

Visual Summary 규약 v0.5

내용시각화 Summary 고도화

2026-08-04 MORDOR 체크인 직후 세션

정원석 (MORDOR)

이 텍스트는 파생물이다. SSoT 는 `mordor-wiki / projects/mordor/visual-summary/` 다.

할 일은 `TODO.md` , 규약은 `spec/schema-v0.5.md` 가 갖는다.

이 세션이 하려는 것

여기서 할 것

규약은 2026-08-02 에 확정됐다.

규약이 닿지 않는 구간을 각자 소관에서 채운다.

- 색인 쪽에 무엇이 필요한가
- 추출 쪽 경로가 있나
- 언제 누가 하나

여기서 안 할 것

확정된 결정을 처음부터 되묻기.

닫힌 것과 열린 것은 슬라이드에 표시했다.

닫힌 것은 `TODO.md` G절, 열린 것은 B절이 SSoT 다.

1분 요약 사용자 입력을 컴포넌트로 바꿀 때 구조가 뭉개지는 것이 지금 품질 문제의 핵심이다. Summary 를 Markdown 에서 구조화 형식으로 바꿔 슬롯을 보존하고, 컴포넌트 쪽에는 의미 태그를 붙여 대조한다.

먼저 오해 하나를 지운다

"사용자 입력이 RLSC 로 변환된다" **아니다**. 경로가 둘이고 별개로 진행된다.

질의 쪽 (query side)

사용자 입력을 검색 조건으로 바꾼다

`Summary`, `structure`

색인 쪽 (index side)

이미 있는 템플릿 컴포넌트를 미리 검색 가능한 형태로 뽑아 둔다

`RLSC`, `Search Signature v2`

질의 쪽 조건



BE 검색 모듈에서 합류



색인 쪽 파생물

`structure` 에서 RLSC 로 가는 화살표는 변환이 아니라 capability 대조다. 그 규약이 `capability-table.md` 다.

왜 하나: BAD 의 85% 가 매칭 문제다

렌더링 문제가 아니다

2026-07-26 BAD 원인 분석

judge 채점 16,819행 집계 +

BAD 200건 직접 분류

→ 85% 가 컴포넌트와 내용이 안 맞는 문제

세부 원인 넷

- 텍스트 용량 초과
- 매핑과 위계 오류
- 슬롯 수 불일치
- 대치 실패

검색이 컴포넌트를 잘 골라도 그 다음이 정의돼 있지 않아 결과가 나빠지고 있다.

2026-08-03 방향 1.5개월 이내에 구조 불일치를 해결한다. 체감 BAD 50% → 20% 이하가 기대치다.

무엇이 뭉개지나: 같은 입력, 두 결과

② Markdown Summary (현행)

2025 사업 성과

- * 총매출 152억
- * 온라인 매출 22억

값과 단위, 계열이 문자열로만 남는다.

「152억」이 숫자인지 텍스트인지 구분이 없다.

②' Visual Summary (YAML)

```
components:
  - type: KPIKeyMetric.BigNumber
    structure:
      entities:
        - { label: 총매출,
            value: 152, unit: 억원 }
```

슬롯이 그대로 남는다.

① 사용자 입력



② Summary



③ 검색 조건



④ 후보 컴포넌트



⑤ 대치

로드맵: 단계마다 넷이 바뀐다

단계	Summary	검색	태그 범위	태그가 가리키는 것
1 현행	Markdown	임베딩 유사도	안 단다	-
2 목표 2026-08-18	Markdown	semantic tag	컴포넌트 하나	이 컴포넌트 어딘가에 그 타입이 있다
3	YAML	tag + structure	컴포넌트 하나	2단계와 같음
4	YAML	tag + node 태그	컴포넌트 안 node 마다	어느 node 인지 특정된다
5 미정	YAML	4단계와 같음	node 마다	4단계 + 카테고리별 property

3단계는 실행하지 않는다

3단계에서 만들 대치 로직이 4단계에서 버려질 수 있다. 정보가 확실한 4단계로 바로 간다 (7/24 회의)

배포 일정은 다시 봐야 한다

2026-08-03 에 제품 배포가 연기됐다. 2단계 마감 8/18 이 그 연기에 포함되는지 티켓 반영이 없다

왜 단계를 나누나: 태그의 모호성

컴포넌트 하나에 태그를 다는 한, 두 경우를 구분할 수 없다.

경우 A

root 자체가 그 타입이라 붙었다

경우 B

하위 어딘가에 그 타입 node 가 있어서 컴포넌트 전체가 받았다

실데이터가 B 의 존재를 증명한다. design_object_id 4757227 에 ComparisonTable, HierarchicalTree, List, HighlightedMessage 넷이 함께 붙어 있다. root 가 넷을 동시에 가질 수는 없다.

복수 태그는 83,771건 중 260건(0.3%) 이고 최대 4개까지 관측된다.

4단계에서 node 단위로 태그를 달면 해소된다. 그 일을 I.I.S. 가 맡는다.

설계의 실증 근거: 검색 풀이 리스트에 쏠려 있다

색인 쪽 (붙어 있는 태그)

모수 83,771 design object 2026-07-22 적재

카테고리	비율
List	81.1%
BigNumber	3.6%
Steps	2.4%
ListingTable	0.5%

상위 8종이 94.5%

질의 쪽 (사용자가 원하는 것)

카테고리	비율
List	31.96%
ListingTable	15.08%
Steps	10.92%
ComparisonTable	7.24%

수요는 다양한데 풀이 쏠려 있다

이 분포가 설계를 정했다. VennDiagram, Tree, NestedScope, Pyramid 는 각 164건 이하, 데이터 차트는 각 51건 이하다. 성립 조건을 엄격하게 걸면 후보가 0이 되어 전부 List 로 강등된다.

v0.5 가 정한 것 (2026-08-02 확정)

항목	결정
최상위 형태	항상 <code>components[]</code> . 하나면 길이 1 배열
<code>type</code> 표기	문자열 하나로 대분류. 중분류. v7 단어장 49종 중 41종
<code>confidence</code>	제거. 생성기가 채우지 않는다
<code>alternatives[]</code>	신설. fallback 순서보다 먼저 본다
<code>entity.id</code>	생성물에서 뺀다. 파싱 뒤 배열 순서로 매긴다
<code>dataset</code>	<code>measure</code> 와 <code>matrix</code> 두 모양. 표 계열도 <code>matrix</code>
<code>relation.type</code>	<code>intersects</code> 와 <code>compares</code> 둘
생성 방식	LLM 이 YAML 을 직접 쓴다. JSON 변환 후 스키마 검증
원소 수 상한	두지 않는다. 원소 간 관계는 자리만 예약

검증 실측 스키마 Draft 2020-12 유효성 통과 · 본문 예시 18종 전량 통과 · 부정 사례 12종 의도대로 거부

type: 단어장 49종 중 41종



왜 뺐나

페이지 역할은 의미 구조가 아니라 페이지에서의 자리다
지도 전용은 지도 데이터가 있어야 성립해 `structure` 가 표현할 대상이 아니다

다시 열린 질문 (B-24)

49종 전량으로 열어야 하지 않나
enum 은 capability 의 파생물인데 capability 자체가 41종이라 파생 이전에 잘려 있다
페이지 역할은 색인에 실재한다 (v3.5)

허재영님 지적 2026-08-03. 함께 볼 것: 대분류를 없애고 중분류만 남겨도 되는가 (B-25)

capability 테이블: 셋을 선언한다

선언	무엇인가	못 채우면
성립 조건	이게 없으면 그 종류로 배치가 불가능한 것	그 종류가 후보에서 빠진다
선택 속성	안 적어도 되는 것. 규약이 기본값을 정해 둔다	기본값으로 그린다
fallback 순서	그 종류가 색인에 없을 때 물러날 곳	순서를 따라 내려가고 마지막은 항상 List

성립 조건을 최소로 줄였다

내용을 요구하는 것은 다섯뿐이다

BigNumber·Gauge 의 수치, Tree·NestedScope 의 `children`,
ComboChart 의 혼합 조건

나머지는 개수와 `matrix` 유무만 본다

기본값 규칙

기본값은 표현 선택에만 준다. 내용에는 주지 않는다.

`apex` (어느 쪽이 꼭대기) → 기본값 OK

`value` (숫자가 얼마) → 없는 데이터를 지어내는 것이므로 성립 조건

fallback: 빈 결과가 사용자에게 가지 않는다



종단이 셋이다

List · LongFormTextLayout · HighlightedMessage

셋 다 개수 조건만 보므로 언제나 하나는 성립한다

→ 대치 실패 정책을 따로 두지 않았다

엣지 55개에 projection 을 정의했다

2026-08-03

손으로 쓰지 않고 규칙 여덟 개를 세운 뒤 각 엣지에 어느 규칙이 걸리는지를 capability 에서 계산했다

→ 표가 capability 와 어긋날 수 없다

계산 중에 결함 하나가 드러났다. 「선택 속성」열이 내용 필드(period , speaker)와 표현 선택(orientation , apex)을 섞어 담고 있었다. 그대로 계산하면 Timeline → Steps 에서 시각(period)을 버리게 된다. 갈라서 내용은 옮기고 표현만 버리도록 고쳤다.

새로 정한 것: 꾸밈구조

2026-08-03. 어제까지의 결정을 뒤집었다.

무엇이 뒤집혔나

항목	2026-08-02	2026-08-03
계층	structure 에 선언하지 않는다	컴포넌트를 nested 구조로 두어 해결한다
꾸밈	질의 쪽 필드를 두지 않는다	꾸밈구조를 둔다. 다만 1단계는 전역 기준만

뒤집은 이유 1

8/2 결정이 판정 책임을 색인 쪽으로 넘겼는데, 색인 쪽에 그것을 받을 선언이 없었다

capability-table.md 에 꾸밈 언급이 0건

뒤집은 이유 2

그 결정의 근거표가 보존 전용 문서에만 있었다

허브가 「v0.5 는 이 철학을 따르지 않는다」고 적어 둔 문서다

SSoT: design/decoration-model.md · 할 일: TODO.md B-2, B-3, B-18~B-25

왜 등급 하나로는 안 되나

7/24 회의의 N / $N+\alpha$ / $N+\beta$ 는 누적 등급이다. $N \subset N+\alpha \subset N+\beta$

1. 다른 것이 섞인다

$N+\beta$ 가 「요소 배치와 장식」이다
배치가 정교한데 그림이 없는 컴포넌트가 있다

2. 조합을 못 적는다

「배경은 원하지만 일러스트는 싫다」에 이름이 없다

3. 끼울 자리가 없다

배경·테두리·아이콘이 생기면 「 $N+\beta$ 안쪽」이라 할 수밖에 없다

4. N 은 애초에 꾸밈 축이 아니다

N 의 「최소 필요 조건」은 capability 성립 조건과 같은 대상이다. 등급 최하단으로 두면 같은 사실이 두 곳에 생긴다.

그래서 축 집합으로 둔다

```
decoration:           # 꾸밈구조 하나 (component.decoration)
  ornament: prefer    # 꾸밈 축 하나 + 선호값 하나
  background: avoid   # 꾸밈 축 하나 + 선호값 하나
```

축은 RLSC role 에 대응한다

축	RLSC role
ornament	Element.Decoration
background	Element.Background

축이 role 하나를 가리키면 추출이 이미 붙인 표식을 쓸 수 있다

선호값은 셋이다

값	뜻
prefer	있으면 가점
indifferent	보지 않는다 (기본값)
avoid	있으면 감점

require 를 두지 않는다. 꾸밈을 필수로 걸면 후보가 0이 된다

minimal (N) 과 rich (N+β) 는 문서와 회의에서 부르는 별명이고 스키마 값이 아니다. 축이 늘 때 뜻이 바뀌기 때문이다.

판정 기준 하나로 축이 갈린다

대치가 내용으로 채우는 노드는 선호가 필요 없다. 내용이 이미 답이다.
대치가 건드리지 않는 노드만 선호가 필요하다.

RLSC role	대치가 채우나	그래서
Element.Title · Subtitle · Description	채운다	내용이다. entity.label 과 entity.components[] 가 담는다 → 축 아님
Element.Highlight	채운다	내용이고 entity.emphasis 가 이미 있다 → 축 아님
Element.Decoration	안 채운다 role 정의: "내용 대치 시 변경 불필요"	입력에 대응물이 없다 → 축 필요
Element.Background	안 채운다	위와 같다 → 축 필요

부수 효과 N+α(타이틀·디스크립션 등급)는 개념 자체가 사라졌다. 표현 밀도가 아니라 내용 유무다. 그리고 계층과 DESC 가 entity.components[] 한 수단을 쓰게 됐다.

1단계 범위: 전역 기준만 잡는다

총	1단계 (지금)	요소 대치 착수 후
Summary (질의)	생성하지 않는다. 스키마에 optional 로 자리만	필요하면 생성기가 적는다
카테고리 기본	미정	41종에 채운다
시스템 전역 기본	여기만 잡는다 예: 꾸밈 많은 것을 우선	유지

왜 미루나

- 지금 소비할 주체가 없다. 요소 대치가 아직 동작하지 않는다
- 색인이 대조할 대상을 못 받는다
- 지금 필요한 것은 전역 기준 하나로 된다

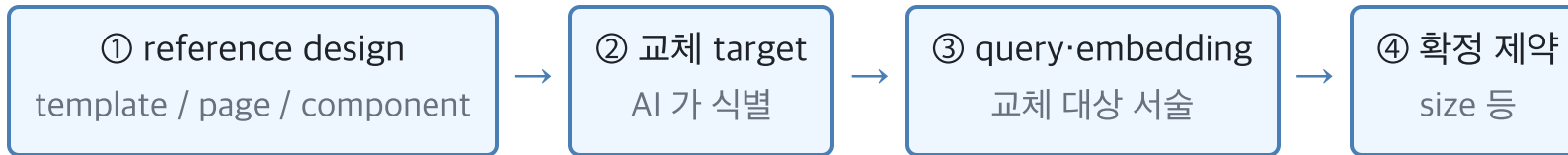
선례가 있다

`entity.components[]` 와 `page relations` 가 모두 「스키마에 열어 두고 생성기는 채우지 않는다」
`decoration` 도 같은 방식으로 연다

내용 대치는 꾸밈에 관여하지 않는다. 꾸밈구조는 앞으로 요소 대치가 참고하는 값이다.

요소 대치: 꾸밈구조를 쓰는 곳

신동한님 쪽 모듈(AI 꾸미기)을 참고해 구현하기로 했다 2026-07-23 회의



난점은 ②와 ③이다

7/23 회의 원문: 「비슷한 요소 찾기 사용. 다만 교체 타겟 결정 및 타겟별 query 생성이 중요」

검색 자체보다 무엇을 바꿀지 정하고 어떻게 서술할지가 문제다

정하지 않은 것

reference design 이 꾸밈 수준을 이미 정해 버리면 꾸밈구조는 요소 대치에 안 쓰이고 검색 단계에만 쓰인다

7/23 기록은 「장식 요소는 유지하고 내용 관련 요소만 대치」다

다섯 구간이 모두 대응해야 동작한다

구간	무엇을 하나	지금 상태
추출	노드에 Decoration · Background role 을 붙인다	role 은 이미 붙는다
색인	그 role 을 얼마나 가졌는지 색인에 담는다	경로가 없다 text_map · text_node_roles · semantic_tags 는 있는데 노드 분류 자리가 없다
검색	기본값과 선호를 대조해 순위를 움직인다	capability 에 꾸밈 축 열이 없다
Visual Summary	그 입력의 선호를 적는다	표기 미정 (B-18)
대치	선호에 맞춰 role 노드를 지우거나 남긴다	요소 대치의 일. 아직 동작하지 않는다

색인 경로가 먼저 풀려야 나머지가 동작한다. 추출이 role 을 붙여도 색인이 못 받으면 검색이 대조할 대상이 없다. semantic_tags denormalize 방식이 선례다. → TODO.md l-4, B-22

누가 무엇을 소유하나

구간	담당
단어장 v7 정의와 유지	김수현
Summary 생성기	스키마 정원석 · 생성 실무 허재영, 황현진
시멘틱 쿼리 추출 LLM · 컴포넌트 태깅	황현진
태깅 결과 수신·검증·적재	류민 → 박지우
노드 분류와 카테고리별 속성	신영빈 · 류민
I.I.S. (4단계)	류민 2026년 6월 중순부터 정지
시그니처·랭킹·대치 계산 (comp-sign)	이석
Visual Summary 규약	정원석
검색과 대치 실행	최운식 (BE)

layouts 태깅은 BE 로 가지 않는다. RAT pod 내부용이고 검색에는 컴포넌트 태깅만 쓴다.

오늘 각자에게 물을 것

강창석님

- `dataset` 을 DataViz Matrix 로 바로 넘길 수 있나 (어댑터 실측)
- `nested_scope` 를 `concentric_rectangles` 로 볼 수 있나
- 스마트 다이어그램 32종 중 2단계에 열 것이 있나

류민님

- I.I.S. 재개 시점
- `tree` children → Graph relations 규칙
- 카테고리별 속성 목록 초안 리뷰

이석님

- semantic 만 빼고 검색·대치가 되나. toggle 이 있나
- 세 축 정렬을 언제 누구와
- `rankCompCandidates` 위계 가중치 A/B
- 41종 fallback 순서 검토

허재영님

- Venn 라벨 오류 프롬프트 225건 중 148건이 오라벨
- `entity.components[ ]` 생성 시점
- 평가셋 38 중분류 확장

급한 순서와 오늘 정할 것

급한 순서 2026-08-02 기준

1. C-1 모듈 버전 2.0 BE 정책 협의
2. C-2 semantic tag 데이터 흐름 공유
3. C-6 태깅 증량분 재전달 9만 → 16만
4. C-7 컴포넌트 분포 재측정
5. I-2 세 축 정렬
6. I-6 위계 가중치 A/B

오늘 정할 것

- 다음 주요 작업 3축 분담 서머리 생성 / 스마트시리즈 / 조합
- 2단계 마감 8/18 유지 여부
- 검색폴 17만 증량분에 judge 를 돌릴지
- 꾸밈구조 표기와 `entity.components[]` 생성 시점
- 41종 vs 49종 (B-24), 대분류 제거 (B-25)
- List 편중 81% 를 두고 2단계를 낼지

참고

먼저 볼 것 둘

`spec/schema-v0.5.md` 규약

`TODO.md` 할 일 SSoT

이 텍스트의 근거

`spec/capability-table.md` 41종

`design/decoration-model.md` 꾸밈구조

`design/pipeline-flow.md` 전체 흐름

`design/content-binding-contract.md`

`reference/semantic-taxonomy-v7.md`

`README.md` 로드맵과 분포

저장 위치 `github.com/miridih/mordor-wiki` → `projects/mordor/visual-summary/`

GitHub 계정이 없으면 Confluence 발행본으로 공유한다 (`TODO.md` C-8)