RichCast - A Voice-Driven Interactive Digital Narrative Authoring System

Christopher Ferraris[1][0000−0002−8870−9908]
and Charlie Hargood[2][0000−0002−5067−5941]

[1] Panivox, UK `chris.ferraris@panivox.com`
[2] Bournemouth University, UK `chargood@bournemouth.ac.uk`

Abstract. We present RichCast a platform for conversational audio Interactive Digital Narrative (IDN). RichCast includes an accessible 'No-code' authoring tool, a community driven library of works, and voice interactive medium for interactive storytelling.

Keywords: Interactive Digital Narrative · Authoring Tool · Voice control · Mobile gaming

1 Introduction and Background

Voice recognition technology grows increasingly pervasive. With the launch of Apple Siri in 2010, Microsoft Cortana in 2013, Amazon Echo in 2014 and Google Assistant in 2016, voice as a user interface is becoming increasingly normalised in our day-to-day lives. Where once specific voice commands were required, today's voice recognition systems consist of, 'conversational AI' which can include Automatic Speech Recognition (ASR), Natural Language Processing (NLP) and Text-to-Speech (TTS) and have been used to create, 'emotive speech' systems, where users can interact with AI agents using natural language [2].

Voice and audio as an interactive medium hold great potential for Interactive Digital Narrative (IDN), providing potentially screen, and hands, free interactive story experiences. Some forms of IDN which lend themselves to a reduced reliance on screens, such as locative narrative where the users physical surroundings are critical, stand to benefit from further research in screenless forms of interaction. Beyond this other forms of IDN may benefit from a conversational approach to storytelling where using voice and audio the player/reader can converse with characters and ask questions in natural language as a voice extension to the NLP powered free agency first explored by Façade [9]. There are some examples of emergent technologies to support this, including the bespoke actors of Charisma [12], or the voice controlled experiences of Fabella [16]. In this work we present one of these emerging technologies in the form of a community centric conversational narrative framework RichCast[1].

RichCast is more than just a means of voice interaction, its community centric approach means authors and authoring tools are a priority. Authorship and

---

[1] RichCast by Panivox as available here https://www.richcast.com/ as of 20/7/22

authoring technologies have long been both a critical area for IDN research and one that presents many challenges [15, 4]. How 'authoring tool' is defined is a matter for debate [14, 5] - but broadly we can refer to applications that are used by authors principally for the creation of IDNs. A range of interface design paradigms can be seen from the classic Hypertext graphs of StorySpace [1], to domain specific languages such as in Inform 7 [11], to faceted approaches as seen in StoryPlaces [10, 7]. However, a principle challenge for all authoring tools remains accessibility [15, 4, 10] and enabling non technical authors to create IDN and work within the medium without the need to hire programmmers. Indeed, much of the success of Twine [8] (arguably one of the more widely used IDN authoring tools) has been attributed to the tools accessibility for non technical creatives [3, 13]. The accessibility of our authoring tools not only impacts usability but can influence the resulting works themselves [6] and potentially the entrance of new creative voices in our community [3]. Consequently ways to bring IDN authorship to the author and improving accessibility of our tools can be considered a long standing priority for IDN authoring tool research. It is also to be noted that IDN is not a single form, with games, locative narratives, Hypertext, parser fiction and more all existing under the IDN umbrella - and these different forms demand bespoke authoring tools [10].

In this demo submission we present both RichCast and it's authoring tool.
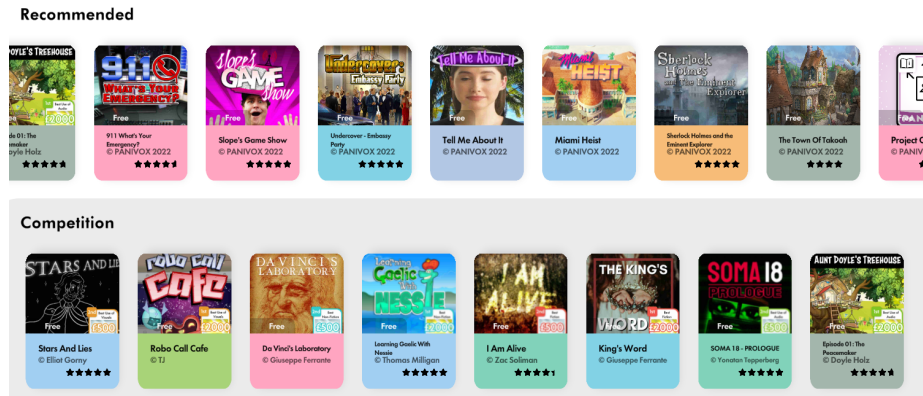

Fig. 1. RichCast supports voice-driven IDN authoring by it's community

2   RichCast

RichCast's key design features are:

- **Conversational Storytelling** Using conversational AI', the Keywords and Fallback systems and customizable AI Voice Actors

– **'Effortless' IDN authoring** With a design philosophy that aims to make IDN creation 'effortless' for new and advanced creators, through design techniques such a 'No code' audio & visuals and supporting visual clarity via the Junction Tile

2.1 Conversational Storytelling

RichCast brings conversational voice interactive AI with accessible authorship to IDN. Utilising the accessibility of voice-driven technology as a user interface, RichCast aims to support community-created content and encourage anyone to create their own voice-driven interactive experiences, as seen in figure 1.

Creating a tool that supports conversational storytelling by anyone, regardless of their IDN or technical experience, is the foundation of RichCast and presents some interesting challenges. With RichCasts's conversational storytelling, players can interact with AI agents using natural language, this can be to ask questions, roleplay conversation, or take actions as seen in an example Sherlock Holmes story in figure 3 on the next page. How can we present a relatively complex system in a format that anyone can use to create voice-driven IDNs? To this end, RichCast's design has focused on reducing barriers traditionally associated with using voice in IDN projects, from coding knowledge to the user interface to the production costs associated with voice input and voice acting, so that we can support a wide variety of creators.
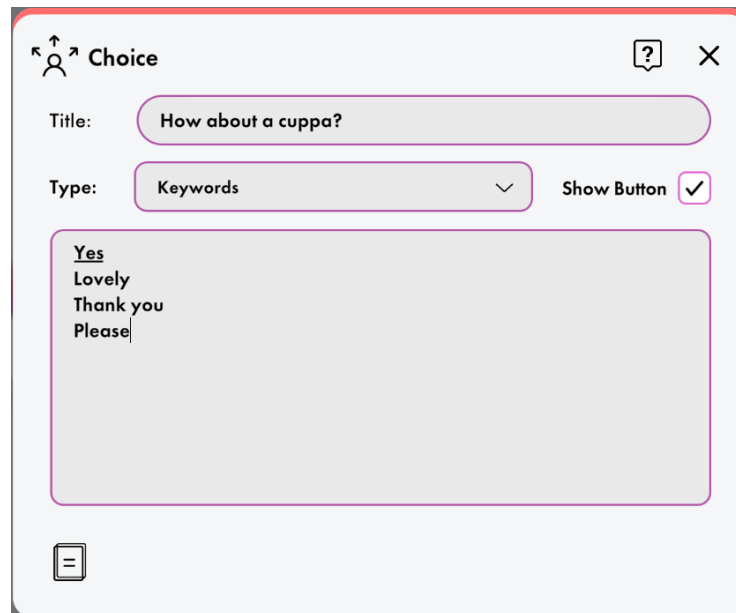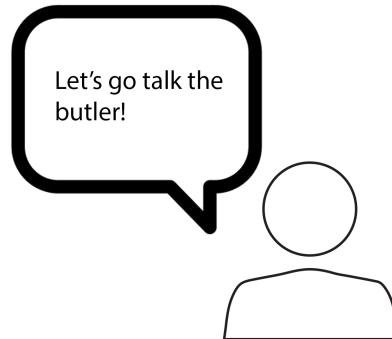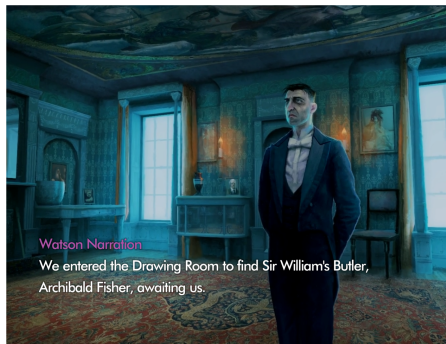

Fig. 2. RichCast's choice tile using keywords to take voice inputs without code.
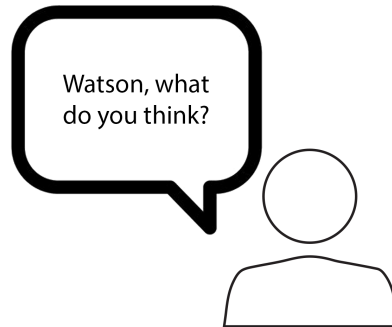
Watson presents the challenge

The player chooses what to do

Watson introduces a character

The players asks Watson for help

Watson suggests an action

The player agrees and continues

Fig. 3. Example of conversational storytelling, with RichCasts's "Sherlock and the Eminent Explorer". The IDN delivers the scene and setting to the player audibly and visually they can issue instructions, ask questions, and converse with the characters.

Keywords & Fallback System When a creator wishes to create a response to specific voice input by a user, they use Keywords. Keywords are a re-imagining of 'intent' words used by systems such as Amazon's Alexa system. RichCast developers have not created the speech recognition libraries themselves. Indeed there are many freely accessible APIs such as Google Speech-to-Text, Microsoft Azure and Amazon Transcribe. Panivox have created a best-in-class system to embrace these technologies by enabling a system of accessibility and ease of use.

In order to build a voice-driven interactive experience, IDN creators need to instruct RichCast to respond to a user's words. The Keyword system is part of the 'Choice' tile. The Choice tile, like all of RichCast tiles, follows the 'no code' design philosophy to support creators with no coding knowledge and presents the creator with an easy to navigate window where they can customise how a user's voice input is managed as seen in figure 2 at the start of this section.

Within the Choice window, Keywords such as 'Yes' are underlined, indicating they are part of a Keyword group. Keyword groups allow creators to easily see all the associated voice responses associated with a given voice input and allow users to provide a variety of responses. For example, the Keyword, 'Yes' also supports the user voice input of, 'ok', 'confirm', 'sure', 'yup', and 'of course' as seen in figure 4.
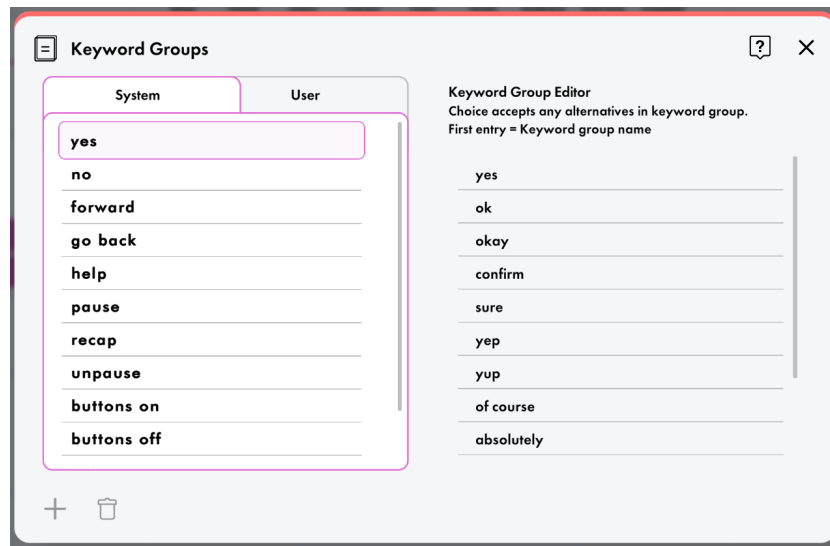

Fig. 4. RichCast's keyword group system grouping varied responses under key actions to aid voice interaction design.

Using voice to interact with a game's cast of characters can create memorable experiences. However, one challenge for IDN creators is to present non-player characters as believable 'living' agents, rather than computer simulations. The

Fallback system supports this goal of presenting characters as living agents by providing creators with the option to customize the automated responses provided by a character in response to unrecognized voice commands as seen in figure 5. This form of elegant failure allows the author to creative immersive experiences that feel conversational rather than filled with dead ends.
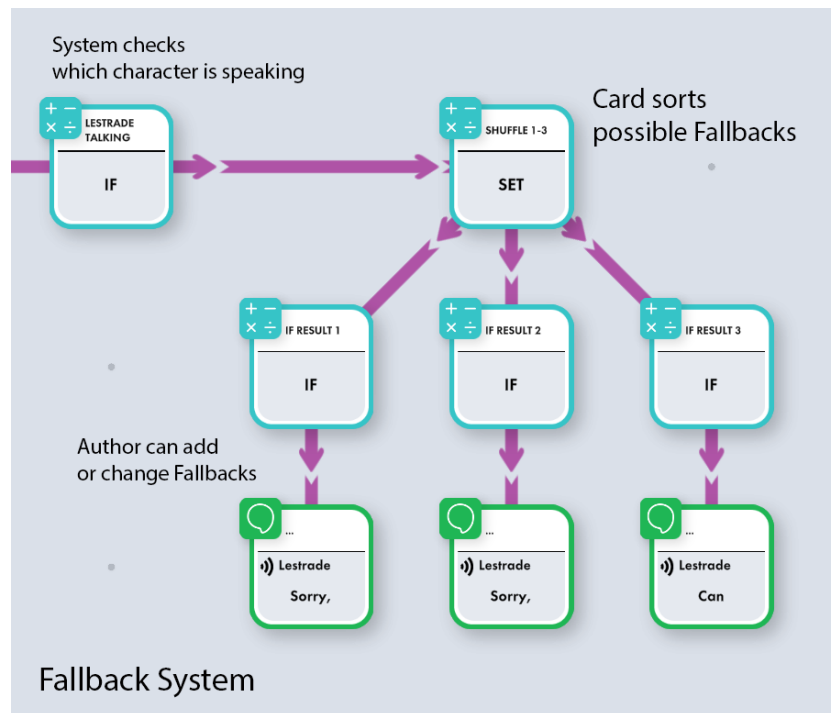

Fig. 5. RichCast's fallback system allowing authors to specify elegant failure.

The Fallback system takes a user's voice input, determines which NPC is being interacted with and then directs the RichCast system to one of a number of responses. The system then uses a 'card shuffle' algorithm to randomize the responses given to the user and minimize the chances of the same response being repeated consecutively. creators can customize their own Fallback responses, either changing the existing texts provided by RichCast or adding their own, as suits the style of experience and character they wish to create.

AI Voice Actors Learning from other AI voice developers, such as IBM, Apple, Amazon and & Google, has allowed RichCast to deliver over 150 AI voice actors, with a wide variety of backgrounds and personalities, for Creators to use in their IDN Projects. Each AI voice actor's voice can be adjusted for pitch and speed by creators to support there IDN projects as depicted in figure 6.
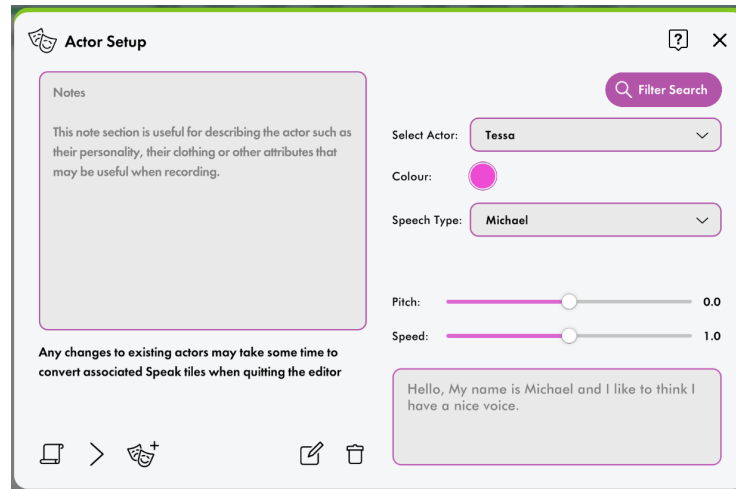
Fig. 6. RichCast's custom actor set up for TTS

Creators can also use RichCast's AI voice system to record their own speech, converting their spoken words to text that can be displayed on screen or voiced by a selected AI voice actor.

To further support creators wishing to create fully voiced projects, RichCast provides the functionality to import recorded voice files by professional voice actors. By providing creators with the tools to populate their IDN projects with fully voiced characters, RichCast moves closer to providing a tool where anyone can create and enjoy fully-voiced IDN projects.

2.2   'Effortless' IDN authoring

RichCast aims to support IDN authors from all backgrounds. To this end, the system incorporates a number of design techniques in an effort to improve the IDN authoring experience.

Project structuring IDN tools make use of Hypertextual graphs in the form of flowcharts to support creators as they structure and create their projects. As a IDN project becomes more complex, its flowcharts can become a complex, intricate network of tiles that can be difficult to understand, even for professional creators.

Flowchart systems permit a value to exit a tile along multiple paths, depending on the condition within the tile. RichCast utilises a binary node system for each tile with, 'one input, one output', where a value passes through a tile if the condition in the tile is true. Visually, logic conditionals are represented as their own discrete tiles, rather than being displayed on connecting path, so that a user only needs to follow and check each title. While this does mean an additional

element is introduced into the flowchart system, it does aim to present a clearer representation of what it happening in a project, as shown in figure 7.
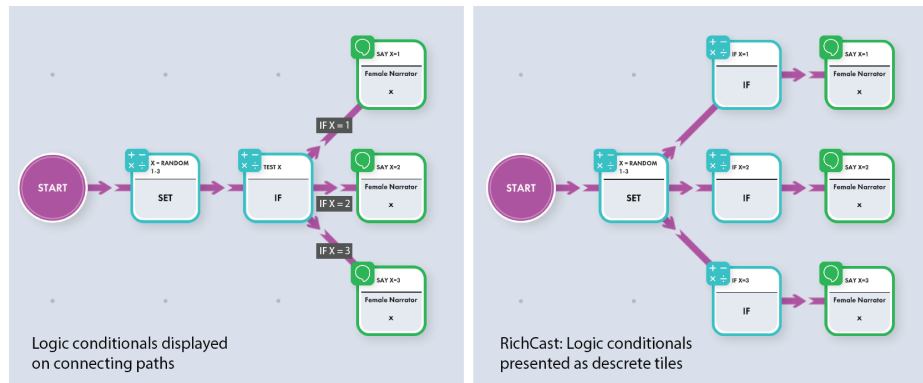

Fig. 7. Improving Flowchart Readability using Binary Nodes

Junction Tile The 'Junction' tile provides users with an 'empty' tile that can be implemented in IDN projects to help them manage the visual representation of their experience. The Junction tile takes any input and connects it to other tiles within the flowchart system as shown in figure 8.

When a flowchart diagram represents a complex system, overlapping connections between nodes can create a representation that is 'messy' and hard to read. Using the Junction tile, creators can better manage their flowchart systems, 'untangling' connections between nodes. This cleaner representation of the flowchart supports the creators and their collaborators directly, by improving readability.
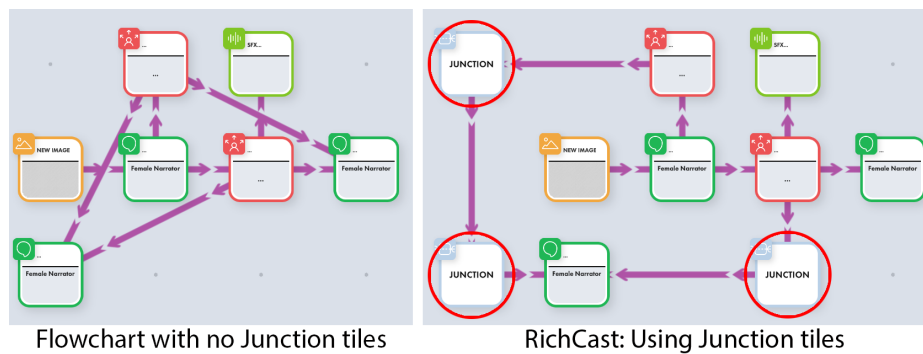

Fig. 8. RichCast's Juction tiles, circled in red, allow the user to organise the layout of their flowchart

3   Conclusion

RichCast presents conversational AI technology for IDN creators, to support a wide range of voice-driven IDN projects, from interactive fiction, to quizes, to cultural heritage projects. With our community-content driven approach to design, barriers that are traditionally associated with IDN projects, such as the need to understand programming, the cost of hiring voice actors and complexity of IDN creatoring tools, are being removed.

With our 'Creators Fund' we encourage anyone with an interest in IDN to explore how RichCast can support their IDN projects. As we continue to engage with our online community, we strive to support would-be creators and professionals alike from using our authoring tool to create engaging, voice-driven interactive experiences, with full 'no code' audio and visuals support.

References

1. Bernstein, M.: Storyspace 3. In: Proceedings of the 27th ACM Conference on Hypertext and Social Media. p. 201–206. HT '16, Association for Computing Machinery, New York, NY, USA (2016)
2. Charles, F., Pizzi, D., Cavazza, M., Vogt, T., André, E.: Emoemma: Emotional speech input for interactive storytelling. In: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2. pp. 1381–1382 (2009)
3. Ellison, C.: Anna anthropy and the twine revolution. The Guardian **10** (2013)
4. Green, D.: Don't Forget to Save! The Impact of User Experience Design on Effectiveness of Authoring Video Game Narratives. Ph.D. thesis (2021)
5. Green, D., Hargood, C., Charles, F.: Define "authoring tool": a survey of interactive narrative authoring tools. In: Authoring for Interactive Storytelling Workshop 2018 (2018)
6. Green, D., Hargood, C., Charles, F.: Use of tools: Ux principles for interactive narrative authoring tools. Journal on Computing and Cultural Heritage (JOCCH) **14**(3), 1–25 (2021)
7. Hargood, C., Weal, M.J., Millard, D.E.: The storyplaces platform: Building a web-based locative hypertext system. In: Proceedings of the 29th on Hypertext and Social Media, pp. 128–135 (2018)
8. Klimas, C.: Twine, https://twinery.org/, accessed: 26/07/2022
9. Mateas, M., Stern, A.: Façade: An experiment in building a fully-realized interactive drama. In: Game developers conference. vol. 2, pp. 4–8 (2003)
10. Millard, D., Hargood, C., Howard, Y., Packer, H.: The storyplaces authoring tool: pattern centric authoring. In: Authoring for Interactive Storytelling Workshop 2017 (2017)
11. Nelson, G.: Natural language, semantic analysis and interactive fiction. IF Theory Reader **141**, 99–104 (2006)
12. Parry, A., Salili-James, B., Gadney, G.: Charisma, https://charisma.ai/, accessed: 26/07/2022
13. Salter, A.: Interactive narrative. Humanities Commons (2020)
14. Shibolet, Y., Knoller, N., Koenitz, H.: A framework for classifying and describing authoring tools for interactive digital narrative. In: International Conference on Interactive Digital Storytelling. pp. 523–533. Springer (2018)

15. Spierling, U., Szilas, N.: Authoring issues beyond tools. In: Joint International Conference on Interactive Digital Storytelling. pp. 50–61. Springer (2009)
16. Wanderword: Fabella, https://fabellacreator.com/, accessed: 26/07/2022