

G4: Define & present initial information story for product: Wireframes, access & architecture

Tara Hulcome

IMT 542 A: Portable Information Structures





| | | |
|-----|-----------------------------------|-----|
| • • | | |
| • • | A NEW DIGITAL INFORMATION PRODUCT | p.1 |
| • • | | |
| • • | INFORMATION WIREFRAME | p.2 |
| • • | | |
| • • | MAIN CONCEPTS | p.3 |
| • • | | |
| • • | BASIC SYSTEM ARCHITECTURE | p.4 |
| • • | | |
| • • | ANNEX | p.5 |
| • • | | |

TABLE OF CONTENTS

FROM PEN TO PIXEL: A NEW DIGITAL INFORMATION PRODUCT

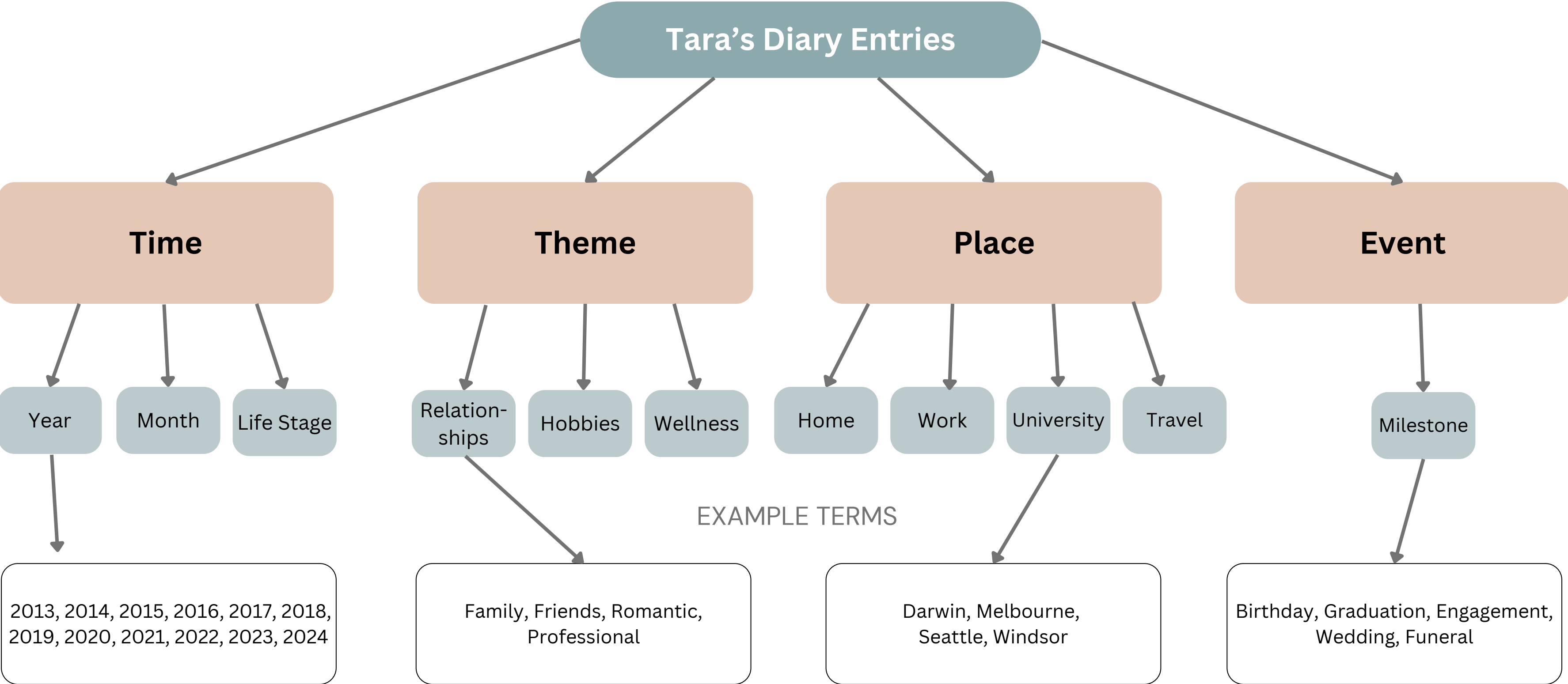
- **What:** Convert handwritten personal diaries into structured, machine-readable format.
- **Goal:** Enhance accessibility & findability, enabling the user (me!) to interact with data in a more enjoyable way.

| EXISTING SYSTEM | Hard copy files ✓ | Digital files ✓ | | New digital information product |
|-------------------|---|---|---|---|
| Scope / Format | <ul style="list-style-type: none">• 22 physical diaries (~8,030 entries) | <ul style="list-style-type: none">• ~1,144 scanned JPEG files. | → | <ul style="list-style-type: none">• Convert all JPEG files to JSON format. |
| Storage Solution | <ul style="list-style-type: none">• Hard copies stored in shoe boxes. | <ul style="list-style-type: none">• Stored in cloud.• Password-protected. | → | <ul style="list-style-type: none">• One consolidated database → easy scale up with new entries. |
| Access / Analysis | <ul style="list-style-type: none">• Very hard to retrieve.• Analysis impossible. | <ul style="list-style-type: none">• Poor folder structure hinders retrieval / analysis. | → | <ul style="list-style-type: none">• Structure information with new taxonomy and topic filters.• Interact through simple API. |
| Risks | <ul style="list-style-type: none">• Vulnerable to loss, exposure or damage. | <ul style="list-style-type: none">• Potential for data breaches. | → | <ul style="list-style-type: none">• Need for manual quality audit.• Security is <u>critical</u>. |

↖ (what I've already got)

(what I want) ↗

GIVING STRUCTURE TO THE STORIES: INFORMATION WIREFRAME



STANDARDIZED EXPRESSION: MAIN CONCEPTS

- Develop controlled vocabulary list (to align with taxonomy in *Slide 2*)
- Four top-level categories
 - Time
 - Theme
 - Event
 - Place
- Defined sub-categories & specific terms
 - *See example* →

AN ITERATIVE PROCESS!

- This vocabulary list is just a starting point
 - will be refined as project progresses.

EXAMPLE CONTROLLED VOCABULARY LIST

| Top-level category | Sub-category | Term | Example tags |
|--------------------|---------------|----------------------------|-------------------------------|
| Time | Year | 2020 | #time-year-2020 |
| Time | Year | 2021 | #time-year-2021 |
| Time | Year | 2022 | #time-year-2022 |
| Time | Year | 2023 | #time-year-2023 |
| Time | Year | 2024 | #time-year-2024 |
| Time | Month | January | |
| Time | Month | February | |
| Time | Month | March | |
| Time | Month | April | |
| Time | Month | May | |
| Time | Month | June | #time-month-june |
| Time | Month | July | |
| Time | Month | August | |
| Time | Month | September | |
| Time | Month | October | |
| Time | Month | November | |
| Time | Month | December | |
| Time | Life Stage | High School | |
| Time | Life Stage | University - Undergraduate | |
| Time | Life Stage | University - Graduate | |
| Time | Life Stage | Cadetship | #time-life_stage-cadetship |
| Time | Life Stage | Career | |
| Time | Life Stage | Van Life | |
| Theme | Hobbies | Music | |
| Theme | Hobbies | Camping | #theme-hobbies-camping |
| Theme | Hobbies | Baking | |
| Theme | Hobbies | Cooking | |
| Theme | Wellness | Hiking | |
| Theme | Wellness | Yoga | #theme-wellness-yoga |
| Theme | Wellness | Pilates | |
| Theme | Wellness | Running | |
| Theme | Wellness | Strength | |
| Theme | Wellness | Aerial | |
| Theme | Wellness | Kayaking | |
| Theme | Wellness | Skiing | |
| Theme | Relationships | Family | |
| Theme | Relationships | Friends | |
| Theme | Relationships | Romantic | #theme-relationships-romantic |
| Theme | Relationships | Professional | |
| Event | Milestone | Birthday | #event-milestone-birthday |

| Top-level category | Sub-category | Term | Sub-Term | Example tags |
|--------------------|----------------|----------------|---------------|----------------------------------|
| Place | Home | Australia | Darwin | #place-home-australia |
| Place | Work | Australia | Canberra | |
| Place | Work | Australia | Gold Coast | |
| Place | Work | China | Beijing | |
| Place | Work | Fiji | Suva | |
| Place | Work | India | New Delhi | |
| Place | Work | Indonesia | Jakarta | |
| Place | Work | Japan | Tokyo | |
| Place | Work | PNG | Port Moresby | |
| Place | Work | Russia | Vladivostok | |
| Place | University | Australia | Darwin | |
| Place | University | Australia | Melbourne | |
| Place | University | Canada | Windsor | #place-university-canada-windsor |
| Place | University | United States | Seattle | |
| Place | Leisure Travel | Australia | Adelaide | |
| Place | Leisure Travel | Australia | Brisbane | |
| Place | Leisure Travel | Australia | Coffs Harbour | |
| Place | Leisure Travel | Australia | Gold Coast | |
| Place | Leisure Travel | Australia | Melbourne | |
| Place | Leisure Travel | Australia | Sydney | |
| Place | Leisure Travel | Australia | Hobart | |
| Place | Leisure Travel | New Zealand | Auckland | |
| Place | Leisure Travel | New Zealand | Queenstown | |
| Place | Leisure Travel | New Zealand | Christchurch | |
| Place | Leisure Travel | New Zealand | Taupo | |
| Place | Leisure Travel | United Kingdom | London | |



FROM PAPER TO THE PLATFORM: BASIC SYSTEM ARCHITECTURE

1

Technologies

- Extract text from JPEG files using Optical Character Recognition (OCR) tools and Python.
- Use simple API (to be determined) to interact with information.

2

Hosting

- To be determined.
- Will be assessed after completion of scoping and testing phase.

3

Information structure

- Aim is to focus on the process rather than building the 'perfect' information product.
- Will experiment with a small subset of files to test and refine the structure.

4

Future expansion

- Full implementation unlikely within scope of IMT 542 project, but plan to pursue later.
- Potential to conduct advanced sentiment analysis and synthesis, leveraging AI tools
 - But requires further thinking re: risk management.

ANNEX



Figure 1: Personal collection of hard copy diaries

| | | |
|---|---|-------------------------------|
| ▼ | 📁 | The Diary Project |
| > | 📁 | 2003 |
| > | 📁 | 2004 |
| > | 📁 | 2005 |
| > | 📁 | 2006 |
| > | 📁 | 2007 |
| > | 📁 | 2008 |
| ▼ | 📁 | 2009 |
| | 📷 | Photo 28-1-20, 8 32 31 pm.jpg |
| | 📷 | Photo 28-1-20, 8 32 38 pm.jpg |
| | 📷 | Photo 28-1-20, 8 32 43 pm.jpg |
| | 📷 | Photo 28-1-20, 8 32 51 pm.jpg |
| | 📷 | Photo 28-1-20, 8 32 56 pm.jpg |
| | 📷 | Photo 28-1-20, 8 33 13 pm.jpg |
| | 📷 | Photo 28-1-20, 8 33 19 pm.jpg |
| | 📷 | Photo 28-1-20, 8 33 25 pm.jpg |
| | 📷 | Photo 28-1-20, 8 33 30 pm.jpg |
| | 📷 | Photo 28-1-20, 8 33 50 pm.jpg |
| | 📷 | Photo 28-1-20, 8 33 56 pm.jpg |
| | 📷 | Photo 28-1-20, 8 34 01 pm.jpg |
| | 📷 | Photo 28-1-20, 8 34 07 pm.jpg |
| | 📷 | Photo 28-1-20, 8 34 35 pm.jpg |
| | 📷 | Photo 28-1-20, 8 34 43 pm.jpg |

Figure 2: Current Dropbox file structure (JPEG files containing scanned diary entries)