their role in supporting the insight. However, the multiple visual elements and associated insights significantly increase the presenting complexity of the relevant information.

To address the first challenge, we construct a knowledge graph to represent visual elements and contextual information in a structured manner. With powerful semantic understanding and rich knowledge, LLMs are used to interpret high-level tactical insights while retrieving the entities and relationships stored in the knowledge graph to perform visual-text linking. To tackle the second challenge, we further utilize LLMs to explain the insights with analytical paths and enhance them with corresponding visual effects according to the reasoning process. Moreover, we design a visualization system to support the navigation and explanation of linking and enhancement results. This allows users to interactively explore tactical insights and understand their analytical reasoning through linked visual elements and enhanced explanations.

In summary, our main contributions are as follows:

- An **extensible pipeline** that supports linking tactical insights with soccer visualizations and provides further reasoning enhancement based on knowledge graphs and large language models.
- An **interactive visualization system** supporting the exploration and explanation of tactical insights with enhanced visualizations.
- **Case studies and a user study** that demonstrate the effectiveness of our linking and enhancement approach.

2 Related Work

2.1 Sports Data Visualization

Visualization serves as a crucial bridge between complex data and valuable insights in sports analytics, drawing significant research interests [39]. Researchers employ visualization techniques to support diverse analytical tasks and address domain-specific needs across var-

- *Z. Liu, X. Xie, A. Cao and H. Zhang are with the Department of Sports Science, Zhejiang University. E-mail: {liuziao, xxie, caoanqi, zhang_hui}@zju.edu.cn. Z. Liu is also with the State Key Lab of CAD&CG, Zhejiang University. X. Xie is the corresponding author.*
- *W. Zhao, Y. Wu and Y. Wu are with the State Key Lab of CAD&CG, Zhejiang University. E-mail: {zhao_ws, wuyihong, ycwu}@zju.edu.cn.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

ious sports [13, 55], such as table tennis [19, 49], badminton [10, 25], baseball [12, 34], basketball [16, 48], and swimming [52].

Soccer, celebrated for its blend of teamwork and individual skill, has captivated researchers worldwide. Existing studies have explored specific visual encoding and designed interactive visualization systems to gain insights and address analytical challenges [7, 39, 50]. For team tactic analysis, SoccerStories [38] utilized facet visualization to present metadata and spatio-temporal data from various perspectives. Beyond coaching applications, visualizations have been applied to assist players in improving their performance. For instance, PassVizor [51] modeled players' event sequences and introduced an interactive system to display variations in passing actions and the evolution of passing patterns throughout a game. Similarly, Action-Evaluator [8] proposed tailored visualizations to connect rich player actions with complex match situations. Andrienko et al. [2] and Seebacher et al. [41] focused on player trajectory analysis and designed visualizations to reveal player movement patterns.

Sports data analytics are not limited to match analysis for experts but are also prominently featured in leading sports news outlets, such as Opta and ESPN, for engaging sports fans. These articles utilize data visualization as evidence and implement data-driven storytelling to effectively communicate insights to their readers [17]. Current research focuses on streamlining the authoring process of these reports. For instance, Fu et al. [15] developed interactive visualization systems, namely GameViz and LineupViz, which assisted NBA writers in delving into analytical insights within extensive datasets. Additionally, SNIL [9] introduced a pipeline that integrates narrative generation with data chart mapping, providing an automated and structured workflow for data report creation. Furthermore, efforts have also been directed towards creating more impactful visual effects. For example, Sporthesia [58] extracted entities from commentary and employed enhanced encoding techniques to significantly augment the visual impact of sports videos. These methods focus on helping analysts create data-driven sports articles. However, improving readers' understanding of these articles by connecting textual tactical insights and sports visualization charts remains an unsolved challenge. To address this challenge, our research introduces an alignment method that enables interactive exploration of the relationship between visualization and textual insights, enhancing users' understanding of complex sports analytical outcomes.

2.2 Visual-Text Linking

The interplay between visualization and text effectively reveals hidden insights within complex data [21, 46, 53]. Extensive research on visual-text linking has inspired numerous downstream applications, such as data storytelling [43], animation generation [42, 54], visual question answering [45], and text annotation [20].

Many studies have designed interactive interfaces to facilitate visual-text linking [3, 31, 57]. For example, Kori [22] explored the design space of the chart-text relationship and developed an authoring tool that allows users to interactively create links. Similarly, Charagraph [31] designs different interaction contexts across three representation levels: chart in text, separate, and text in chart, forming an interactive system. Additional research has introduced innovative interactive methods to enrich user exploration. InkSight [27] offers a sketch-based interaction that allows users to interact directly with charts to analyze interesting areas in computational notebooks. Furthermore, structural information of visualizations tends to yield better performance in linking tasks [24, 56]. For instance, Data Playwright [43] employs filtered SVG abstraction as visual input to better capture associations between charts and natural language commands. Similarly, Cai et al. [6] constructed a knowledge graph to represent relationships between data objects and visual elements, enhancing the effectiveness of visual linking.

LLMs, with their extensive domain knowledge and powerful multimodal comprehension capabilities, have also attracted significant research interest to advance visual-text linking. Recent studies have leveraged LLMs to establish deeper connections between textual descriptions and visual data. FinFlier [18] utilizes the financial knowledge-grounded LLMs to bind text and data, adding graphical overlays to data charts. Data Playwright [43] applies prompt engineering to filter visual elements based on narration and data tables. These approaches primarily focus on various types of charts. However, sports visualizations extend beyond conventional chart-based representations and inherently incorporate contextual elements, such as the field and player position. Although Metoyer et al. [32] explored a valuable method of visual-text linking in sports, their approach lacked sufficient consideration of contextual information. To address this limitation, our research utilizes LLMs and integrates knowledge graphs to link contextual elements and design an interactive system to enhance sports visualizations.

3 Background

In this section, we first introduce the background and key concepts relevant to our work. We then describe a preliminary study conducted to identify users' challenges and summarize users' requirements.

3.1 Background and Concept

Soccer is a team sport where eleven players collaborate on the pitch, aiming to score goals and secure victory. Player decisions are highly dependent on the spatial awareness, which has led to the widespread use of pitch-based visualizations that integrate contextual information to present data features, such as passing frequency or shot distributions. Analysts and journalists reveal and interpret these patterns using their domain knowledge, summarizing them into concise descriptions known as tactical insights [9]. Thus, we focus on linking these tactical insights with their corresponding visualizations. The primary contextual information involved is described as follows.

- **Pitch zones** serve as fundamental units for organizing tactical play and execution, and can be categorized into standard zones (e.g., the penalty area) and tactical zones. Existing research [14] defined tactical zones as three vertical zones (defensive third, middle third, and attacking third) and three horizontal zones (left wing, central area, and right wing) parallel to the touch line.
- **Player positions** define the roles of players within a team, which are generally classified into goalkeeper, defender, midfielder, and forward. They can be further specified based on field placement, such as left-back, right-forward, etc.
- **Formation** establishes a team's lineup and playing style in the game, defining players' distribution on the pitch. For example, a 4-4-2 formation includes four defenders, four midfielders, and two forwards, shaping both offensive and defensive strategies.

3.2 Preliminary Study

To explore the challenges users face when interpreting soccer data visualizations and analytical text, we conducted a preliminary study. First, we collected visualization-text pairs from various online analysis articles. Then, we categorized and analyzed the dataset based on the chart type and content type. Finally, we conducted interviews with soccer fans to summarize and confirm our findings.

Collection. We collected 204 pitch-based visualizations from soccer tactical analysis articles published between January 2024 and February 2025 using Google Search. To minimize bias and improve diversity, we sourced articles from various top leagues (e.g., *the Premier League*) and major continental and international tournaments (e.g., *Copa América* and *the Olympics*) across multiple sports media outlets, such as Opta, The Athletic, Total Football Analysis, and Cannon Stats. For each visualization, we extracted the corresponding analytical text to form visualization-text pairs and recorded the title, website, and publisher.

Analysis. We conducted a quantitative analysis of our dataset, beginning with a systematic review of the collected visualizations. This review focused on key aspects, including information representation, visualization techniques, and analytical processes. To further categorize the visualization-text pairs, we recruited three volunteers with expertise in both data visualization and soccer analytics. They were first introduced to a structured annotation phase to establish a shared understanding of the categorization criteria. The volunteers then independently labeled each pair based on two dimensions: chart type (for data visualization) and content type (for analytical text). Any discrepancies were discussed and resolved through a consensus process.

We analyzed the distribution of chart types (Fig. 1(A)) and content types (Fig. 1(B)) separately. Node-link diagrams, scatter plots, and heatmaps were the three most common, aligning with standard visualization formats. Arrow-based representations were also prevalent, mainly for pitch trajectory visualization rather than traditional charts.

Some visualizations combined multiple chart types. Moreover, as most of these visualizations were event data-based, we classified content into specific action types such as pass, shot, carry, and touch.

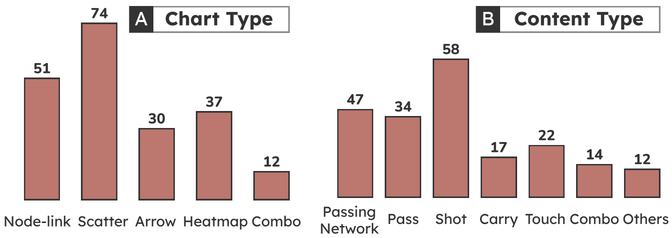

Fig. 1: Statistics of the collected data. (A) shows the distribution of chart types. (B) presents the distribution of content types.

Interview. We conducted semi-structured interviews with six soccer fans recruited via public online postings, including four novice readers of tactical analysis (P1-P4, casual viewers with no formal background) and two experienced readers (P5 and P6, both PhD researchers with expertise in soccer data analysis). The interviews began with general questions about participants' habits and challenges in interpreting tactical visualizations. Next, we presented them with a sample of visualization-text pairs featuring different chart types and content types. Participants were instructed to interpret the visualizations using the corresponding text, following a think-aloud protocol. Finally, we conducted follow-up discussions to identify where they struggled and why. The experienced readers were additionally invited to reflect on their initial learning experiences and the efforts they took to become proficient. We recorded their thought process and discussions and summarized them as the following key findings.

F1 Visualization type affects comprehension. All participants agreed with the classification of chart type and context type, while highlighting the impact of chart type on comprehension. Arrow-based visualizations were the easiest to interpret due to their simple and intuitive design. In contrast, scatter plots and heatmaps posed greater challenges, with node-link diagrams being the most difficult to understand. P3 noted, *"Node-link diagrams are common in video analysis, but their high information density makes them harder to process."* P5 added, *"Combo visualizations help provide more data features to improve understanding, but too many overlapping elements make the visualization overwhelming."*

F2 Text presents conclusions with limited reasoning. Text typically highlights the key takeaways of articles but often omits the underlying reasoning, creating a gap between text and visualization. Readers are thus required to infer the reasoning process themselves to achieve a comprehensive understanding. P1 noticed, *"While I can understand the textual insights, I want to see the thought process and find supporting evidence in the visualization."* P5 reflected, *"I often had to look up multiple articles with similar types of analysis and go back and forth between the text and the visuals to figure out how the conclusions were being drawn."*

F3 Visualizations provide limited contextual information. Soccer analysis often includes additional contextual details (e.g., a player's position or a team's formation), which are typically absent from the visualization. This omission makes it harder for readers to fully understand the tactical context. P4 highlighted, *"A player's position can change from match to match, which affects their role. But that kind of information doesn't show up in the visualization."* Furthermore, terminology in text complicates comprehension. P3 mentioned, *"I only have a rough idea of terms like defensive third and midfield, so it's hard to fully understand what they mean."*

F4 Locate visualization with related text needs effort. Participants reported difficulty in identifying and interpreting visual elements referenced in the analysis text. P2 noted, *"The text often points to just one part of the visualization, so I have to search for it myself. When there are multiple things to find, it takes even more time."* P6 reflected, *"I struggled to map terminology in the insights to the visual elements. Terms like 'defensive line' or 'half-space' didn't immediately translate into what I saw on the chart."*

3.3 Requirement Analysis

Based on the findings of our preliminary study, we summarize the requirements into six key aspects and categorize them into visual-text linking (R1-R3) and visual-text enhancement (R4-R6).

R1 Integrate domain knowledge. Domain knowledge encompasses a deep understanding of soccer terminology, tactics, and other analytical concepts. Integrating relevant knowledge can help clarify technical terms, bridging the semantic gap between expert analysis and reader comprehension effectively.

R2 Incorporate contextual information. Soccer analysis relies heavily on contextual details such as pitch zones and player positions, which are closely linked to a player's expected performance. Including this information provides a more comprehensive analytical and reading experience.

R3 Highlight linked elements. In soccer analysis, both text and visualization convey extensive information. Highlighting relevant parts in both can filter out unrelated details and enhance visual clarity to help users focus on key insights.

R4 Present reasoning process. Insights are often concluded from data exploration and observation. Providing the reasoning behind these insights helps users better understand the analytical process, making it easier to connect text and visualization.

R5 Integrate additional visual effects. Text often provides insights beyond what a single visualization can show. To better support analysis, additional visual elements should be added when needed. For example, adding a player's heatmap in a passing network can better illustrate his positional coverage.

R6 Enable interactive exploration. Users have diverse preferences when engaging with visual-text analysis. Introducing interactive exploration allows them to navigate visualizations based on their interests, enhancing the analytical experience and efficiency.

3.4 Data Description

This work focuses on three common soccer visualization types: passing networks (node-link diagrams), action heatmaps, and shot maps (scatter plots). Since the collected visualizations were not fully reproducible due to limited data access, we constructed visualizations using StatsBomb open data [44] from 83 matches involving 40 teams in *Copa América 2024* and *EURO 2024*. Each match consists of two main data components: metadata and event data. Metadata provides contextual information such as player lists, positions, and team formations, while event data records detailed in-game actions (e.g., passes, shots, and carries) with timestamps, locations, involved players, and outcomes. The visual encoding and data processing approaches are presented below. All data-bound SVG elements are denoted using the <*> format.

- **Passing network** visualizes the team and player passing patterns in each match across the full pitch (<path>). Each circle (<circle>) marks a player's average passing location, and line (<line>) widths represent the number of passes between players. We extract pass events before the first substitution, compute each player's average coordinates for both passing and receiving, and count pairwise passes.
- **Shot map** visualizes each team's shot distribution throughout all matches in the attacking half (<path>). Shots are represented by <circle> elements: open circles for goals and crosses for misses. Shot events are extracted from event data, mapped to a standardized attacking direction, and positioned using their starting coordinates with shooter names from metadata.
- **Action heatmap** visualizes the intensity of player or team actions in each match across the full pitch (<path>). The pitch is divided into an 8×12 grid, and the locations of actions (e.g., carry, pass) are extracted and aggregated within each cell, which is represented as a rectangle (<rect>) in the SVG. The resulting counts are then normalized to compute action intensity and encoded using a color gradient.

Each SVG visualization and its relevant metadata serve as the input into GPT-4o [35] to generate tactical insights. Following the analytical perspectives (identification, comparison, and summarization) summarized in existing studies [4, 9, 26], for each tactical visualization, the LLM is prompted to generate both player-level and team-level insights under each perspective. As shown in Fig. 2, the insight *"Montiel frequently supports attackers."* serves as a player-level summarization. The prompt template is provided in the supplementary materials.

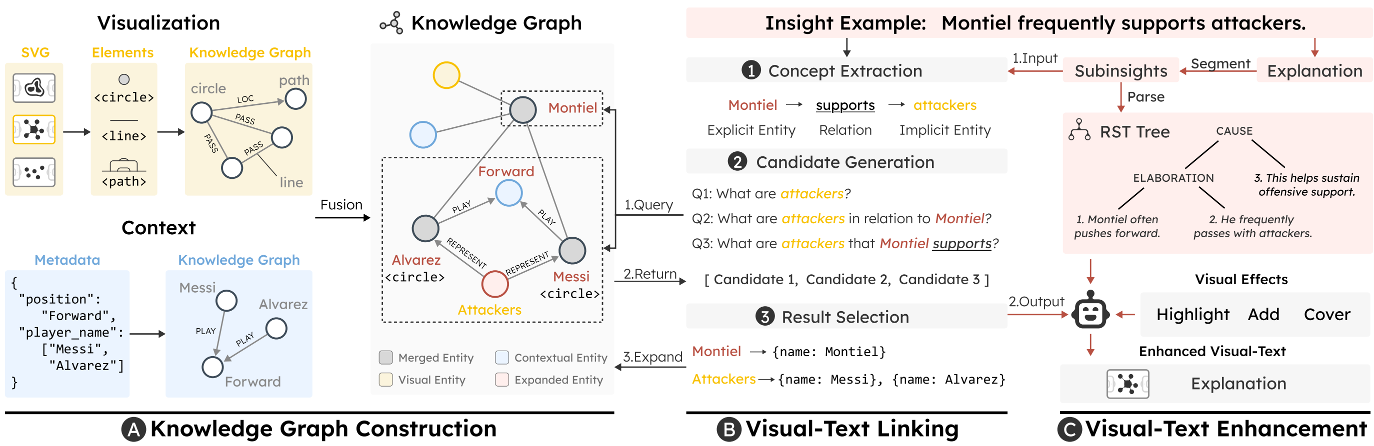

Fig. 2: The pipeline of InsightChaser. (A) Knowledge graph construction from visualization and context separately, and fusion into a unified knowledge graph. (B) Visual-text linking performs an LLM-based linking process, including concept extraction (B1), candidate generation (B2), and result selection(B3). (C) Visual-text enhancement generates explanatory reasoning paths, segments into subinsights, builds an RST tree, and applies visual effects to form an enhanced visual reasoning path.

4 InsightChaser

In this section, we propose InsightChaser, a pipeline designed to support the linking and enhancement of tactical insights and sports visualizations. InsightChaser consists of three main components (Fig. 2). In knowledge graph construction, we structure both visualization and contextual information into a unified knowledge graph. In visual-text linking, we introduce a multi-step LLM-based analysis to link insights to corresponding entities in the knowledge graph. Finally, in visual-text enhancement, we generate explanatory reasoning of insights and apply visual effects to form an interpretable visual reasoning process. We employ Neo4j [1] as our graph database and GPT-4o as the LLM.

4.1 Knowledge Graph Construction

Visual-text linking requires effectively capturing relationships among visual elements [6]. However, sports tactical visualizations often encode complex spatial and contextual information, resulting in intricate relationships among these elements. To address this, InsightChaser first constructs a separate KG for each tactical visualization. Specifically, it independently builds a visualization KG from visual contents and a context KG from metadata, which are then fused into a unified structure (Fig. 2(A)). This structured representation lays a solid foundation for subsequent visual-text linking and reasoning tasks.

KG of visualization. To construct a structured representation of tactical visualizations, we first extract SVG elements from the original visual content (as detailed in Sec. 3.4), including <circle>, <line>, <path>, and <rect>. These graphical elements are then assigned to entity types based on their semantic roles in different tactical visualizations (e.g., <circle> corresponds to entity *Player* in passing networks and to entity *Shot* in shot maps). Relevant attributes are also extracted, including average player location and pass count in passing networks, shooter, result, and location in shot maps, action intensity in heatmaps, and region bounding boxes for standard pitch zones (e.g., the penalty area). The entities, attributes, and relations are summarized in Table 1.

KG of context. Contextual information is organized into a knowledge graph to enable integration with visual elements. We define five types of entities: *Player*, *Team*, *Position*, *Formation*, and *Tactical_Zone* (e.g., middle third-central area). Each entity is assigned a *name* attribute extracted from the metadata, and for each *Tactical_Zone*, we additionally include a *region* attribute representing its bounding box. Relationships are used to connect semantically related entities. For instance, *Player* entity "Messi" and "Alvarez" are connected to *Position* entity "Forward" through the *Play* relationship (Fig. 2(A)). The detailed entities and relations are shown in Table 1.

KG fusion. We fuse the two separately constructed KGs for each visualization based on the information summarized in Table 1. For the passing network, KG_p and KG_c are merged by aligning the *name* attribute of E_{player}. For the shot map, each E_{shot} in KG_s is linked to a player in KG_c by matching the *shooter* name with the *name* attribute of E_{player} in KG_c. For the action heatmap, each E_{cell} is linked to either

Table 1: Entities (E), attributes, and relations (R) defined in each KG. KG_p, KG_s, and KG_a are constructed from the passing network, shot map, and action heatmap, while KG_c is constructed from contexts. Attributes are denoted using variable names. In particular, *region* specifies the bounding box $(x_{min}, y_{min}, x_{max}, y_{max})$ of each pitch zone. Relations are expressed as directed triples $(E_1, R_{1\to 2}, E_2)$, except for R_{pass}, which is undirected to represent passing connections between players.

KG	Entities (E): Attributes	Relations (R)
KG_p	E_{player}: $name, x_{avg}, y_{avg}$ E_{line}: $count, source, target$ E_{zone}: $name, region$	$(E_{player}, R_{loc}, E_{zone})$ $(E_{player}, R_{pass}, E_{player})$
KG_s	E_{shot}: $shooter, result, x, y$ E_{zone}: $name, region$	$(E_{shot}, R_{loc}, E_{zone})$
KG_a	E_{cell}: $intensity, x, y$ E_{zone}: $name, region$	$(E_{cell}, R_{loc}, E_{zone})$
KG_c	E_{player}: $name$ E_{team}: $name$ $E_{position}$: $name$ $E_{formation}$: $name$ $E_{tactical\_zone}$: $name, region$	$(E_{team}, R_{have\_player}, E_{player})$ $(E_{player}, R_{play}, E_{position})$ $(E_{team}, R_{deploy}, E_{formation})$

E_{player} or E_{team}, depending on whether it represents player or team actions. In all cases, $E_{tactical\_zone}$ entities are added to the existing E_{zone} set, and the relations are recomputed based on the updated zones.

4.2 Visual-Text Linking

Most visual-text linking methods focus on visual patterns (e.g., trends or extrema in line charts) derived from data tables. However, tactical insights are semantic summaries without explicitly defined data features. To address this, we incorporate LLMs to semantically interpret these abstract insights (Fig. 2(B)). Specifically, the linking task is divided into three subtasks: concept extraction, candidate generation, and result selection. Guided by structured prompts, the LLM first identifies explicit and implicit entities in concept extraction, then queries the KG for linking candidates during candidate generation, and finally selects the best matches to form text-graph pairs as linking results. Prompt templates are provided in the supplementary materials. Examples of visual-text linking results are shown in Fig. 3.

Concept extraction. Given the KG schema and the tactical insight as input, the LLM is guided through a three-step reasoning process within a single prompt. It first identifies explicit entities based on the schema (e.g., "Montiel" is identified as a *Player* entity), then extracts

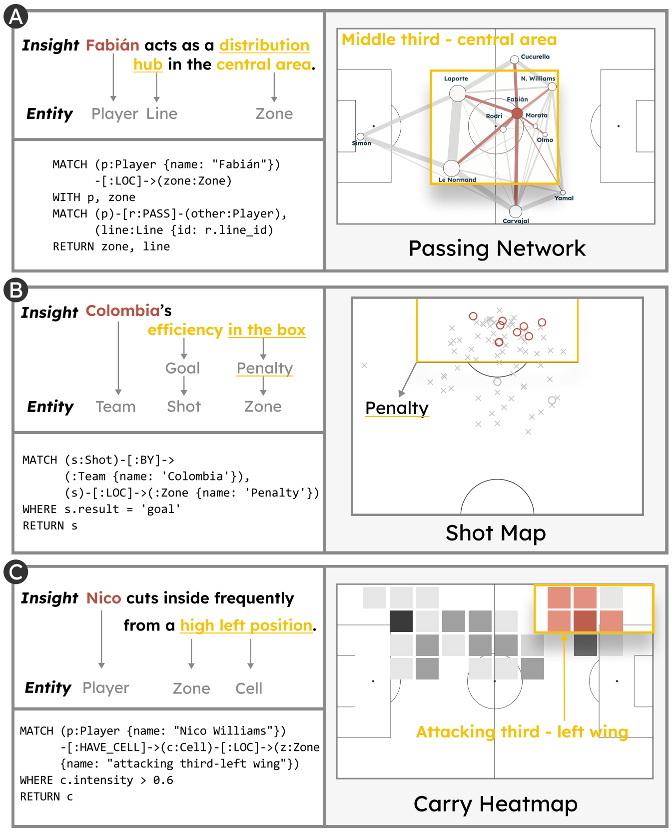

Fig. 3: Examples of visual-text linking. (A) presents a passing network. (B) displays a shot map. (C) presents a carry heatmap.

implicit entities, which are abstract or context-dependent concepts not directly present in the KG (e.g., "attackers"), and finally establishes their semantic relations with the extracted explicit entities (Fig. 2(B1)). All entities and relations are returned in a structured format.

Candidate generation. To link extracted implicit entities to relevant nodes or subgraphs in the KG, we use predefined templates to dynamically generate natural language questions for each extracted implicit entity. Each question follows a fixed structure, with the implicit entity, explicit entity, and their relation filled into placeholders. As shown in Fig. 2(B2), Q1 asks for the definition, Q2 explores its semantic relation to the explicit entity, and Q3 focuses on the contextual interaction between both entities. The LLM first interprets each question using domain knowledge, then translates it into executable Cypher to query the KG, and finally generates explanations for each returned candidate. This process ensures that all suggested entities are grounded in the KG.

Result selection. To determine the final links for each implicit entity, the LLM is prompted to reason over the candidate entities using domain knowledge and select the most relevant one. Based on the LLM's selection, we structure the output as text-graph pairs, where each textual entity is linked to its corresponding entity or subgraph in the KG. These text-graph pairs are then incorporated into the existing graph for further visual reasoning.

4.3 Visual-Text Enhancement

Visual-text linking helps users capture key insight concepts in tactical visualizations. However, clear analytical paths and relevant contexts are often needed to better understand the summarized insights. To address this, we introduce a visual-text enhancement process consisting of two steps: textual reasoning and visual mapping. For each insight, textual reasoning generates an explanatory narrative, segments it into subinsights, and organizes them into a reasoning structure based on Rhetorical Structure Theory (RST) [29]. Visual mapping then assigns appropriate visual effects to each subinsight to support visual reasoning.

Textual reasoning. The reasoning process significantly improves users' comprehension of tactical insights. We first prompt the LLM to generate explanatory reasoning based on the existing links between insights and the knowledge graph. To structure the explanation, we adopt RST to present the semantic relationships between subinsights in a hierarchical form. Specifically, we segment the explanation into coherent subinsights [33] and employ the fine-tuned LLaMA-2-70B model from Maekawa et al. [28] to apply a top-down parsing approach to construct RST trees. In this tree, the leaf nodes represent segmented subinsights, while the internal nodes capture rhetorical relations among them. For example, as shown in Fig. 2(B3), the explanation is segmented into three numbered subinsights. The first subinsight introduces Montiel's forward movement. The second is linked to the first via the "ELABORATION" relation, describing his passing connections. These two are jointly linked to the third subinsight through the "CAUSE" relation, which introduces Montiel's tactical contribution.

Visual mapping. Each subinsight is treated as a normal insight to undergo the visual-text linking process for text-graph pairs. These pairs and the RST tree are provided as input to the LLM via structured prompts to determine appropriate visual effects. We formulate various visual enhancement effects as callable functions, such as highlight, add, and cover. Specifically, the highlight function emphasizes existing elements in the visualization, such as visual markers (e.g., players or shots), events (e.g., passes), or important pitch zones. The add function introduces additional contextual information that may be implicit or missing, such as filters on visual elements, player positions, or team formations. Lastly, the cover function overlays new visual representations onto existing content. For example, the distribution of actions, key space on the pitch, or action heatmaps. These visual effects are then applied according to the structure of the analytical path, enabling intuitive visual reasoning.

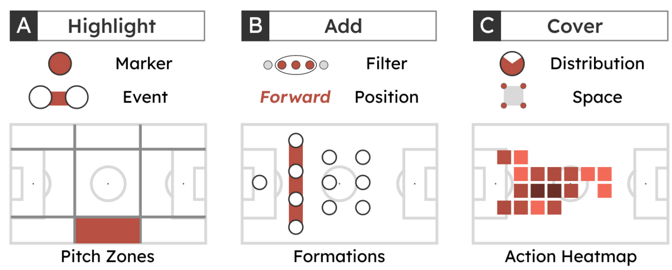

Fig. 4: Visual effects used in our approach. (A) depicts the highlight type. (B) illustrates the add type. (C) displays the cover type.

5 Visual Design

Our system is designed based on the summarized requirements of casual fans and contains three views: linking view, enhancement view, and edit view. Users can explore and establish visual-text links in the linking view (Fig. 5(A)). Then users can navigate these links and obtain visual reasoning explanations in the enhancement view (Fig. 5(B)). Users can further customize visual effects and chat with LLMs in the edit view (Fig. 5(C)).

5.1 Linking View

The linking view (Fig. 5(A)) consists of three components. The insight view allows users to select textual entities for exploration (**R3**). The context view provides relevant contextual information to assist users in the linking process (**R1, R2**). The visualization view displays corresponding visual elements associated with selected insights (**R3**).

The insight view (Fig. 5(A1)) presents textual insights. We use distinct colors to represent different entity types to enhance readability: red for explicit entities and yellow for implicit entities. We also provide a corresponding legend at the top-right corner for clarity.

The context view (Fig. 5(A2)) uses a node-link diagram to represent a knowledge graph, which provides relevant contextual information and associated visual elements. Different colors are applied to distinguish entities clearly. Directed and undirected links represent relationships between entities. Additionally, node dragging and filtering are implemented to facilitate user exploration of the knowledge graph.

The visualization view (Fig. 5(A3)) incorporates three main types of tactical visualization: passing networks, shot maps, and action

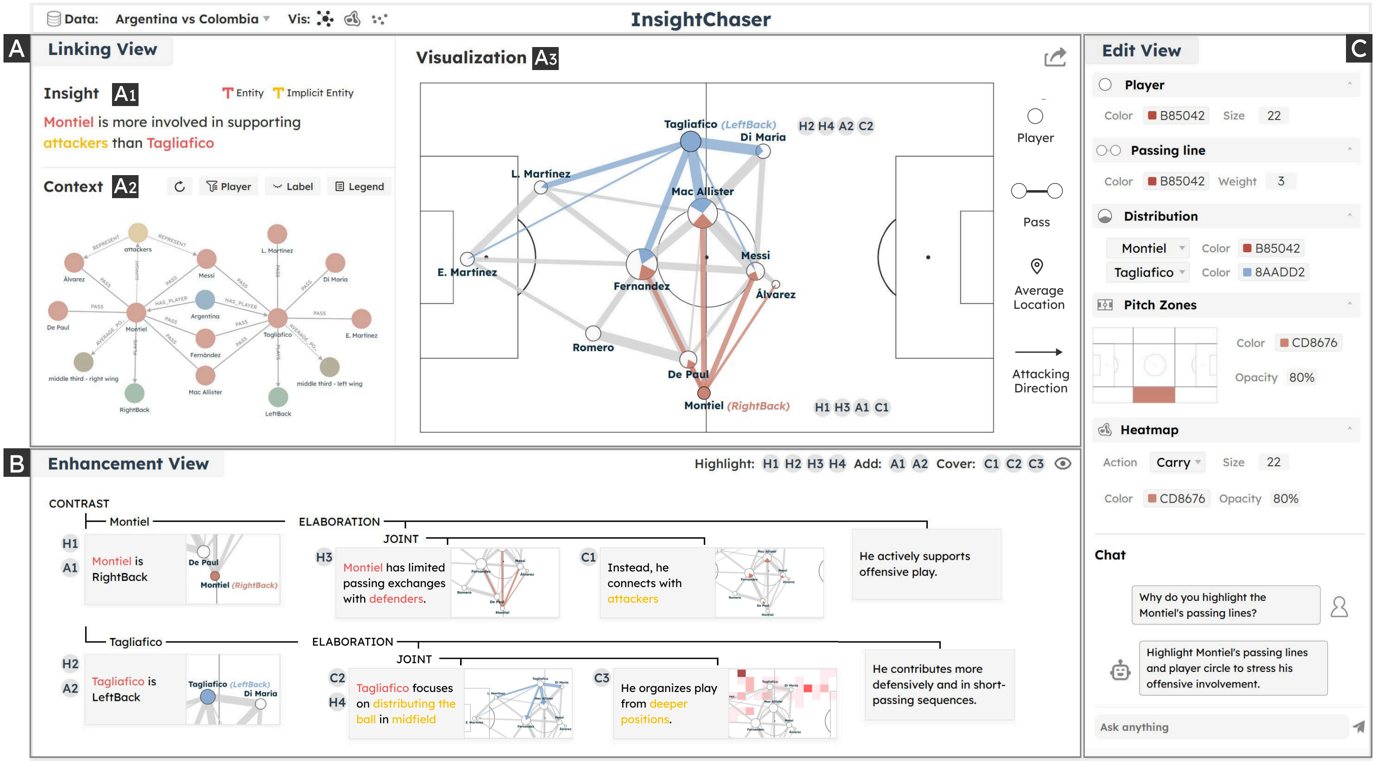

Fig. 5: The system interface of InsightChaser. (A) The linking view supports users in exploring visual-text links with highlighted tactical insights, contextual knowledge graphs, and corresponding visual elements. (B) The enhancement view explains tactical insights with a structured reasoning path and corresponding visual effects. (C) The edit view allows users to edit visual effects' styles and interact with LLMs for a deeper understanding.

heatmaps. Each visualization is implemented by following established conventions and widely adopted design practices. The visual encoding for these tactical visualizations is detailed in Sec. 3.4. To help users capture the visual context quickly, a clear title combined with the team name and visualization type is provided. The legends of visual elements and necessary contexts are placed to the right for clarity.

Interactions of the linking view are as follows.

- *Hover over entities.* Users can hover over the colored entities in the insight view to highlight relevant subgraphs in the context view and linked visual elements in the visualization view.
- *Add new links.* Users can first brush over unlinked text and then select one or more relevant nodes in the context view or visual elements in the visualization view to establish the visual-text link.
- *Hover over contexts.* Users can hover over nodes or edges in the context view to reveal detailed properties.
- *Filter contexts.* Users can click the filter button in the context view and select specific entity types to refine the displayed contexts.
- *Reset contexts.* Users can click the reset button in the context view to restore the default layout.

5.2 Enhancement View

The enhancement view (Fig. 5(B)) presents a visual reasoning path based on visual-text links (**R4**). The reasoning path follows an RST tree structure, with each leaf node represented by a card that displays linked textual explanations and enhanced visual effects (**R5, R6**).

RST tree organizes the reasoning steps into a hierarchical tree structure. We adopt a horizontal layout to make efficient use of vertical space. Rhetorical relationships between explanations are indicated using uppercase text labels, while straight lines illustrate the branching structure. Nodes aligned at the same horizontal level represent elements at the same rhetorical layer within the tree. To provide users with contextual guidance, we insert a designated node beneath the root to briefly introduce the relevant player or team. This representation provides users with an intuitive navigation of the reasoning process.

Reasoning cards fill each leaf node of the RST tree. Each card has two parts. The left part contains a linked textual explanation. For visual consistency, linked entities are labeled using the same visual encoding as in the insight view. The right part displays a thumbnail preview of the applied visual effects. To effectively highlight relevant visual elements, the visualization automatically zooms to the appropriate canvas area, forming a coordinated overview-detail [11] display in combination with the linking view. Moreover, we introduce effect labels to denote the specific type of visual effects, numbered consistently within the visualization view, allowing users to associate the detailed explanation with the resulting visual enhancements.

Interactions of the enhancement view are as follows.

- *Hover over cards.* Users can hover over individual cards to highlight relevant subgraphs in the context view and corresponding visual effects in the visualization view.
- *Select cards.* Users can click on cards to activate or deactivate visual effects in the visualization view.
- *Hover over effect labels.* Users can hover on effect labels to display tooltips, providing specific applied visual effects.
- *Hide effect labels.* Users can click the hide button to toggle the visibility of all visual effects and their corresponding effect labels in the visualization view.

5.3 Edit View

The edit view (Fig. 5(C)) enables users to interactively customize the visual effects applied to selected visual elements. To reduce the potential confusion caused by the excessive use of colors in visual effects, we allow users to exert fine-grained control over individual visual elements. The edit view lists multiple visual elements as editable objects, each offering specific style parameters tailored to their characteristics for precise adjustment. Additionally, the chat view allows users to interact directly with LLMs. Users can engage in conversations to search for unfamiliar terminology and quickly access related domain knowledge. They can also explore contextual information like the team's playing styles with richer explanations.

Interactions of the edit view are as follows.

- *Pick colors.* Users can select and apply colors using the color picker in the edit view to customize visual elements.
- *Input queries.* Users can type questions to interact with LLMs and receive responses in the chat view.

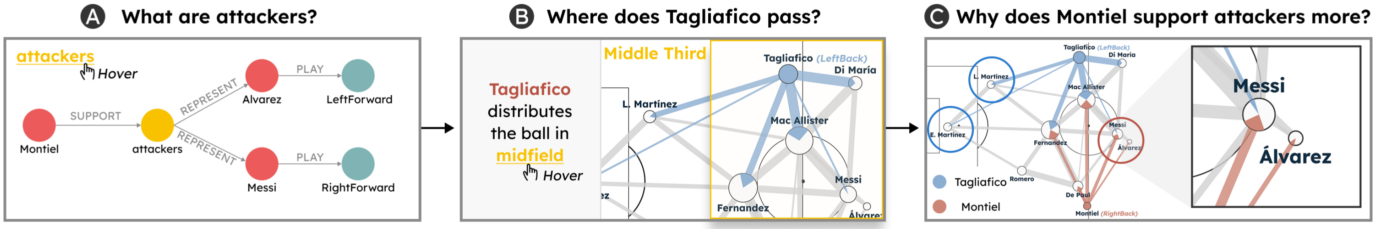

Fig. 6: The pipeline of our first case study. (A) presents the linked knowledge graph in the context view. (B) displays the reasoning cards related to Tagliafico's passes. (C) illustrates the passing distributions of Tagliafico and Montiel.

6 System Evaluation

We evaluate InsightChaser through three approaches. First, we performed a model evaluation to quantitatively assess the performance of our visual-text linking method. Second, we conducted expert case studies to examine how the system supports visual reasoning and collected qualitative feedback. Finally, we conducted a user study with soccer fans to assess the usability and effectiveness of InsightChaser.

6.1 Model Evaluation

To assess the performance of visual-text linking, we conducted a model evaluation using GPT-4o as the LLM.

Data preparation. We randomly sampled several matches from the dataset and generated a total of 36 tactical insights covering three visualization types, following the data generation method described in Sec. 3.4. Each insight corresponds to a specific combination of analytical perspective (identification, comparison, or summarization) and content focus (player-level or team-level), ensuring a diverse range of tactical interpretations. Using our linking approach, we got 145 visual-text links from these insights. We then invited two experts with extensive experience in soccer tactical visualizations and expertise in visual-text linking to validate our generated links. The experts judged these pairs independently and considered two aspects: whether the generated links were accurate and whether all relevant concepts had been extracted completely. Any disagreements between experts were resolved through discussion to reach a consensus.

Result analysis. The detailed evaluation results are presented in Table 2. In our evaluation, accurate linking fully accounts for one-to-many relationships (e.g., a single entity such as "pass" may correspond to multiple lines in a passing network). Therefore, incomplete or partially correct links are considered incorrect. Expert validation identified 122 correct links, resulting in an overall precision of 84.14%. Additionally, experts found 14 missing links, leading to an overall recall of 89.71%. Consequently, the final F1-score for the linking evaluation is 86.83%.

We further analyzed model performance across the three visualization types to investigate their influence on linking accuracy, with detailed results shown in Table 2. The passing network achieved the highest precision of 84.44%. Given the structural complexity of the passing network, the high precision indicates effective knowledge graph construction and interaction between LLMs and the graph. Nonetheless, specific semantic ambiguities sometimes caused linking inaccuracies. For example, the phrase "Messi has top connections" is intended to reflect all passing interactions from Messi, but the entity "top connections" is incorrectly linked only to the player with whom Messi interacted most frequently. The shot map and carry heatmap exhibit relatively high recall values of 87.50% and 91.84%, respectively, suggesting effective decomposition of insights. However, their corresponding precision scores are lower: 82.35% for the shot map and 81.82% for the carry heatmap. Upon examining these failure cases, we observed that the LLM occasionally extracted abstract tactical concepts that lacked concrete visual counterparts. For example, the term "transition play" refers to a tactical notion rather than any entity type represented in our KG. This gap highlights the need for introducing visual reasoning.

6.2 Case Study

We conducted our case study and follow-up interviews with two domain experts, employing a think-aloud protocol throughout the process. The studies are based on real-world soccer data, as detailed in Sec. 3.4.

Table 2: Metrics (%) of Model Evaluation

Visualization	Precision	Recall	F1 Score
Passing Network	84.44	86.36	85.39
Shot Map	82.35	87.50	84.85
Carry Heatmap	81.82	91.84	86.54
Overall	**84.14**	**89.71**	**86.83**

6.2.1 Case 1: Player positions in passing network

We invited a soccer player (E1) with professional playing experience to the study. Given his enthusiasm for Argentina and their victory in the *2024 Copa América* final against Colombia, he selected this match for analysis. E1 specifically focused on the passing network (Fig. 5(A3)) to explore tactical insights related to Argentina's defenders (Fig. 5(A1)).

Upon interacting with the linking view, the highlighted textual entities immediately attracted his attention. While hovering over several automatically extracted keywords, he observed corresponding updates in the context and visualization views. One particular keyword, "attackers," attracted his curiosity. Exploring the context view further (Fig. 6(A)), he noted that "attackers" specifically referred to "Alvarez" and "Messi." After reviewing the contextual information, he recognized both players as forwards and remarked, "*The term "attackers" now explicitly refers to specific players, which helps clarify its tactical meaning.*" Nonetheless, he mentioned it was still challenging to directly infer these insights from the visualization alone.

Therefore, E1 turned to the enhancement view to examine the reasoning tree (Fig. 5(B)). Inspecting the reasoning structure, he quickly noticed differences between Montiel and Tagliafico, despite their shared defensive roles. Following Montiel's reasoning path, he inspected the highlighted passing network and found Montiel interacting frequently with Messi and Alvarez (the previously defined attackers). He commented, "*The labeling of attackers in the passing network is very clear and intuitive.*" Next, exploring Tagliafico's reasoning path and passing interactions (Fig. 6(B)), E1 observed that Tagliafico mostly initiated passes within midfield. When he hovered over "midfield," the visualization highlighted the tactical zone (middle third). Then, by querying "*What's the meaning of a right-back passing frequently in the middle-third?*" in the chat view, E1 remarked, "*The chat allowed me to effectively connect the midfield tactical area with Tagliafico's role.*" Reviewing the subsequent pass heatmap card, he identified Tagliafico's frequent passes initiated from deeper defensive positions and grasped the player's tactical style: "*Tagliafico passes mostly in deeper areas and excels at building up play from the back.*"

After applying visual effects, E1 noticed the passing network in the visualization view (Fig. 6(C)) clearly displayed pass distributions from Tagliafico and Montiel. Within Messi's circular node, the sector representing pass counts with Montiel was significantly larger than the one with Tagliafico, while no sector was present in Alvarez's circle for Tagliafico. E1 quickly realized that Montiel contributed more to supporting the attackers through his passes and considered the summarized reasoning particularly useful, "*This summary clearly highlights the differences between the two players, significantly improving my understanding.*" Ultimately, E1 agreed that the visual effects automatically suggested by the LLM effectively aligned with the identified tactical insights, and he chose to retain them.

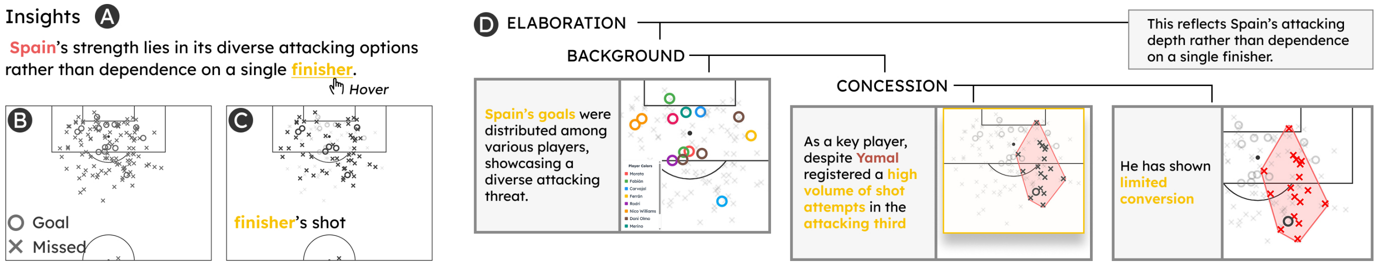

Fig. 7: The pipeline of our second case study. (A) presents the insight content. (B) displays Spain's shot map. (C) shows the linked visual elements with "finisher". (D) presents the visual reasoning path, including three rhetorical relationships: ELABORATION, BACKGROUND, and CONCESSION.

6.2.2 Case 2: Team attacks in shot map

We invited the second expert (E2), a die-hard soccer fan with many years of reading tactical analysis articles, to participate in this study. Motivated by Spain's victory in *EURO 2024*, he selected team Spain to investigate their shooting performance across their seven matches in the shot map (Fig. 7(B)) and chose an interesting insight (Fig. 7(A)).

E2 first explored the linked textual concepts. He first hovered over the highlighted entity "Spain" to view player information in the context view. Then he switched to the implicit entity "finisher", and explored the linked player positions displayed in the context view. He remarked, "*Finisher usually feels like an abstract concept, but here it's connected directly to player positions based on domain knowledge.*" Following the links from player positions, he identified specific players labeled as finishers. Simultaneously, the visualization highlighted shots by these players (Fig. 7(C)). Observing this closely, E2 noted, "*Shots by finishers indeed represent most of the team's attempts. The connection between the text and visualization is so concrete.*"

Next, E2 switched his attention to the enhancement view (Fig. 7(D)). Upon examining the tree structure, he first reviewed the card summarizing Spain's playing style and expressed agreement with its description. E2 further noticed that the reasoning tree primarily utilized "ELABORATION" relationships to convey insights. By clicking on individual cards, he accessed "BACKGROUND" information. Different players' goals were distinctly color-encoded, and he observed many players scoring goals but few achieving multiple goals. This strongly supported the insight regarding Spain's diverse attacking options. E2 commented, "*This visual effect clearly shows how goals are distributed among players. It's intuitive and convincing.*" Continuing his exploration, the user interacted with the subsequent card, which displayed a "CONCESSION" relationship. Through iterative interactions, he discovered that Yamal, identified as a finisher, accounted for a large proportion of Spain's shots yet scored only once. This visual observation matched the textual explanation about Yamal's low conversion rate. He commented, "*I didn't expect this detail when initially reading the insight. Yamal's example strongly supports this insight.*" In the edit view, the user modified the emphasis color of missed shots to red.

Finally, the user asked in the chat view about Yamal's performance in *EURO 2024*. The LLM replied that Yamal had provided four assists during the tournament, indirectly confirming Spain's collective attacking style and further enriching his knowledge.

6.2.3 Expert Interview

Following the case studies, we conducted an open discussion to gather user feedback, summarized as follows.

Usability. Participants provided positive feedback, confirming that our system effectively supported visual reasoning in tactical visualization. First, they appreciated the hover linking and low learning cost. E2 commented, "*The relationships among text, context, and visualization are really clear and easy to explore.*" Second, the visual reasoning process significantly improved their understanding. E1 said, "*The visual reasoning closely aligns with the insight. It effectively supports and expands the insight.*" E2 added, "*The visual effects are intuitive. Sometimes, just seeing the thumbnails helps me understand the text immediately.*" Finally, participants highlighted the additional flexibility provided by the edit view. Both E1 and E2 actively used this feature, with E1 commenting, "*I can easily adjust styles to my preference, making it convenient to share insights with others.*"

Suggestions. Experts also offered valuable suggestions for further improvements. E1 proposed an export feature, "*It would be great if the reasoning process could be exported as documents or videos.*" E2 expressed interest in analyzing multiple visualizations simultaneously, "*Combining multiple charts would help me better understand relationships between multiple objects and broaden the use cases.*"

6.3 User Study

We conducted a user study to further evaluate the system's usability and effectiveness from the perspective of soccer fans.

6.3.1 Experiment Setup

Participants and Data. We recruited 12 soccer fans (S1-S12, 9 males and 3 females) as participants. All participants regularly watched soccer matches but had limited familiarity with tactical analysis visualizations. None had participated in the development or previous interviews for InsightChaser. The user study was conducted under three scenarios with different matches, passing network and shot map from the earlier case studies, and a third scenario involving a heatmap visualization.

Procedure. Each study lasted about 50 minutes. We began with a 10-minute tutorial, introducing the study background, visual encodings used in tactical visualizations, and the interactions of our system. For each scenario (10 minutes per scenario), participants were asked to complete tasks following two stages. The linking stage focused on exploring linked insights and visualizations, including three tasks: capturing key concepts **(T1)**, finding contextual information **(T2)**, and locating corresponding visual elements **(T3)**. The reasoning stage emphasized comprehension of the reasoning process underlying the tactical insights, specifically involving understanding of the narrative structure **(T4)** and describing the enhanced visual effects **(T5)**. During the process, participants were allowed to freely explore the system and were encouraged to think aloud. To minimize potential bias caused by visualization types, we randomized the scenario order for each participant. After completing all tasks, participants rated their experiences for the five tasks **(T1-T5)** and additionally answered a comparative question **(T6)** regarding whether the reasoning stage enhanced their understanding compared to the simple linking stage. All responses used a 7-point Likert scale. Moreover, they filled out the System Usability Scale (SUS) questionnaire [5]. Finally, we conducted a 10-minute semi-structured interview to collect user feedback. All participants were compensated according to our school's guidelines.

6.3.2 Result Analysis

We present the results of the questionnaire in Fig. 8 and analyze them from two perspectives: usability and effectiveness.

Usablity. We calculated the average SUS score, which was 84.38, placing our system within the top 10% of ratings [5]. Detailed results are shown in Fig. 8(A). Additionally, we computed separate scores for usability (86.46) and learnability (76.04), further confirming the ability of our system [23]. Considering that most participants were soccer fans without extensive experience using interactive analysis systems, it is acceptable that some required assistance with system interactions. Nevertheless, all participants expressed willingness to use our system for understanding sports tactical visualization and highlighted the coherence and relevance of information presented in our system.

Effectiveness. We analyzed users' task ratings and their feedback, summarizing the following key findings on system effectiveness.

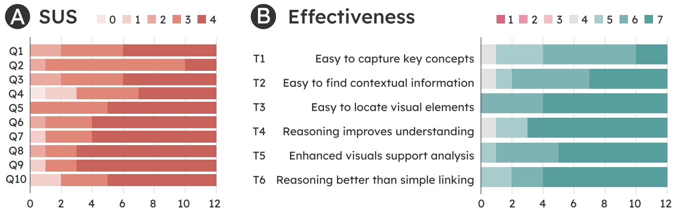

Fig. 8: Results of our user study. (A) shows scores of the System Usability Scale. (B) displays the results from our questionnaire.

Linking lays understanding foundation. Fig. 8(B) illustrates that participants generally gave positive ratings to the experience involving visual-text linking tasks (T1-T3). Participants S1, S3, and S7 particularly appreciated the highlighted concepts, noting that these allowed them to quickly grasp essential keywords and initiate analysis. However, S10 raised a concern, "*Highlighted concepts drew my attention, but sometimes distracted me from fully understanding the original meaning.*" Regarding contextual information, participants praised the hover interactions starting from key concepts, with S4 mentioning reduced cognitive load and S12 appreciating the smooth exploration across views. All participants positively agreed with the linked visualizations. Specifically, S7 emphasized, "*the linking is precise and intuitive, helping me quickly find relevant visual elements.*" Despite the generally positive feedback, S2 pointed out his confusion, "*Although these interactions help me identify keywords, they only clarify the sentences rather than truly helping me understand the deeper insights.*"

Reasoning leads to deeper comprehension. The experience when exploring reasoning-related tasks (T4-T6) was well-received overall (Fig. 8(B)). Participants indicated that the textual explanations effectively structured the reasoning narratives. S11 appreciated the structured tree representation and clearly labeled relationships, noting that "*this clearly helped me break down and understand the explanations*". However, S8 felt that when linking alone was already clear to interpret insights, the reasoning part mainly served as a confirmation rather than providing additional insights. S3 positively highlighted the thumbnail visualizations within reasoning cards, which provided concise overviews of enhanced visual effects. S6 appreciated how visual enhancements closely matched textual explanations, "*These visual effects made insights more concrete in the visualizations.*" Most participants agreed on the necessity and effectiveness of visual reasoning, with S5 and S2 specifically noting, "*The enhancement view provided deeper logical chains and substantially improved tactical understanding*."

7 Discussion

In this section, we discuss our work from two aspects as follows.

7.1 Visual Reasoning in Sports Tactical Visualization

Extensible pipeline. InsightChaser integrates structured knowledge graphs with LLM-based reasoning to build an extensible, modular framework. This extensibility is demonstrated in three main aspects: *1. Expanding visualization types.* Although tactical visualizations differ greatly in structure and visual encoding, our methodology readily adapts to various formats by adjusting the knowledge graph schema and LLM prompting strategies. *2. Incorporating richer data.* Contextual information such as team tactical patterns and player styles can be embedded to enable more tailored analytical approaches. *3. Enhancing visual effects.* By incorporating existing research (e.g., sports-related enhancements from [38]), the framework continues to broaden both analytical perspectives and presentation quality. Built on the synergy between knowledge graphs and LLMs, this framework design leaves ample room for integrating future composite tactical visualizations.

Generalizability. InsightChaser currently focuses on pitch-based soccer visualizations. Many other sports, such as badminton and basketball, also involve contextualized tactical visualizations. Constructing KGs using these sport-specific data can adapt to different domains. However, handling highly customized visualizations (e.g., complex glyphs) remains challenging. Developing a comprehensive representation of such visualization is a promising direction for improving cross-sport applicability. Beyond sports, our approach generalizes to domains such as finance and experimental sciences, where high-level insights are often inferred from visualizations without explicit references to chart elements in some scenarios. In addition, it may generalize to related tasks such as chart question answering [30], chart generation [45], which require aligning visual elements with textual insights.

Applicability. InsightChaser is designed for casual soccer fans and novice report readers who may struggle to interpret complex visualizations in tactical analysis articles. While the visualizations and insights in our study are automatically generated from data, this setup is not intended to simulate content creation. Instead, it allows us to control the availability of underlying data (e.g., detailed event records) and contextual information (e.g., player positions), which are essential for building knowledge graphs and performing visual-text linking. When such structured data is accessible, users can directly upload visualizations and textual insights to use InsightChaser for visual reasoning and exploration, without requiring synthetic generation.

7.2 Knowledge Graphs and LLMs Integration

Knowledge graph integration. InsightChaser treats the KG as a structured repository of visual and contextual information to support linking, rather than as interpretable knowledge. This design choice limits its ability to reason effectively in context-sparse scenarios [37]. Existing methods have explored combining search engines [40] and encoding additional textual data [36] to enrich external knowledge bases. However, due to the large scale of sports data and the complexity of tactical interpretation, integrating such knowledge requires rigorous validation, which remains an important direction for future work.

LLM performance. While our evaluations demonstrate generally reliable performance, LLM-based visual-text linking still faces several challenges. First, the model may extract irrelevant implicit entities or phrases that cannot be grounded in visual elements, leading to failure cases. To mitigate this, we provide detailed definitions and examples of implicit entity types in the prompt to guide extraction. Further improvements may be achieved by collecting more annotated tactical insights and fine-tuning a dedicated model. Second, the LLM is prone to hallucination when handling unfamiliar players or uncommon tactical terms, due to insufficient training coverage and limited contexts in the KG. Expanding the KG with richer data and context as an external knowledge source is a potential solution to improve grounding and reduce hallucinations. Third, the LLM lacks internal mechanisms to evaluate the plausibility of the results [47]. For example, the LLM fails to refine or reject the result when all generated queries fail. Introducing agent-based architectures with validation or re-querying strategies offers a promising direction for improving robustness.

In addition, InsightChaser leverages the advanced capabilities of GPT-4o throughout the pipeline. While effective, this reliance introduces instability due to model updates and limits the transparency of its behavior. Developing a unified framework or training an open-source LLM presents a potential solution to improve system stability and reduce reliance on black-box models. What's more, in our evaluation, we used the LLM-generated insights as input. While these insights were designed to ensure diversity across analytical perspectives and content levels, and were validated through expert feedback, it remains an open question how well they reflect the visualizations and how they differ from insights produced by real-world analysts. Addressing this requires collecting a large corpus of expert-authored insights and developing metrics to evaluate their quality across dimensions such as accuracy and specificity. This presents a promising direction for future research on the evaluation of LLM-generated data insights.

8 Conclusion

We introduced InsightChaser, a novel approach to bridge tactical insights and soccer visualizations through visual-text linking and reasoning enhancement. By integrating knowledge graphs and large language models, our method establishes explicit visual-text links and further enhances them with explanatory reasoning paths and corresponding visual effects to form interpretable visual reasoning. This enables fans to directly link abstract textual insights to contextualized sports visualizations, deepening their engagement. An interactive system further supports navigation and explanation. Case studies and a user study demonstrate the effectiveness of our approach.

ACKNOWLEDGMENTS

The work was supported by Zhejiang Provincial Natural Science Foundation of China under Grant No. LD25F020003 and Key "Pioneer" R&D Projects of Zhejiang Province (2023C01120).

REFERENCES

[1] Neo4j Graph Database. https://neo4j.com/. Accessed: 2025-03-31. 4

[2] G. L. Andrienko, N. V. Andrienko, G. Anzer, P. Bauer, G. Budziak, G. Fuchs, D. Hecker, H. Weber, and S. Wrobel. Constructing Spaces and Times for Tactical Analysis in Football. *IEEE Transactions on Visualization and Computer Graphics*, 27(4):2280–2297, 2019. doi: 10.1109/TVCG.2019.2952129 2

[3] S. K. Badam, Z. Liu, and N. Elmqvist. Elastic Documents: Coupling Text and Tables through Contextual Visualizations for Enhanced Document Reading. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):661–671, 2019. doi: 10.1109/TVCG.2018.2865119 2

[4] M. Brehmer and T. Munzner. A Multi-Level Typology of Abstract Visualization Tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2376–2385, 2013. doi: 10.1109/TVCG.2013.124 3

[5] J. Brooke et al. SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996. 8

[6] X. Cai, D. Weng, T. Fu, S. Fu, Y. Wang, and Y. Wu. Linking Text and Visualizations via Contextual Knowledge Graph. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–14, 2024. doi: 10.1109/TVCG.2024.3412241 2, 4

[7] A. Cao, X. Xie, R. Zhang, Y. Tian, M. Fan, H. Zhang, and Y. Wu. Team-Scouter: Simulative Visual Analytics of Soccer Player Scouting. *IEEE Transactions on Visualization and Computer Graphics*, 31(1):1–11, 2025. doi: 10.1109/TVCG.2024.3456216 2

[8] A. Cao, X. Xie, M. Zhou, H. Zhang, M. Xu, and Y. Wu. Action-Evaluator: A Visualization Approach for Player Action Evaluation in Soccer. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):880–890, 2024. doi: 10.1109/TVCG.2023.3326524 2

[9] L. Cheng, D. Deng, X. Xie, R. Qiu, M. Xu, and Y. Wu. SNIL: Generating Sports News From Insights With Large Language Models. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–14, 2024. doi: 10.1109/TVCG.2024.3392683 2, 3

[10] X. Chu, X. Xie, S. Ye, H. Lu, H. Xiao, Z. Yuan, C. Zhu-Tian, H. Zhang, and Y. Wu. TIVEE: Visual Exploration and Explanation of Badminton Tactics in Immersive Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):118–128, 2022. doi: 10.1109/TVCG.2021.3114861 2

[11] A. Cockburn, A. Karlson, and B. B. Bederson. A review of overview+detail, zooming, and focus+context interfaces. *ACM Computing Surveys*, 41(1):2:1–2:31, 2009. doi: 10.1145/1456650.1456652 6

[12] C. Dietrich, D. Koop, H. T. Vo, and C. T. Silva. Baseball4D: A Tool For Baseball Game Reconstruction & Visualization. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 23–32, 2014. doi: 10.1109/VAST.2014.7042478 2

[13] M. Du and X. Yuan. A Survey of Competitive Sports Data Visualization and Visual Analysis. *Journal of Visualization*, (1):47–67, 2021. doi: 10.1007/s12650-020-00687-2 2

[14] J. Fernandez-Navarro, L. Fradua, A. Zubillaga, P. R. Ford, and A. P. McRobert. Attacking and Defensive Styles of Play in Soccer: Analysis of Spanish and English Elite Teams. *Journal of Sports Sciences*, 34(24):2195–2204, 2016. doi: 10.1080/02640414.2016.1169309 2

[15] Y. Fu and J. Stasko. Supporting Data-Driven Basketball Journalism through Interactive visualization. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–17, 2022. doi: 10.1145/3491102.3502078 2

[16] Y. Fu and J. Stasko. HoopInSight: Analyzing and Comparing Basketball Shooting Performance Through Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):858–868, 2024. doi: 10.1109/TVCG.2023.3326910 2

[17] Y. Fu and J. Stasko. More Than Data Stories: Broadening the Role of Visualization in Contemporary Journalism. *IEEE Transactions on Visualization and Computer Graphics*, 30(8):5240–5259, 2024. doi: 10.1109/TVCG.2023.3287585 2

[18] J. Hao, M. Yang, Q. Shi, Y. Jiang, G. Zhang, and W. Zeng. FinFlier: Automating Graphical Overlays for Financial Visualizations with Knowledge-Grounding Large Language Model. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–17, 2024. doi: 10.1109/TVCG.2024.3514138 2

[19] Y. He, J. Lv, L. Cheng, L. Meng, D. Deng, and Y. Wu. ProTAL: A Drag-and-Link Video Programming Framework for Temporal Action Localization. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pp. 1–18, 2025. doi: 10.1145/3706598.3713741 2

[20] T. Kauer, D. Akbaba, M. Dörk, and B. Bach. Discursive Patinas: Anchoring Discussions in Data Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 31(1):1246–1256, 2025. doi: 10.1109/TVCG.2024.3456334 2

[21] D. H. Kim, V. Setlur, and M. Agrawala. Towards Understanding How Readers Integrate Charts and Captions: A Case Study with Line Charts. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, p. 1–11. ACM, 2021. doi: 10.1145/3411764.3445443 2

[22] S. Latif, Z. Zhou, Y. Kim, F. Beck, and N. W. Kim. Kori: Interactive Synthesis of Text and Charts in Data Documents. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):184–194, 2022. doi: 10.1109/TVCG.2021.3114802 1, 2

[23] J. R. Lewis and J. Sauro. The Factor Structure of the System Usability Scale. In *Human Centered Design: First International Conference, HCD 2009, Held as Part of HCI International 2009, San Diego, CA, USA, July 19-24, 2009 Proceedings 1*, pp. 94–103. Springer, 2009. doi: 10.1007/978-3-642-02806-9_12 8

[24] H. Li, Y. Wang, A. Wu, H. Wei, and H. Qu. Structure-aware Visualization Retrieval. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, article no. 409, 14 pages. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3491102.3502048 2

[25] T. Lin, A. Aouididi, Z. Chen, J. Beyer, H. Pfister, and J.-H. Wang. VIRD: Immersive Match Video Analysis for High-Performance Badminton Coaching. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):458–468, 2024. doi: 10.1109/TVCG.2023.3327161 2

[26] T. Lin, C. Zhu-Tian, Y. Yang, D. Chiappalupi, J. Beyer, and H. Pfister. The Quest for Omnioculars: Embedded Visualization for Augmenting Basketball Game Viewing Experiences. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):962–972, 2023. doi: 10.1109/TVCG.2022.3209353 3

[27] Y. Lin, H. Li, L. Yang, A. Wu, and H. Qu. InkSight: Leveraging Sketch Interaction for Documenting Chart Findings in Computational Notebooks. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):944–954, 2024. doi: 10.1109/TVCG.2023.3327170 2

[28] A. Maekawa, T. Hirao, H. Kamigaito, and M. Okumura. Can we obtain significant success in RST discourse parsing by using Large Language Models? In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 2803–2815. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.eacl-long.171 5

[29] W. C. Mann and S. A. Thompson. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text - Interdisciplinary Journal for the Study of Discourse*, 8(3):243–281, 1988. doi: 10.1515/text.1.1988.8.3.243 5

[30] A. Masry, D. X. Long, J. Q. Tan, S. Joty, and E. Hoque. Chartqa: A benchmark for question answering about charts with visual and logical reasoning. *arXiv preprint arXiv:2203.10244*, 2022. 9

[31] D. Masson, S. Malacria, G. Casiez, and D. Vogel. Charagraph: Interactive Generation of Charts for Realtime Annotation of Data-Rich Paragraphs. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, article no. 146, 18 pages. Association for Computing Machinery, 2023. doi: 10.1145/3544548.3581091 2

[32] R. Metoyer, Q. Zhi, B. Janczuk, and W. Scheirer. Coupling Story to Visualization: Using Textual Analysis as a Bridge Between Data and Interpretation. In *Proceedings of the International Conference on Intelligent User Interfaces*, pp. 503–507, 2018. doi: 10.1145/3172944.3173007 2

[33] P. Muller, C. Braud, and M. Morey. ToNy: Contextual embeddings for accurate multilingual discourse segmentation of full documents. In *Proceedings of the Workshop on Discourse Relation Parsing and Treebanking 2019*, pp. 115–124. Association for Computational Linguistics, 2019. doi: 10.18653/v1/W19-2715 5

[34] J. P. Ono, C. Dietrich, and C. T. Silva. Baseball Timeline: Summarizing Baseball Plays Into a Static Visualization. In *Proceedings of the Computer Graphics Forum*, vol. 37, pp. 491–501, 2018. doi: 10.1111/cgf.13436 2

[35] OpenAI. GPT-4o. https://openai.com/index/hello-gpt-4o/,

2024. Accessed: 2025-03-28. 3

[36] S. Pan, L. Luo, Y. Wang, C. Chen, J. Wang, and X. Wu. Unifying Large Language Models and Knowledge Graphs: A Roadmap. *IEEE Transactions on Knowledge and Data Engineering*, 36(7):3580–3599, 2024. doi: 10.1109/TKDE.2024.3352100 9

[37] L. Peng, Z. Lin, N. Andrienko, G. Andrienko, and S. Chen. Contextualized Visual Analytics for Multivariate Events. *Visual Informatics*, 9(2):100234, 2025. doi: 10.1016/j.visinf.2025.100234 9

[38] C. Perin, R. Vuillemot, and J.-D. Fekete. SoccerStories: A Kick-off for Visual Soccer Analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2506–2515, 2013. doi: 10.1109/TVCG.2013.192 2, 9

[39] C. Perin, R. Vuillemot, C. D. Stolper, J. T. Stasko, J. Wood, and S. Carpendale. State of the Art of Sports Data Visualization. In *Proceedings of the Computer Graphics Forum*, vol. 37, pp. 663–686, 2018. doi: 10.1111/cgf.13447 1, 2

[40] Y. Qin, S. Hu, Y. Lin, W. Chen, N. Ding, G. Cui, Z. Zeng, X. Zhou, Y. Huang, C. Xiao, C. Han, Y. R. Fung, Y. Su, H. Wang, C. Qian, R. Tian, K. Zhu, S. Liang, X. Shen, B. Xu, Z. Zhang, Y. Ye, B. Li, Z. Tang, J. Yi, Y. Zhu, Z. Dai, L. Yan, X. Cong, Y. Lu, W. Zhao, Y. Huang, J. Yan, X. Han, X. Sun, D. Li, J. Phang, C. Yang, T. Wu, H. Ji, G. Li, Z. Liu, and M. Sun. Tool learning with foundation models. *ACM Computing Surveys*, 57(4):1–40, 2024. doi: 10.1145/3704435 9

[41] D. Seebacher, T. Polk, H. Janetzko, D. A. Keim, T. Schreck, and M. Stein. Investigating the Sketchplan: A Novel Way of Identifying Tactical Behavior in Massive Soccer Datasets. *IEEE Transactions on Visualization and Computer Graphics*, 29(4):1920–1936, 2023. doi: 10.1109/TVCG.2021.3134814 2

[42] Z. Shao, L. Shen, H. Li, Y. Shan, H. Qu, Y. Wang, and S. Chen. Narrative Player: Reviving Data Narratives with Visuals. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–15, 2025. doi: 10.1109/TVCG.2025.3530512 1, 2

[43] L. Shen, Y. Zhang, H. Zhang, and Y. Wang. Data Player: Automatic Generation of Data Videos with Narration-Animation Interplay. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):109–119, 2024. doi: 10.1109/TVCG.2023.3327197 2

[44] StatsBomb. StatsBomb Open Data. `https://github.com/statsbomb/open-data`, 2024. 3

[45] Y. Tian, W. Cui, D. Deng, X. Yi, Y. Yang, H. Zhang, and Y. Wu. ChartGPT: Leveraging LLMs to Generate Charts from Abstract Natural Language. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–15, 2024. doi: 10.1109/TVCG.2024.3368621 2, 9

[46] J. Wang, X. Li, C. Li, D. Peng, A. Z. Wang, Y. Gu, X. Lai, H. Zhang, X. Xu, X. Dong, Z. Lin, J. Zhou, X. Liu, and W. Chen. AVA: An Automated and AI-Driven Intelligent Visual Analytics Framework. *Visual Informatics*, 8(2):106–114, 2024. doi: 10.1016/j.visinf.2024.06.002 2

[47] S. Wu, Y. Yang, W. Wu, R. Li, Y. Zhang, G. Wang, H. Tan, Z. Liu, and L. Shi. Visual Analysis of LLM-Based Entity Resolution from Scientific Papers. *Visual Informatics*, p. 100236, 2025. doi: 10.1016/j.visinf.2025.100236 9

[48] Y. Wu, D. Deng, X. Xie, M. He, J. Xu, H. Zhang, H. Zhang, and Y. Wu. OBTracker: Visual Analytics of Off-ball Movements in Basketball. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):929–939, 2023. doi: 10.1109/TVCG.2022.3209373 2

[49] Y. Wu, J. Lan, X. Shu, C. Ji, K. Zhao, J. Wang, and H. Zhang. iTTVis: Interactive Visualization of Table Tennis Data. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):709–718, 2018. doi: 10.1109/TVCG.2017.2744218 2

[50] Y. Wu, X. Xie, J. Wang, D. Deng, H. Liang, H. Zhang, S. Cheng, and W. Chen. ForVizor: Visualizing Spatio-Temporal Team Formations in Soccer. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):65–75, 2019. doi: 10.1109/TVCG.2018.2865041 2

[51] X. Xie, J. Wang, H. Liang, D. Deng, S. Cheng, H. Zhang, W. Chen, and Y. Wu. PassVizor: Toward Better Understanding of the Dynamics of Soccer Passes. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1322–1331, 2021. doi: 10.1109/TVCG.2020.3030359 2

[52] L. Yao, R. Vuillemot, A. Bezerianos, and P. Isenberg. Designing for Visualization in Motion: Embedding Visualizations in Swimming Videos. *IEEE Transactions on Visualization and Computer Graphics*, 30(3):1821–1836, 2024. doi: 10.1109/TVCG.2023.3341990 2

[53] Y. Ye, J. Hao, Y. Hou, Z. Wang, S. Xiao, Y. Luo, and W. Zeng. Generative AI for Visualization: State of the Art and Future Directions. *Visual Informatics*, 8(2):43–66, 2024. doi: 10.1016/j.visinf.2024.04.003 2

[54] L. Ying, Y. Wang, H. Li, S. Dou, H. Zhang, X. Jiang, H. Qu, and Y. Wu. Reviving static charts into live charts. *IEEE Transactions on Visualization and Computer Graphics*, 31(8):4314–4328, 2025. doi: 10.1109/TVCG.2024.3397004 2

[55] M. Zheng, D. Lillis, and A. G. Campbell. Current State of the Art and Future Directions: Augmented Reality Data Visualization to Support Decision-Making. *Visual Informatics*, 8(2):80–105, 2024. doi: 10.1016/j.visinf.2024.05.001 2

[56] Z. Zhou, H. Wang, Z. Zhao, F. Zheng, Y. Wang, W. Chen, and Y. Wang. ChartKG: A Knowledge-Graph-Based Representation for Chart Images. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–15, 2024. doi: 10.1109/TVCG.2024.3476508 2

[57] C. Zhu-Tian and H. Xia. CrossData: Leveraging Text-Data Connections for Authoring Data Documents. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–15, 2022. doi: 10.1145/3491102.3517485 2

[58] C. Zhu-Tian, Q. Yang, X. Xie, J. Beyer, H. Xia, Y. Wu, and H. Pfister. Sporthesia: Augmenting Sports Videos Using Natural Language. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):918–928, 2023. doi: 10.1109/TVCG.2022.3209497 2