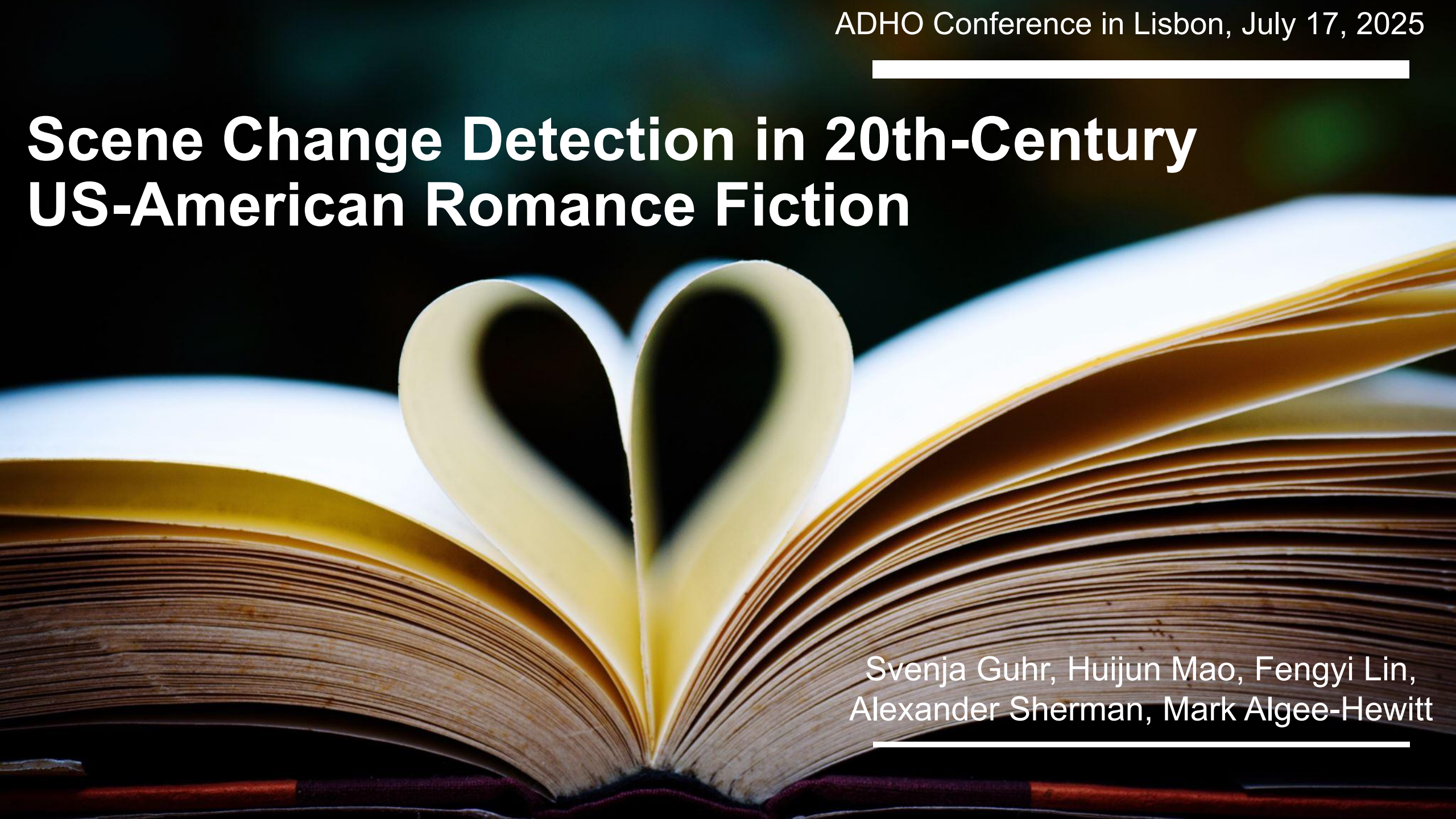
ADHO Conference in Lisbon, July 17, 2025

Scene Change Detection in 20th-Century US-American Romance Fiction



Svenja Guhr, Huijun Mao, Fengyi Lin,
Alexander Sherman, Mark Algee-Hewitt

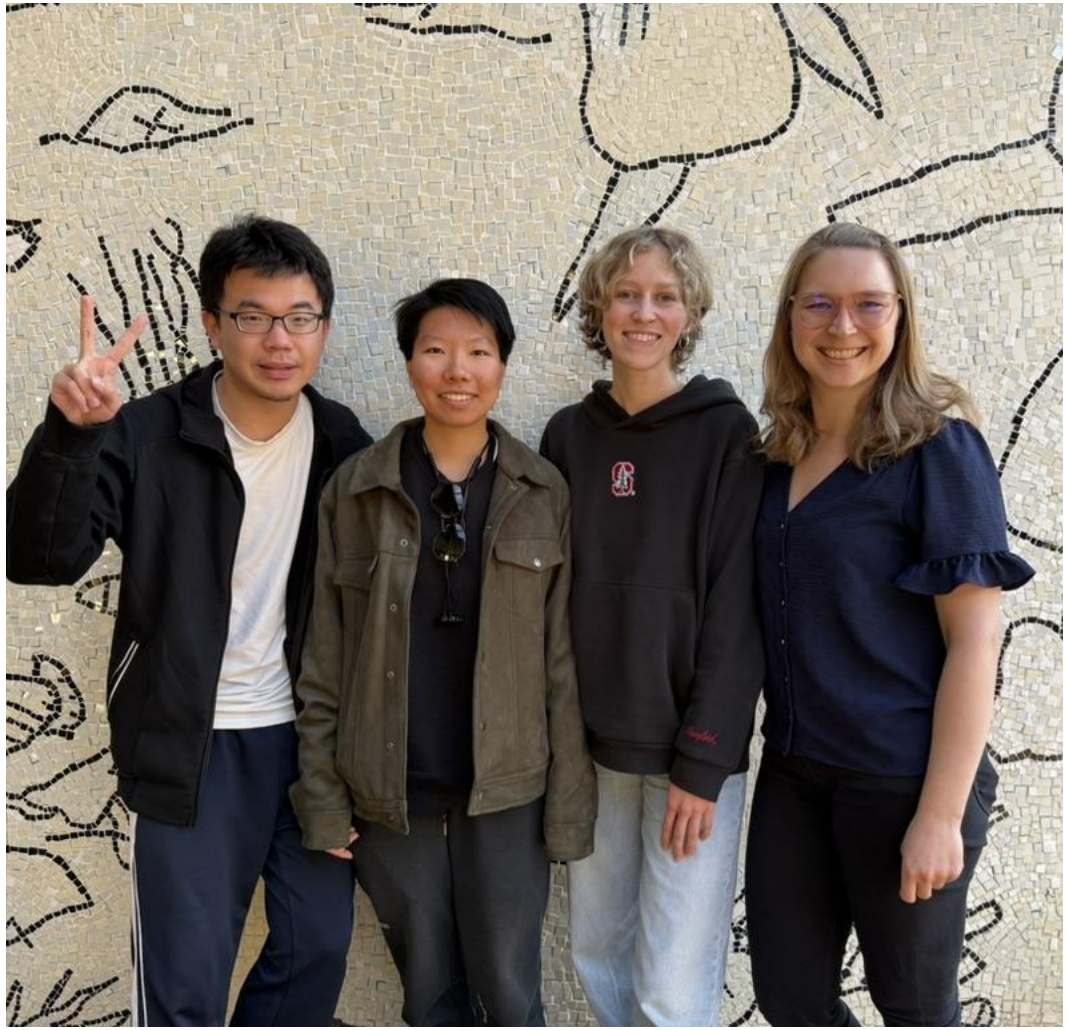
-Project:

Romance Novels.

Segmentation, Domestic Space, Gendered Stereotypes

+ student annotators:
Jessica Monaco, Alexander Sherman, Mallen Clifton, Ellen Yang

GitHub Repo with Code & Slides →
github.com/literarylab/scene_segmentation



Core research team:
Fengyi Lin, Huijun Mao, Agnes Hilger, Svenja Guhr

Research Questions

- How can scene segmentation enhance scalable reading of Romance novels, particularly in relation to narrative structure, scene length, character dynamics, and spatial settings?
- To what extent can scene segmentation be automated using manually annotated data and transformer-based language models or generative models (e.g., BERT, LLMs)?
- How do Romance novels engage with gender(ed) stereotypes, and how does this representation evolve in the 20th century?
- What is the relationship between gender(ed) stereotypes and domestic space?

Project Milestones

Scene Segmentation

- Manual annotation
- Automated detection of scene change positions

Domestic Space

- Automated classification of scene segments
- Comparative analysis of domestic scenes

Gendered Stereotypes

- Automated character recognition
- Analysis of gendered stereotypes

Project Milestones

Scene Segmentation

- Manual annotation
- Automated detection of scene change positions

Domestic Space

- Automated classification of scene segments
- Comparative analysis of domestic scenes

Gendered Stereotypes

- Automated character recognition
- Analysis of gendered stereotypes

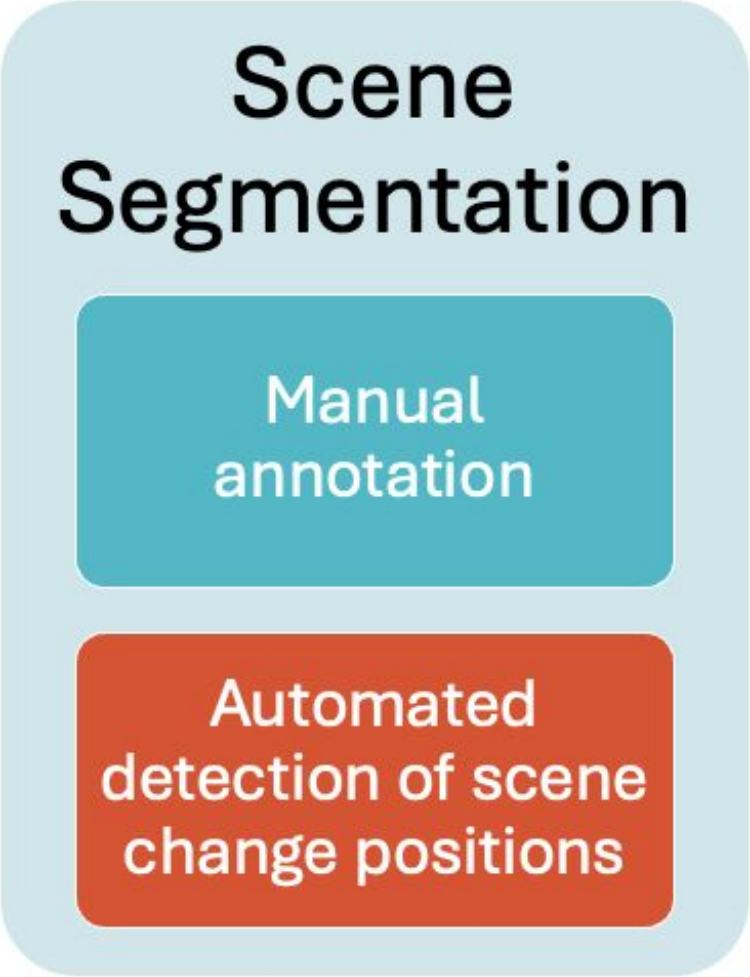
Work in Progress
“Workshop report”

Segmenting Novels into Scenes

Why do scenes matter?

- Literary critical importance
  - Formal
  - Historical
  - Generic
- Computational importance
  - Meaningful semantic units
  - Especially for space, time, and character

First Milestone: Scene Segmentation



1. Manual annotation of scene segments (CATMA)
2. Discussion and measurement of IAA (GitMA)
3. Preparation of gold annotations
4. Data preprocessing and upsampling for automation
5. Automation of scene segmentation through scene boundary detection with LM and LLMs

What is a scene segment in a novel?

1.1 Definition "Scene"

A scene is defined as *a segment of the* ***discourse*** *(presentation) of a narrative which presents a part of the* ***histoire*** *(chronologically ordered, causally connected events in the narrated world) in such a way that*

a) *time is more or less similar in discourse and histoire,*

b) *place stays - more or less - the same,*

c) *it centers around a particular action, and*

d) *the character configuration is - again: more or less - equal*

based on [Gius et al. 2019].

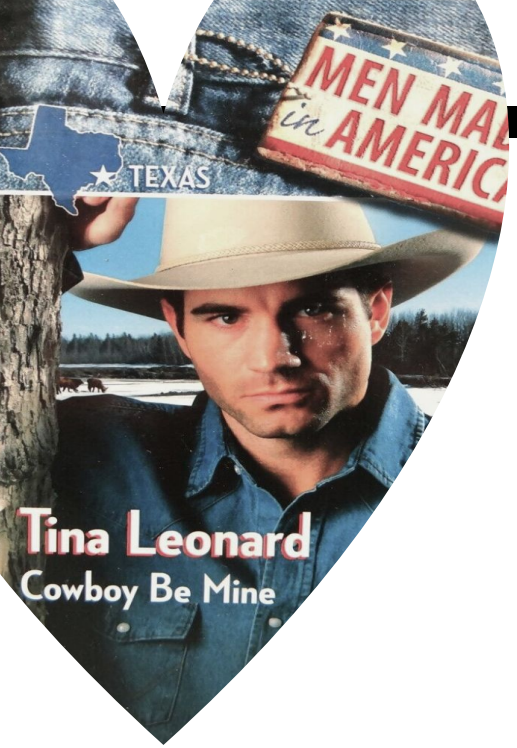
Guidelines for Scene Annotation by Gius et al. (2021, 3)

Concentration on Scene Change Positions

He tipped his hat to her. "Be seeing you." "Goodbye," she murmured through stiff lips. She saw him to the door, managing a frozen smile as she closed the door behind him. Then she put her head in her hands and told herself she wouldn't cry.
She wouldn't. Not over Michael Wade.

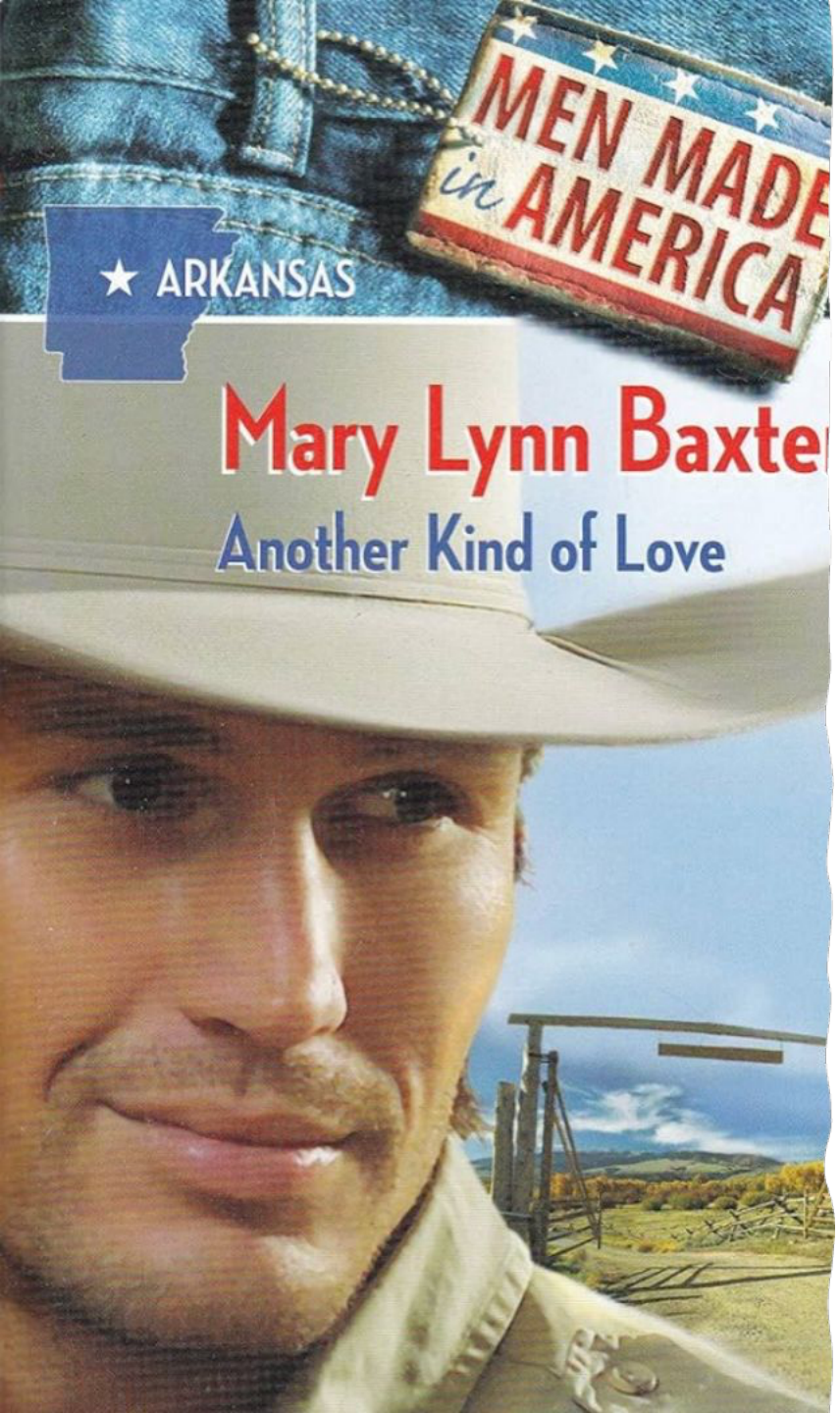
<scene-change/>

The doorbell rang. Bailey stiffened, wondering if Gunner might have quickly decided to snatch back one or both of his offers. She pulled the door open again, looking out cautiously. Chili Haskins stood on her porch, his white, bushy mustache like icicles above his lips. "Howdy, Bailey."



Corpus

Harlequin Series
Men Made in America



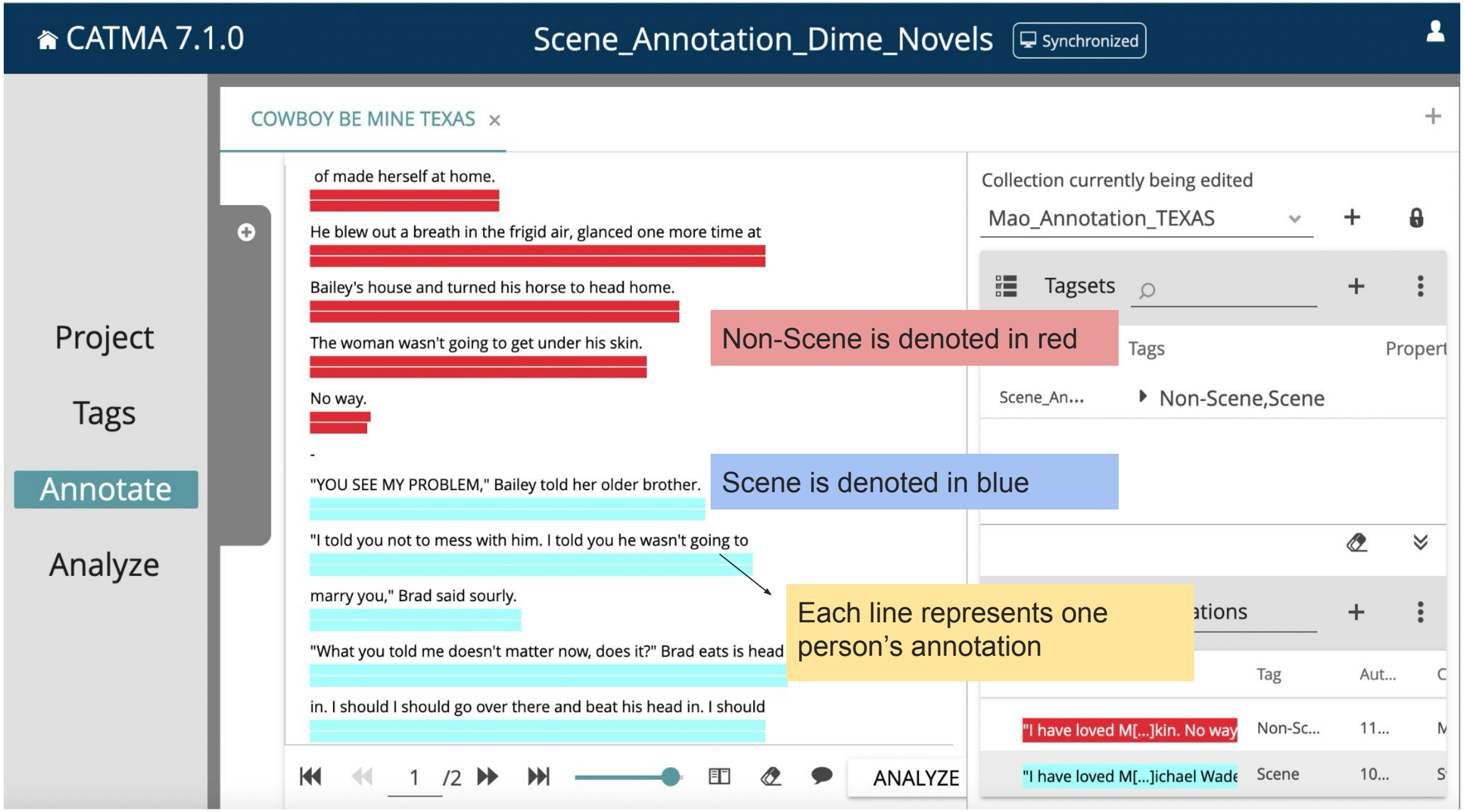
Corpus

Harlequin Series *Men Made in America*

- US-American Romance novels
- written by 40 female authors
- 50 States – 50 novels (local settings)

Title	Author	Date	Series
Cowboy Be Mine	Tina Leonard	2000	MEN MADE in AMERICA
To Love a Dreamer	Ruth Langan	1985	MEN MADE in AMERICA
Send Me a Hero	Rita Herron	1998	MEN MADE in AMERICA
Warrior's Embrace	Peggy Webb	2000	MEN MADE in AMERICA
Murdock's Family	Paula Detmer Riggs	1994	MEN MADE in AMERICA
Mysterious Stranger	Patricia Rosemoor	2002	MEN MADE in AMERICA
The Bargain	Patricia Coughlin	1988	MEN MADE in AMERICA
Another Kind of Love	Mary Lynn Baxter	1983	MEN MADE in AMERICA

Manual Annotation with CATMA


CATMA 7.1.0
Scene_Annotation_Dime_Novels
Synchronized
COWBOY BE MINE TEXAS
Project
Tags
Annotate
Analyze
of made herself at home.
He blew out a breath in the frigid air, glanced one more time at
Bailey's house and turned his horse to head home.
The woman wasn't going to get under his skin.
No way.
"YOU SEE MY PROBLEM," Bailey told her older brother.
"I told you not to mess with him. I told you he wasn't going to
marry you," Brad said sourly.
"What you told me doesn't matter now, does it?" Brad eats is head
in. I should I should go over there and beat his head in. I should
1 /2
ANALYZE
Non-Scene is denoted in red
Scene is denoted in blue
Each line represents one person's annotation
Collection currently being edited
Mao_Annotation_TEXAS
Tagsets
Tags
Propert
Scene_An...
Non-Scene,Scene
ations
Tag
Aut...
"I have loved M[...]kin. No way
Non-Sc...
11...
"I have loved M[...]ichael Wade
Scene
10...

Inter-Annotator Agreement Calculation with GitMA

Author (Date)	Title	Cohen's k	Mathet's γ	Scene changes in gold annot.	words
Ferrarella (2000)	*Found: His Perfect Wife* AK	0.49	–	59	65,421
Broadrick (1986)	*Deceptions* CA	0.3	–	69	42,605
Stuart (1984)	*Tangled Lies* HI	0.3	0.31	88	69,124
Palmer (1985)	*Love By Proxy* IL	0.3	–	76	42,063
Campbell (1987)	*Pros and Cons* MA	0.38	–	91	75,527
Webb (2000)	*Warrior's Embrace* MS	0.21	0.34	118	58,903
McKenna (1984)	*Too Near the Fire* OH	0.78	0.41	39	43,319
Leonard (2000)	*Cowboy Be Mine* TX	0.52	0.53	65	62,421
Neggers (1989)	*Finders Keepers* VT	0.51	–	90	52,869
Cassidy (1997)	*Midnight Wishes* WY	0.39	–	100	60,655
10 Romance novels	from "Men Made in America"	ø0.4	ø0.4	795	572,907
20 translated novels	reuse from Zehe et al. (2025)	–	ø0.7	1,250	597,659
30 novels	total training set	–	–	2,045	1,170,566

Inter-Annotator Agreement Calculation with GitMA

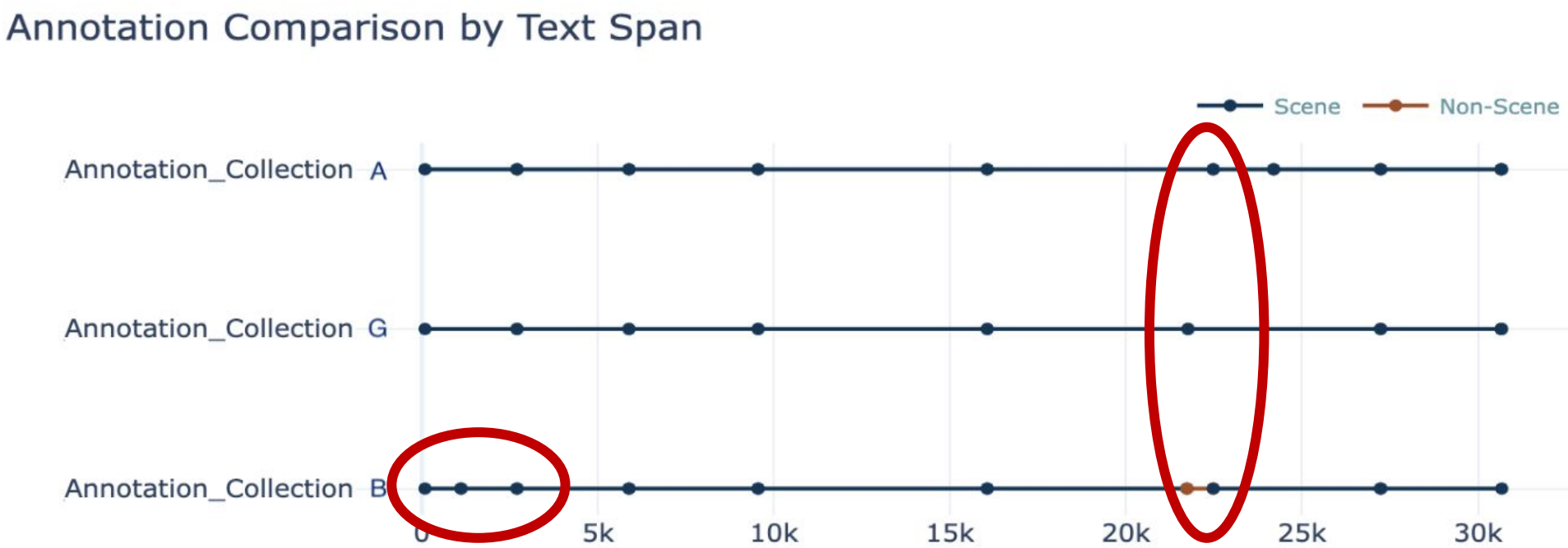


Figure 1: Comparison of two independent annotations (A+B, 0.35γ) with the gold annotation (G) in the middle. Visualization created with GitMA (Vauth et al. 2022) to demonstrate the gold annotation creation process.

General Observations

- There are more scenes than non-scenes
- Scenes tend to be longer than non-scenes
- The distributions and lengths of scene/non-scene differ for each novel
- Disagreements tend to appear when there are large chunks of reflective passages (e.g., psychonarration), particularly long/short scenes, non-scenes intersecting each other
- In general, our annotation agreement is high

Decisions for the Automation

Concentrate on scene change position detection

Combination of smaller scenes and non-scenes to sequences

Treat “scene“ and “non-scene“ distinction as another classification step

Delete typographical markers for scene changes like “* * *” and “Chapter”

Automation – How?

The easiest way: LLM-Experiments

IN A WORD: LLMs DO NOT PERFORM WELL (EXCEPT FOR DeepSeek)

Used Models:

- ChatGPT: GPT4, GPT4-o
- Claude: Claude 3.5-sonnet, Claude 3.5-haiku
- Gemini Pro
- Llama 3.2
- DeepSeek (new model from February)

Prompting:

- Step 1: Ask the model to learn the annotation guidelines
- Step 2: Give the text to the model
- Step 3: Evaluate model performance

LLM Experiment Example: GPT 4-o

Scene 4:
Before she could collect her wits, something struck her bent head and fell to her feet. Afraid to discover what it might be, she dared a glance and saw a black cowboy hat.

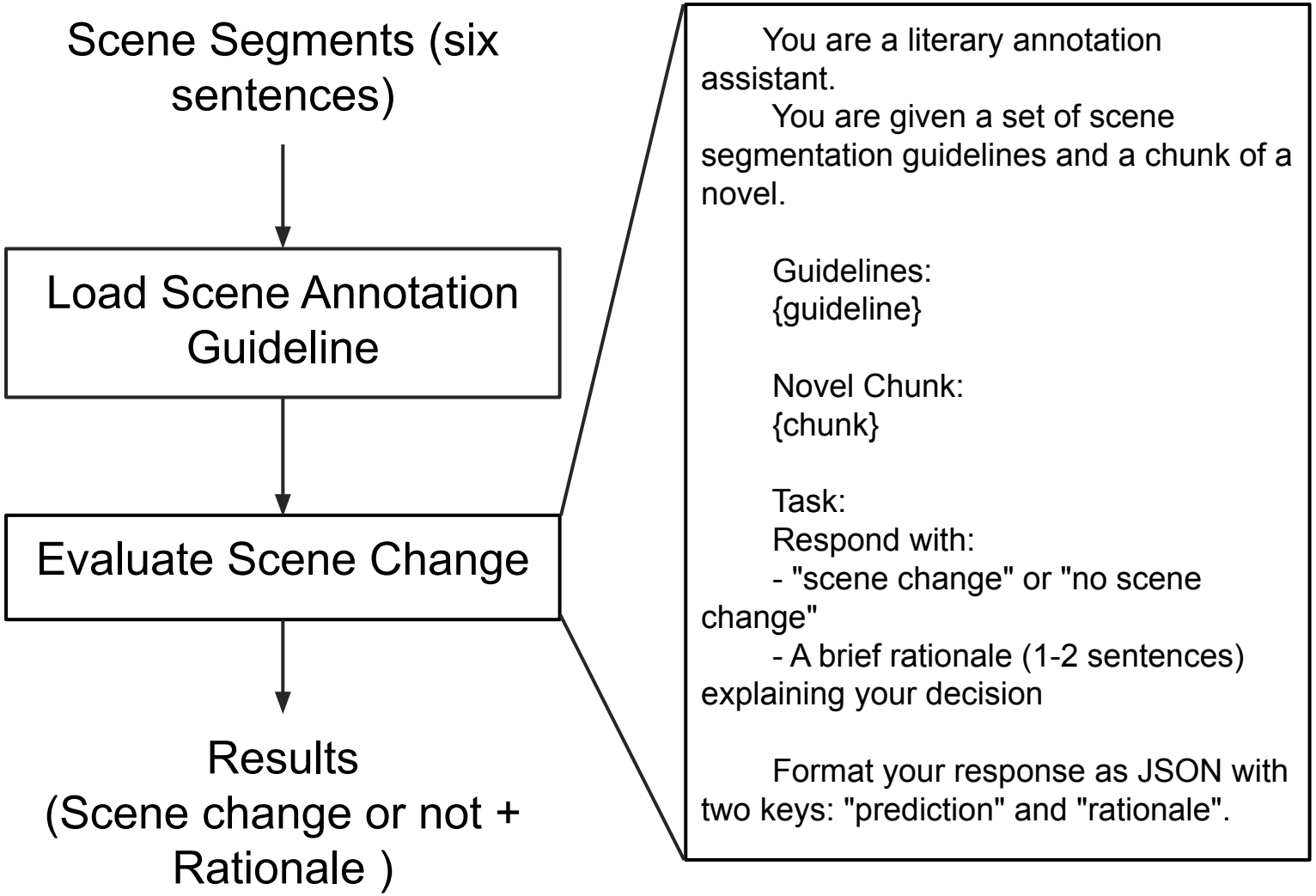
Scene 5:
"Where did this come from?" She picked up the hat and stared at it as she straightened. Gingerly holding the dusty object, she looked to the arena where the last rider stood waving at the cheering crowd—hatless.
"Hang on to it," Beth said over the din.

Scene 6:
Jules stared at her. "You hang on to it," she said, shoving the hat at her friend.

New LLMs Update: Automated workflow via API Calls

- Web interface → API Calls
- Constructed an automated scene change detection workflow
- Framework: Langgraph
- Model: GPT 4.1

Pipeline


Scene Segments (six sentences)
Load Scene Annotation Guideline
Evaluate Scene Change
Results (Scene change or not + Rationale)
You are a literary annotation assistant.
You are given a set of scene segmentation guidelines and a chunk of a novel.
Guidelines:
{guideline}
Novel Chunk:
{chunk}
Task:
Respond with:
- "scene change" or "no scene change"
- A brief rationale (1-2 sentences) explaining your decision
Format your response as JSON with two keys: "prediction" and "rationale".

Initial Results

- The last five scenes the last 5 of the 10 annotated novels, 493 scene segments
- > 80% accuracy

	ground_truth	prediction	rationale
ming Mrs. Michael Wade." She coul	no	no scene change	The passage maintains a consistent time, place (Bailey's bedroom), character focus (Bailey), and c
l a catch. Michael Wade would never	no	no scene change	The passage is a summary and reflection on the ongoing relationship between Michael and Bailey,
and some soup. Very casually she ha	no	no scene change	The passage describes a continuous action centered around the female character visiting and cari
few friends but many wary acquainta	no	scene change	The passage shifts from the woman's desire to tame Michael (implied reflective moment) to a direc
hat's what one of his high school ac	no	no scene change	This passage is reflective and summarizes background information about Michael's past and his c
ending more than physical pleasure.	no	no scene change	The passage remains focused on the narrator's thoughts about Bailey Dixon, with no change in tir
ssessed something in common, afte	no	no scene change	The paragraph maintains the same setting, time frame (a single February day), and focuses on Mic
e end of a long week. She hadn't be	no	no scene change	The passage consists of a continuous sequence of thoughts and reflections from a single characte
sn't matter now, does it?"	yes	scene change	There is a clear shift in both character constellation and action: the passage moves from a charact
o the children wouldn't understand t	no	no scene change	There is no significant change in time, location, character constellation, or independent action; the

Future Directions

- Enlarge the input. More context might increase accuracy
- Potential to support more flexible outputs, including scene summarization, scene and non-scene classification, etc.
- End to end pipeline: novel → disjoint set of scenes

LLM Experiment Example: GPT 4-o

Scene 1:
"I'm not sure this was such a good idea."
Jules Vandeveer didn't realize she had spoken as she stared across the dirt-floored indoor arena of the Agri-Plex. From her front-row seat next to her best friend, she watched the cowboy in the brilliant blue shirt position himself on the back of a horse.
"We can leave if you'd rather not stay," Beth Anders told her.
Jules was tempted, but strengthened her resolve. "No," she said, shaking her head. "No, I need to do this." She knew the dangers involved in any sport involving animals, but running away would not solve her problem. It was time to face her fears, and although she had trouble believing that anyone would be insane enough to make a career of tempting fate on the back of a bucking, twisting animal, she knew her reaction was based on those fears. "We'll stay," she said, determined to see this through.
Beth placed a hand on her arm, concern still in her eyes. "Hang in there. It's nearly over. This could be one of the best bareback bronc rides of the night."

The classic way: LM-Classifier

Transformer-based Classifier Approach

Aim:

- train a model of the BERT family to detect points of **SCENE CHANGE** positions instead of detecting entire scene segments

Preparation of Annotated Data

Model Design

- **Input:** a passage of n sentences
- **Output:** scene-change probability predicted for each passage

- **How to process a text:**

  Step 1: Input text

  Step 2: Format the text into several n-sentence passage

  Step 3: For each n-sentence passage:

  - Feed the passage to the model
  - Record any instances of scene change

  Step 4: Output result

Two Approaches

6-Sentences Approach (binary classification):

input: a 6-sentence passage

output: scene-change probability predicted for each passage

2-Sentences Approach (binary classification, but overfits too easily):

input: a 2-sentence passage

output: scene-change probability predicted for each passage

6-sentence segments + rolling window

30 manually annotated novels

→

data_preparation
txt files to xlsx
{ 6-sentence segment | Annotation }

→

fine-tuning pretrained English USE model for detecting the probability of a present scene-change

annotations in the training data

[...] She wouldn't. Not over Michael Wade.
<scene-change/>
The doorbell rang. Bailey stiffened, [...]

1	Chunk	scene-change
63	To Leah's relief, Apache got the	no
64	Leah sat in one of the old frayec	yes
65	In the rear of the main pumper	no
66	She lifted her chin and met his l	no
67	Hell, if you weigh over one-fort	no
68	"No way, honey, no way." He ra	no
69	Sam and Apache feel the same.	no
70	She gasped softly, turning in his	no
71	Gil looked around and then bac	no
72	If we have to make a run tonigh	no
73	"I don't know what you're talki	no

USE preprocessor and encoder from hub.KerasLayer google/universal-sentence-encoder-cmlm/multilingual

n epochs = 10

F1 Score: 0.70
0.45

6-sentence segments + rolling window

	Chunk	scene-cha	no	yes
85	TANNER O'BRIEN spotted t	yes	0.3038125	0.6961875
472	Chapter 2 Jules regarded th	yes	0.38467273	0.6153273
273	JULES WATCHED the cowb	yes	0.39314163	0.6068583
272	It didn't take much imagina	yes	0.40534353	0.5946565
471	He was sorry he probably w	yes	0.40981913	0.59018075
84	It was obvious she needed t	yes	0.44388056	0.55611944
78	Noticing the crowd had thi	no	0.57378095	0.42621905
29	When the shrill sound of a p	no	0.583957	0.41604295
386	Considering what had happ	no	0.6519512	0.34804875
332	The face of fourteen-year-o	no	0.68773824	0.3122617
57	"Jules spun around to see B	no	0.68976265	0.31023744
379	"I'll meet you at the car. "Ju	yes	0.7005487	0.29945132
35	With a wave of her hand, Be	yes	0.7208424	0.27915752
377	""Dr. Anders," someone cal	no	0.722908	0.2770919
499	"Seeing a group in obviously	no	0.7308313	0.26916867

```
Pulling Jules through the stragglers leaving the stands, Beth said with a
wicked grin, "We're going to a party." Jules smiled. She could handle a
little quiet mingling and a glass of wine to clear the dirt from her
throat, a place where she could relax and get her heart slowed to a more
normal pace. It was obvious she needed this vacation if a cowboy could
leave her tongue-tied.

<scene-change/>

TANNER O'BRIEN spotted the woman he'd seen in the stands with Beth Anders
the minute he walked into the noisy bar. Country music played at full
volume by a local band assaulted his ears, and multicolored lights flashed
in his eyes as a throng of well-wishers and words of congratulations
swamped him. And he still couldn't keep his eyes off her. Walking across
the wood-plank floor, he felt a friendly whack to his back and wordlessly
accepted the praise that accompanied it with a smile and a nod, while
someone else pressed a frosty mug of beer into his hand. Rodeoers and fans
were one big family, no matter what part of the country.
He thanked, smiled and nodded his way through the boisterous crowd to a
familiar face. Pulling up a chair, he straddled it.
```

Evaluation USE vs. GPT-4.1

Model Comparison:

Metric	USE Model	LLM Model	Better
Accuracy	0.6288	0.8418	LLM
Precision	0.3154	0.6481	LLM
Recall	0.8085	0.3723	USE
F1 Score	0.4537	0.4730	LLM

Summary: USE predicts scene changes more frequently while LLM is very cautious and only predicts scene changes when highly confident

Observations and Outlook

- USE Model detects the beginning of a new scene rather than the point of change between two scenes
- LLM predicts scene-change presence but fails to detect the beginning of a scene if given not enough context
- Refine the prediction to add <scene-change/> XML TEI element at the “right” position
- Even human annotators often disagree on the exact position of a scene-change (+/- 3 sentences)
- Next step: Analysis of the detected segments

Outlook:

Scene Segmentation

- Manual annotation
- Automated detection of scene change positions

Domestic Space

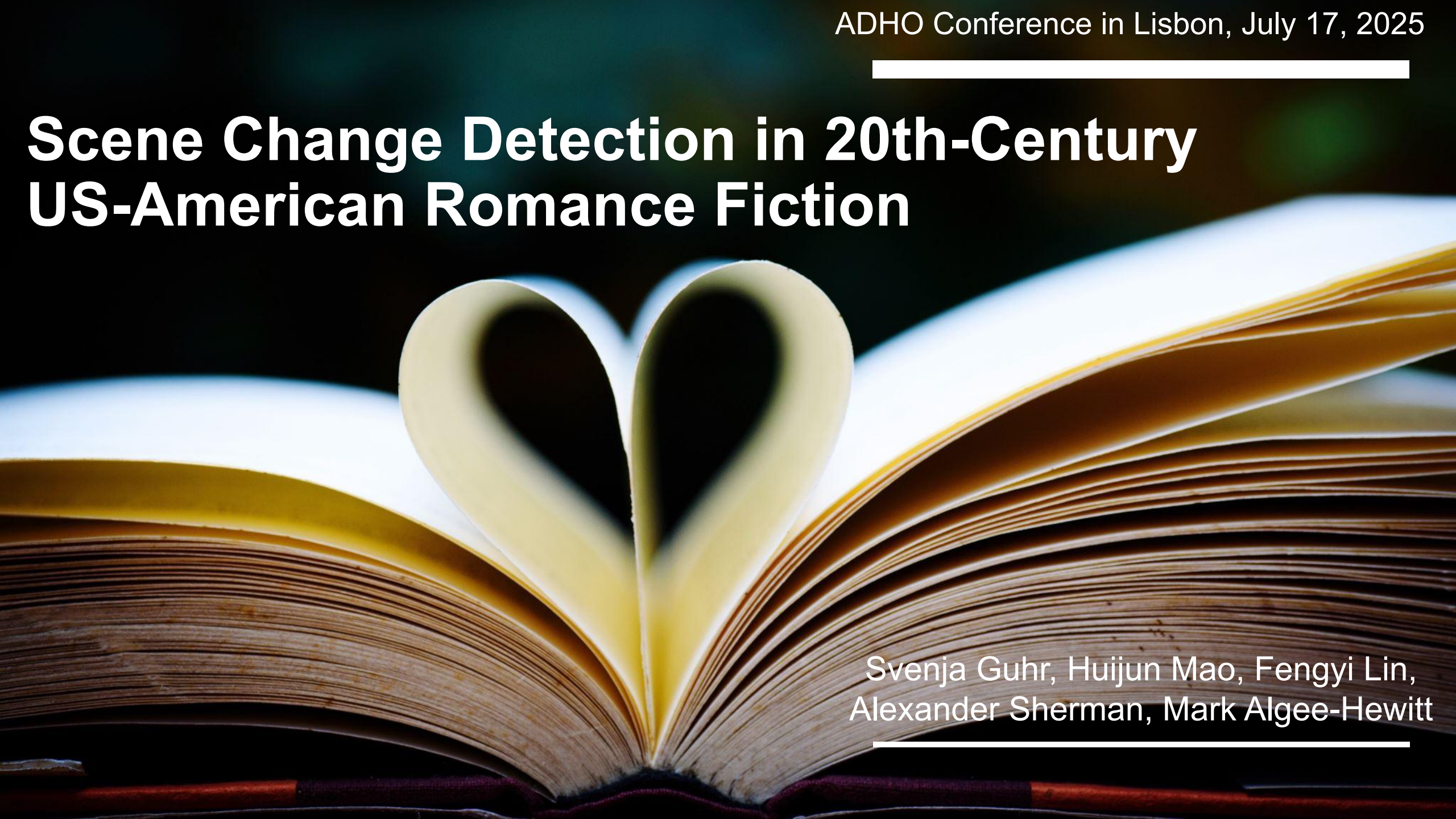
- Automated classification of scene segments
- Comparative analysis of domestic scenes

Gendered Stereotypes

- Automated character recognition
- Analysis of gendered stereotypes

ADHO Conference in Lisbon, July 17, 2025

Scene Change Detection in 20th-Century US-American Romance Fiction



Svenja Guhr, Huijun Mao, Fengyi Lin,
Alexander Sherman, Mark Algee-Hewitt

References

Gius, Evelyn, Fotis Jannidis, Markus Krug, Albin Zehe, et al. “Detection of Scenes in Fiction.” Paper presented at DH2019, Utrecht. Book of Abtracts, 12.07 2019. https://dev.clariah.nl/files/dh2019/boa/0608.html.

Gius, Evelyn, Carla Sökefeld, Lea Dümpelmann, Svenja Guhr, et al. “Guidelines for Detection of Scenes.” Zenodo, 2021. https://doi.org/10.5281/ZENODO.4457177.

Guhr, Svenja, Huijun Mao, and Fengyi Lin. “Rethinking Scene Segmentation. Advancing Automated Detection of Scene Changes in Literary Texts.” *Proceedings of the 9th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature (LaTeCH-CLfL 2025)*, Association for Computational Linguistics, 2025, 79–86. https://doi.org/10.18653/v1/2025.latechclfl-1.8.

Guhr, Svenja, Jessica Monaco, Alexander J. Sherman, Matt Warner, and Mark Algee-Hewitt. *Making BERT Feel at Home : Modelling Domestic Space in 19th-Century British and Irish Fiction*. CCLS Conference Reader, 2025. https://doi.org/10.26083/TUPRINTS-00030145.

Kurfalı, Murathan, and Mats Wirén. “Breaking the Narrative: Scene Segmentation through Sequential Sentence Classification.” *Proceedings of the Shared Task on Scene Segmentation*, 2021, 49--53. https://ceur-ws.org/Vol-3001/paper6.pdf.

Sherman, Alexander J., Svenja Guhr, Jessica M. Monaco, Matt Warner, Annie K. Lamar, and Mark Algee-Hewitt. “A Home without a `House’: Modelling Domestic Space in Nineteenth-Century British Fiction.” Paper presented at Digital Humanities, Washington D.C. *Book of Abstracts*, ADHO, 2024.

Zehe, Albin, Leonard Konle, Lea Katharina Dümpelmann, et al. “Detecting Scenes in Fiction: A New Segmentation Task.” *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, Association for Computational Linguistics, April 2021, 3167–77. https://www.aclweb.org/anthology/2021.eacl-main.276.

Zehe, Albin, Leonard Konle, Svenja Guhr, et al., eds. *Proceedings of the Shared Task on Scene Segmentation*. Vol. 3001. CEUR Workshop Proceedings. CEUR, KONVENS 2021, 2021. https://ceur-ws.org/Vol-3001/.

Zehe, Albin, Leonard Konle, Svenja Guhr, et al. “Shared Task on Scene Segmentation @ KONVENS 2021.” In *Proceedings of the Shared Task on Scene Segmentation*, edited by Albin Zehe, Leonard Konle, Lea Dümpelmann, et al., vol. 3001. CEUR Workshop Proceedings. CEUR, 2021. http://ceur-ws.org/Vol-3001/#paper1.

Zehe, Albin, Leonard Konle, Lea Dümpelmann, Evelyn Gius, et al. “Detecting Scenes in Fiction: A New Segmentation Task.” Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, Association for Computational Linguistics, April 2021, 3167–77. https://www.aclweb.org/anthology/2021.eacl-main.276.

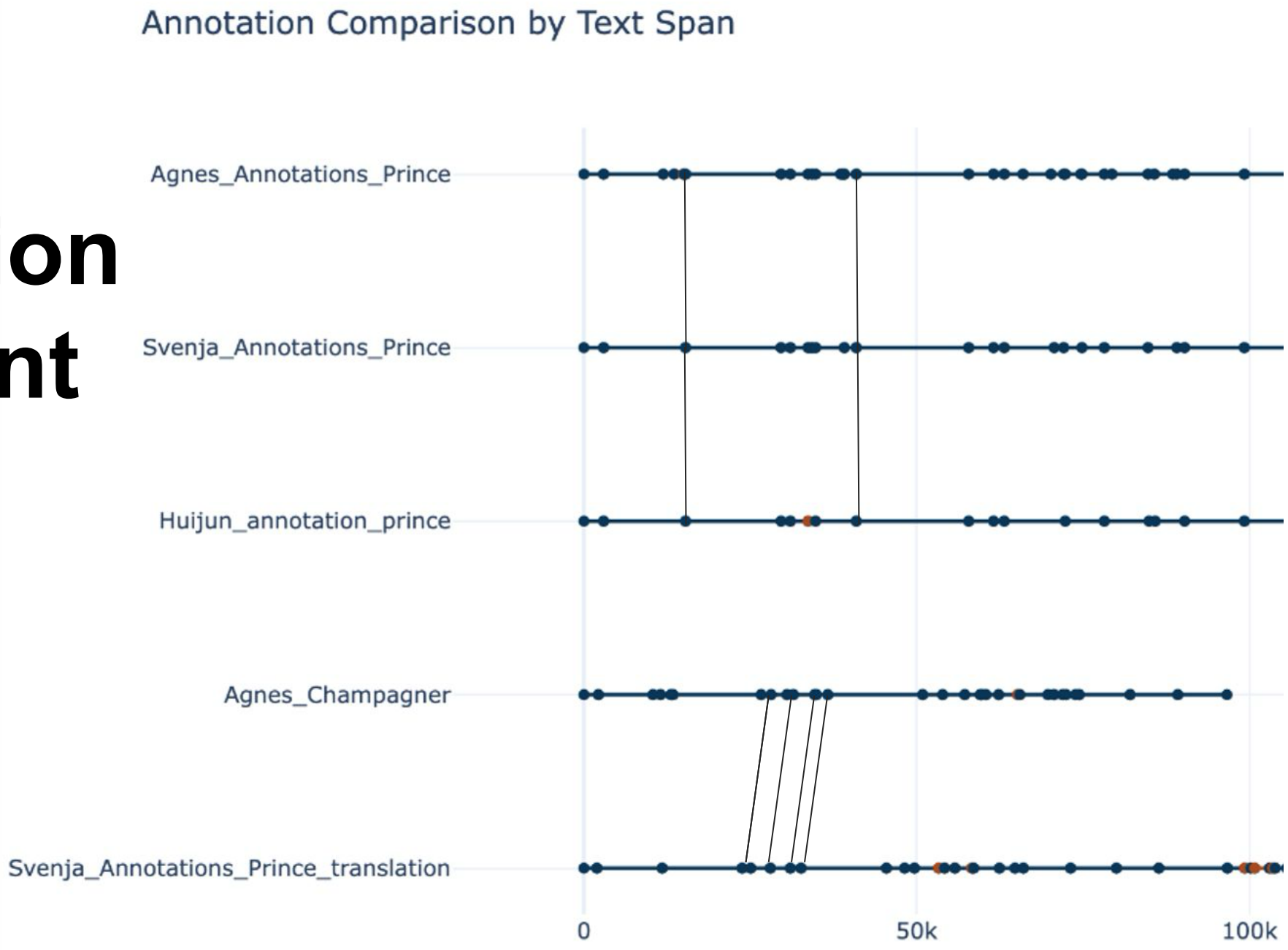
Zehe, Albin, Elisabeth Fischer, and Andreas Hotho. “Assessing the State of the Art in Scene Segmentation.” *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, Association for Computational Linguistics, 2025, 9922–41. https://doi.org/10.18653/v1/2025.naacl-long.500.

Appendix

Corpus
50 States - 50 “Men Made in America”

Corpus “Men Made in America”	
female authors	47
romance novels	50
words in total	5,5 Mio.
training set texts	10
words in total	578,562
manually annotated texts	8
words in total	464,929
scene changes in gold annotation	667
texts in annotation progress	2

Outlook: Translation Alignment



Observations (1/6): reflective passages are non-scenes

There was no controlling Bailey—she was as resilient and

headstrong as her mother had been. She'd do whatever she wanted to

do, and if she'd thrown him over for Gunner, then there wasn't a

whole lot he could do about it except hope his insides didn't feel

like worms were tunneling through them forever. He didn't think

he could stand it.

When he left the barn, he refused to look at the rambling house

again. It hurt too much. Keeping his gaze down as he strode to his

porch, he jerked off his leather gloves finger by finger, as if

he couldn't remove them without carefully observing his hands.

Observations (2/6): summaries are non-scenes

offer. She'd gone with him to see his office setup and learn everything her job would entail.

And had been horribly ill not ten minutes after she'd walked into the enormous King mansion. The cook had been preparing sausage links and beef tacos for the hands, and though she'd tried to fight off the green sensation stealing through her insides, she had barely made it to the bathroom Gunner swiftly helped her to.

Gunner had brought her home and assisted her to the worn red-and-white check sofa to sit. Her mother had loved to sit here and gaze out the big window at her children playing while she folded laundry. Bailey felt a twinge sitting in her mother's place, almost as if she could feel her mother's presence.

I've let Mother down, Bailey thought sadly.

Gunner stood, staring at her with concerned eyes she could hardly meet. She had to tell

Observations (3/6): there can be cliffhanger – not every chapter boundary is also a scene change position

Michael's eyes went wide. "Mother?"

"Hello, son. You needn't sound like I'm a nightmare you've just awakened to, though from your tone I can tell you weren't expecting me?" She gazed at him steadily, then slowly eyed Bailey from toe to head, settling, it seemed, on her lips. Nodding as if she'd seen all she needed to, she set a large suitcase on the floor with a thump. "Now what's this I hear about you needing a chaperon?"

Chapter 8

Michael was so stunned he couldn't move. Bailey poked him purposefully. He walked forward to awkwardly kiss his mother's cheek. With his mother's

Observations (4/6): time changes indicate scene changes

"Thanks. Goodbye, Brad. We'll celebrate when you get home." Bailey was breathless as she ushered the children into the black truck. She didn't care what anyone in Fallen thought - she couldn't handle another moment of being on display with Michael. Not when it felt like her heart was being torn right out of her.

TWO HOURS LATER, Michael sat in his den trying to figure out what had happened. The Dixons' roof was finished and secure, and they had packed and disappeared before he'd returned. If he'd known today was going to blow up in his face, he wouldn't have hired the extra work crew. He'd feel better if

Observations (5/6): space changes indicate scene changes

There was no reason to say more. It was up to Michael and her to resolve how

they were going to deal with the baby they were expecting.

MICHAEL SENT the bunkhouse door flying open. He strode to the den where the

fence-sitters lay pitifully, propped on pillows, their faces pasty. Fred still

looked worse than any of them. Michael nearly softened as he looked at the

men, but then he thought about what they'd done to him and got angry all over

Observations (6/6): character changes can indicate scene changes

"Aw," Chili said, helping himself to the bowl of peanuts Deenie had set out,

as well as a bottle of wine chilling in a bucket. "She who lies in wait in the

dark ought not be too upset when she gets ambushed better." He laughed. "Miss

Fancy Pants wants herself a

cowboy, but she ain't gonna get him, is she, boys?"

"Nope," Fred and Curly agreed.

"BABY, IS BAILEY HOME?" Michael had hoped Bailey would open the door, but Baby

liked to greet every-one. The lamb bleated a welcome, too, dancing nervously

on the worn hardwood floor.