

# MongoDB 정리

아래는 **MongoDB 컬렉션**(예: `inference_results` ) 요약 문서입니다.

## 목적

- 사용자가 업로드한 **구강/엑스레이 이미지**에 대한
  - 설문(survey)
  - 모델 추론 결과(객체 감지/임플란트 분류)
  - 생성된 오버레이 이미지 경로
  - AI 요약 결과/조언를 단일 문서로 기록.

## 스키마(관찰 기준)

필드	타입	설명
<code>_id</code>	ObjectId	기본 키
<code>user_id</code>	string	사용자 식별자(예: RDB의 <code>user.register_id</code> )
<code>image_type</code>	string	이미지 타입 ( "xray" 등)
<code>survey</code>	object<string, any>	설문 문항→응답 맵. 값이 문자열/숫자 혼재(예: <code>3</code> , "7회" )
<code>original_image_path</code>	string	원본 이미지 경로
<code>model1_image_path</code>	string	모델1 오버레이 경로
<code>model2_image_path</code>	string	모델2 오버레이 경로
<code>model1_inference_result</code>	object	객체 감지 결과
<code>used_model</code>	string	사용 모델 파일/버전

└ predictions[]	array	감지 목록
└ class_id	int	클래스 ID
└ class_name	string	클래스명(임플란트/보철물/근관 치료 등)
└ confidence	double	신뢰도(0~1)
└ bbox	number[4]	[x1,y1,x2,y2] (float)
└ summary	string	요약(없음/결과 설명)
implant_classification_result[]	array	임플란트 제조사 분류 결과
└ original_image	string	원본 경로
└ bbox	int[4]	[x1,y1,x2,y2] (여기선 int로 저장됨)
└ predicted_manufacturer_class	int	제조사 클래스 ID
└ predicted_manufacturer_name	string	제조사명
└ confidence	double	신뢰도
timestamp	date	업로드/처리 시각(UTC ISO)
AI_result	string	사람 읽기용 요약/조언(마크다운 포함 가능)

메모: bbox가 일부는 **float**, 일부는 **int**로 섞여 저장되어 있습니다.

## 대표 쿼리 예시

### 1. 사용자 최신 결과 1건

```

1 db.analysis_results.find(
2   { user_id: "121212" }
3 ).sort({ timestamp: -1 }).limit(1)
4

```

## 2. 특정 일자(KST)·타입별 건수

```

1 db.analysis_results.aggregate([
2   { $match: {
3     timestamp: {
4       $gte: ISODate("2025-08-26T00:00:00+09:00"),
5       $lt:  ISODate("2025-08-27T00:00:00+09:00")
6     }
7   }},
8   { $group: { _id: "$image_type", cnt: { $sum: 1 } } }
9 ])
10

```

## 3. 시간대 분포(KST 기준 0~23시)

```

1 db.analysis_results.aggregate([
2   { $match: {
3     timestamp: { $gte: startKST, $lt: endKST }
4   }},
5   { $project: {
6     hour: { $hour: { date: "$timestamp", timezone: "Asia/Seoul" } }
7   }},
8   { $group: { _id: "$hour", cnt: { $sum: 1 } } },
9   { $sort: { _id: 1 } }
10 ])
11

```

## 4. 감지 결과에서 특정 클래스만(예: 임플란트, **conf**≥0.7)

```

1 db.analysis_results.find({
2   "model1_inference_result.predictions": {
3     $elemMatch: { class_name: "임플란트", confidence: { $gte: 0.7 } }
4   }
5 })
6

```

## 5. 임플란트 제조사명 포함 검색

```

1 db.analysis_results.find({
2   implant_classification_result: {
3     $elemMatch: { predicted_manufacturer_name: /덴티움/i }
4   }
5 })
6

```

## 권장 인덱스(조회 패턴 기준)

```

1 db.analysis_results.createIndex({ user_id: 1, timestamp: -1 })
2 db.analysis_results.createIndex({ timestamp: -1 })
3 db.analysis_results.createIndex({ image_type: 1, timestamp: -1 })
4 // 필요 시 제조사명 검색이 잦다면

```

```
5 db.analysis_results.createIndex({ "implant_classification_result.predicted_manufacturer_name":  
6
```

## RDB 연계 메모

- 이 컬렉션의 `user_id` 는 MySQL의 `user.register_id` (문자열)와 매칭되는 패턴이 일반적입니다.  
필요하면 RDB의 `consult_request.id` (또는 `register_id`)를 별도 필드로 함께 저장해 조인 키를 명확히 하세요(예: `consult_request_id`).

## 데이터 품질/문서화 포인트

- **타입 일관성:** bbox는 한쪽이 float, 한쪽이 int → 가능하면 통일(예: 모두 float) 권장.
- **오버레이 경로:** `model1_image_path / model2_image_path` 대신  
`{ overlays: { xmodel1: <path>, xmodel2: <path> } }` 같은 맵 구조가 관리에 용이.
- **설문(survey):** 현재 한글 질문문이 key로 사용되고 값도 숫자/문자 혼재.  
문서에는 *질문 코드표*를 별도 부록으로 두고, 저장 시 `survey.{QID}` 형태로 기록하면 다국어/분석에 유리.
- **타임존:** 운영 기준(KST)로 일/시간대 집계 시 `timezone: "Asia/Seoul"` 을 명시.

필요하면 위 내용을 **운영 위키 템플릿**(개요 → 스키마 → 쿼리 → 인덱스 → 연계 → 비고) 형태로 내보내는 마크다운 파일도 만들어줄게요.