

잠시 후

# 숭실(SSU)데이터톤 2024

오리엔테이션이 시작됩니다.  
안쪽부터 차례로 착석부탁드립니다

사전설문  
참여하기



# 2024 SSU 데이터톤

효율적인 서가관리 정책제안

오리엔테이션  
2024. 12. 10 (화)  
전산관 다솜홀



# 환영사

황규백 지식정보처장 / 컴퓨터학부 교수

# CONTENTS

---

01

공모주제 및  
데이터 상세소개

02

Data OASIS  
서비스 활용

03

진행일정 및  
유의사항





# 01.

# 공모주제 및 데이터 개요

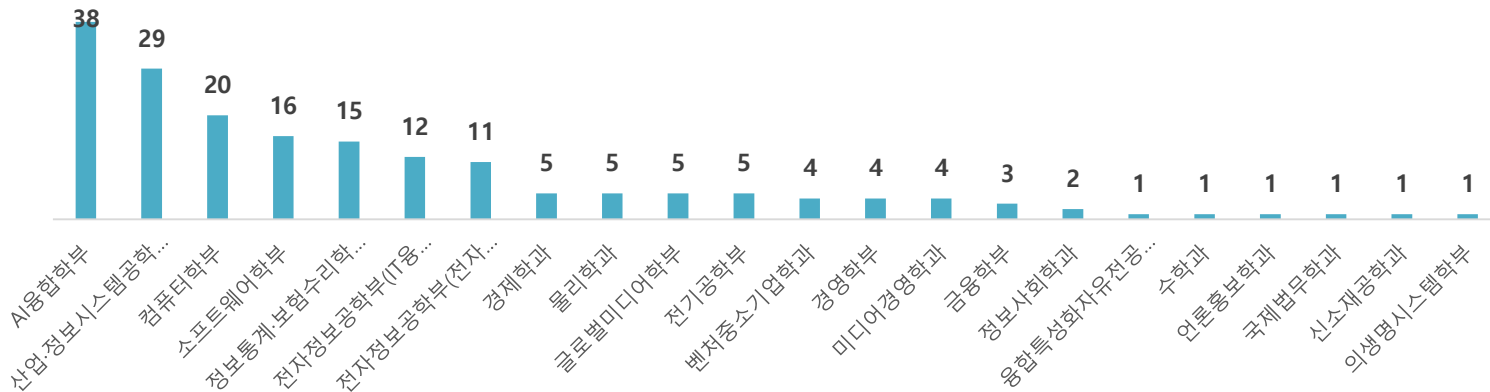
김덕성 팀장 / 중앙도서관 학술정보운영팀

## 1. 공모주제 개요 [1]

### ➤ 데이터로 바꾸는 도서관 - 효율적인 서가관리 정책제안 -

- 일정: 2024년 12월 10일(화) 12:00 ~ 2025년 02월 04일(화)
- 상금: 대상 1개팀: 150만원, 최우수상(2팀) 각 70만원, 우수상(3팀) 각 60만원
- 상금은 세금 공제 후 지급하며, 평가 결과에 따라 6팀 미만 시상 가능

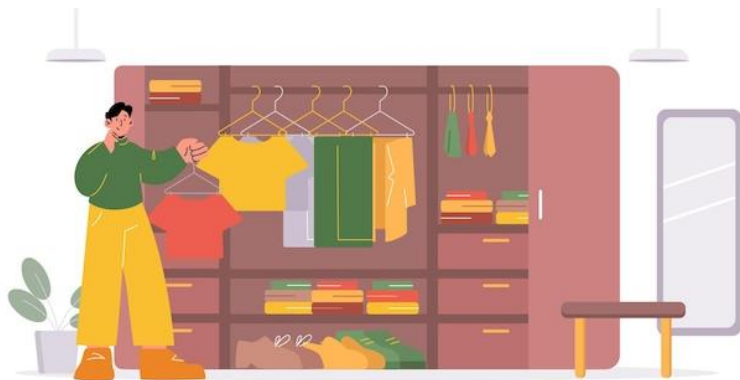
### ➤ 참가현황: 64개팀 184명



## 1. 공모주제 개요 [2]

### ➤ 분석목적

- 도서관은 한정된 공간을 효율적으로 활용하기 위해, 이용 빈도가 낮거나 오래된 책을 **'보존서고'**로 이동시킵니다.
- 보존서고는 단순 보관이 아니라, 자료를 언제든지 활용할 수 있도록 유지하는 공간입니다.
- **'자주 입는 옷을 걸어두고, 잘 쓰지 않는 계절 옷을 옷장에 넣어두는 것'** 처럼 생각하면 쉽습니다.





## 1. 공모주제 개요 [3]

### ➤ 분석목적1. 보존서고로 옮길 책을 선정하는 것

#### - 자료의 최신성 및 이용가능성을 고려한 보존서고 이동 대상 도서를 선정

- 도서관 서가가 꽉 차서 새로 들어오는 책을 놓을 공간이 부족한 상황을 생각해 보세요.
- 도서관이 보존서고로 옮길 책을 정하려면 **책이 얼마나 오래되었는지** 또는 **얼마나 자주 대출되었는지**를 확인해야 합니다
- 예를 들어, "5년 동안 한 번도 빌려간 적 없는 교재"는 **보존서고로 옮길 후보**가 될 수 있습니다.
- 이렇게 하면 서가에서 자주 읽히는 최신 책들이 더 쉽게 접근 가능하도록 자리를 만들 수 있어요.



## 1. 공모주제 개요 [4]

### ➤ 분석목적2. 미래에 필요한 책을 예측하는 것

#### - 서가 선반의 여유공간 확보를 위한 주제별 도서 입수 예측

- 시험 준비를 위해 특정 교재를 빌리러 갔는데, 항상 대출 중이라 찾기 힘들었던 경험이 있을 거예요.
- 데이터를 분석하면 어느 과목이 인기가 많고, 학생들이 어떤 책을 자주 찾는지를 알 수 있습니다.
- 예를 들어, "새로운 전공 과목이 추가될 때 해당 분야의 책이 더 필요할 것" 또는 "다음 학기에는 문학 책이 더 많이 대출될 것"이라고 예측할 수 있습니다.
- 이런 데이터를 활용하면 도서관은 미리 책을 준비해 서가에 배치하고, 학생들이 원할 때 책을 빌릴 수 있게 도울 수 있어요.

## 2. 중앙도서관 데이터 개요

### ➤ 단행본(도서)정보

- 개요: 단행본(이하 도서)의 정보를 표시하는 정보로 도서 각 권당 부여된 정보
- 데이터 총 건수 및 등록일 범위: 301,151권 / "1968-01-11" ~~ "2024-11-13"
- 데이터형식 및 파일명: 콤마(,)로 구분된 텍스트 정보 / 단행본(도서)정보.txt (35,803KB)

### ➤ 대출정보

- 개요: 대출한 도서 정보와 언제 대출했는지 일시를 알 수 있는 정보
- 데이터 총 건수 및 대출일시 범위: 916,709 대출 건 / "2004-11-01" ~~ "2024-10-31"(20년간)
- 데이터형식 및 파일명: 콤마(,)로 구분된 텍스트 정보 / 대출정보.txt (29,453KB)

구분	데이터 유형	설명	형식 및 예
도서ID	문자열(9)	도서 1권당 붙이는 고유번호	형식: SS_ + 숫자(6) / 예: SS_233793
등록일자	문자열(10)	도서를 중앙도서관에 입수처리한 날짜	형식: YYYY-MM-DD / 예: 2008-10-13
수서방법	문자열(6)	도서를 입수할 때 어떤 사유로 입수했는지 확인하는 정보	*수서정보없음: 입수 방법 정보를 확인할 수 없는 자료 (주로 과거자료) *이용자희망: 희망도서 신청으로 입한 한 자료 주로 학생의 구매 요청 *사서선정: 최신 자료 및 타대학 구매도서를 참고해서 사서가 직접 구매한 자료 *학과신청: 학부(부)의 교원이 직접 또는 학과를 통해 구매 요청한 자료 *수업지정: 수업에 사용할 목적으로 구매한 자료로 최소(1-2)권 대출불가처리함 * 기타: 행정처리 등 기타 목적으로 구매
분류코드	문자열(16)	10진분류법으로 도서의 정보를 분류하기위한 정보로 본교는 DDC를 기반으로 본교에 맞게 일부 수정하여 사용함. 사서가 도서의 정보를 확인한후 분류 코드 부여하며 해당 코드와 저자 정보를 오름차순으로 정렬하여 서가에 배가함.	형식: 숫자(3) + "." + 숫자(12)이내 예: 160 , 796.357 150.195092 016.330917671 016.72094920904 (값 없음)

## 3. 단행본(도서) 정보 [2]

구분	데이터 유형	설명	형식 및 예
ISBN	문자열(13)	International Standard Book Number의 약자 각 출판사가 출판한 각각의 도서에 국제적 표 준화하여 붙이는 고유 도서번호	형식: 숫자(13) 2007년 발생 도서는 대부분 13자리 숫자형식 을 가지만 이전 정보는 10자리 또는 잘못된 정보가 들어가 있을 수 있다
서명	문자열(250)	도서의 제목	형식: 한글, 외국어, 한자 표시 예) 가을, 秋
저자	문자열(150)	도서를 저작한 사람 또는 기관 정보	예) 길선주 Spillmann, Hans Otto (값 없음)
출판사	문자열(25)	도서를 출판한 출판사의 정보	예) 홍성사 / 강릉시 (값 없음)
출판년도	문자열(25)	도서의 출판년도를 표시하는 정보로 주로 도 서의 책등과 표지에 있는 정보참고	형식: 문자 또는 숫자형태로 구성된 정보 예: 2023, 2022-2023
소장위치	문자열(4)	현재 도서가 어느 위치에 있는지 알려주는 위 치 정보	4층인문: 현재 도서관 4층에 정리된 도서 (직접 도서를 확인하여 대출 가능) 보존서고: 지하 보존서고에 정리된 도서 (보존서고 신청을 통해 대출 가능) 2222: 위치 확인 불가도서

## 4. 대출 정보

구분	데이터 유형	설명	형식 및 예
도서ID	문자열(9)	도서 1권당 붙이는 고유번호	형식: SS_ + 숫자(6) / 예: SS_233793
대출일시	문자열(19)	도서를 대출한 일시 정보	형식: YYYY-MM-DD HH24:MI:SS 예: 2004-11-15 12:33:01

※ 중앙도서관 데이터는 금일 오리엔테이션 종료 후 팀장 메일로 발송예정

※ 본 데이터는 숭실(SSU)데이터톤 2024 공모전 목적으로만 제공됩니다.  
이외의 목적으로 활용할 수 없으며, 공모전 종료 즉시 삭제(파기)해야 합니다.



02.

Data Oasis 활용

## 1. 도서관 홈페이지 → 연구학습지원 → 데이터 활용안내 → Data Oasis

### DATA Oasis

데이터 오아시스(DATA Oasis)는 송실인을 위한 DATA와 IT 관련 자원을 소개하는 공간입니다.  
 각종 국가통계 자료에서부터 딥러닝을 위한 데이터까지 연구·학습에 필요한 다양한 정보원과 데이터 활용법을 제공합니다.

**각 제공처 별 이용정책 준수!**

34 건

번호	바로가기	전자정보원	
1		문화 공공데이터관장 문화체육관광부 소속기관 및 타 부처 기관에서 생산·제공하는 모든 문화정보를 한 곳에 모아 서비스하는 문화정보 개방플랫폼입니다.	<a href="#">바로가기</a> <a href="#">상세정보</a>
2		국회 회의록 빅데이터 국회 회의록을 발언자나 키워드 등을 이용하여 검색하고 다운로드할 수 있으며, 인포그래픽을 제공하는 사이트입니다.	<a href="#">바로가기</a> <a href="#">상세정보</a>
3		DataON(국가연구데이터플랫폼) 연구데이터를 체계적으로 공유·관리하고 검색·분석·활용을 지원하는 국가연구데이터플랫폼입니다.	<a href="#">바로가기</a> <a href="#">상세정보</a>
4		Papers With Code 딥러닝 분야의 State-of-the-Art(최신기술)과 Research Trend 등을 논문, 데이터, 코드와 함께 제공하는 사이트입니다.	<a href="#">바로가기</a> <a href="#">상세정보</a>
5		KOSIS 국가통계포털 국내·국제·북한의 주요 통계를 한 곳에 모아 이용자가 원하는 통계를 한 번에 찾을 수 있도록 통계청이 제공하는 One-Stop 통계 서비스입니다.	<a href="#">바로가기</a> <a href="#">상세정보</a>
6		공공데이터포털 공공기관이 생성 또는 취득하여 관리하고 있는 공공데이터를 한 곳에서 제공하는 통합 창구입니다.	<a href="#">바로가기</a> <a href="#">상세정보</a>
7		AI Hub AI 기술 및 제품·서비스 개발에 필요한 AI 인프라를 지원함으로써 누구나 활용하고 참여하는 AI 통합 플랫폼입니다.	<a href="#">바로가기</a> <a href="#">상세정보</a>





03.

# 진행일정 및 유의사항

# 진행일정 및 유의사항

03



## 1. 활동 및 주차별 보고서 제출

➤ 활동기간: 2024.12.10.(화) ~ 2025 02.04.(화)

※ 활동기간 중 **주차별 활동보고 제출 필수 / 매 주 화요일 14시까지 / (메일수신시간 기준)**

※ 제출주소: [dataton@ssu.ac.kr](mailto:dataton@ssu.ac.kr) / 기한 준수여부, 내용 충실성 등은 우수작 평가에 반영함.

활동주차	제출일자
1주차 : 12/10(화)~12/16(월)	<b>12/17(화)</b>
2주차 : 12/17(화)~12/23(월)	<b>12/24(화)</b>
3주차 : 12/24(화)~12/30(월)	<b>12/31(화)</b>
4주차 : 12/31(화)~01/06(월)	<b>01/07(화)</b>

활동주차	제출일자
5주차 : 01/07(화)~01/13(월)	<b>01/14(화)</b>
6주차 : 01/14(화)~01/20(월)	<b>01/21(화)</b>
<b>7주차 : 01/21(화)~01/27(월)</b>	<b>제출없음</b>
최종결과물 제출일	<b>02/04(화)</b>

\* 주차별 활동보고 양식은 Fun시스템 게시판 참고

➤ 팀장 대상 대화채널(카카오톡 오픈 채팅방) 운영 / **12/10 OT 후 팀장 메일로 링크발송**

※ 각 활동팀의 팀장만 입장 / 입장 시 **"성명(팀명)"으로 입장** / ex) 김승실(상금내꺼)

## 2. 최종 결과물 제출 (1)

- 결과보고서 내용이 일정 수준에 미달할 경우 시상이 안 될 수 있음
  - ※ 대상(1팀), 최우수상(2팀), 우수상(3팀) 총 우수작 6팀 선발 예정
  - ※ 결과보고서 평가 시 일정 수준 미달의 경우 6팀 미만 시상 가능
- 결과 보고서 양식 (PPT 양식으로 제출 / 30슬라이드 내외)
  - ※ 포함내용: 정책 아이디어 배경 및 개요 / 데이터 분석 (수집,전처리,EDA) / 기대 효과
- 결과 프로그램 및 소스(해당자에 한함) 제출
  - ※ 프로그램 실행 실시 및 구동 환경 설명 포함. / 프로그램 실행파일 및 소스
  - ※ 가능한 데이터 반입 후 프로그램 안에서 결과 제공 가능하도록 구성



## 2. 최종 결과물 제출 [2]

- 결과물에 사용된 폰트 이미지는 반드시 **무료 만 가능**.
  - ※ 무단 사용으로 인한 **저작권 문제시 제출자에게 법적 책임** 있음
- 결과물이 국내 ·외 에 이미 공개된 작품일 경우 입상 후라도 입상이 취소됨
  - ※ 단, 비슷한 내용이라도 연구 및 접근 방법이 다르다고 인정되는 경우는 제외
  - ※ **저작권 관련 문제 발생시 입상 금액 환수 조치**되며, 이에 대한 **법적 책임은 제출자**에 있음
- 결과물에 대한 저작권은 출품자에게 있으며, 사용권 및 관련 저술 등 판권은 주최측이 사용을 득함
  - ※ 행사 진행시 촬영된 사진, 동영상 등의 저작권/초상권 및 사용권은 주최측에 있음
  - ※ 제출된 결과물은 관련자료 소개, 전시, 책자 및 교육자료 등 주최측에서 활용 할 수 있음

## 2. 최종 결과물 제출 (3)

- 제출기한: 2025.02.04.(화) **14:00까지 (메일수신시간 기준)**
  - ※ 제출주소: [dataton@ssu.ac.kr](mailto:dataton@ssu.ac.kr)
  - ※ 최종결과물 미제출 팀은 중도포기로 간주하여 평가대상에서 제외,  
Fun시스템 미수료 처리 및 지식정보처장명의 수료증 미발급

## 3. 우수팀 선발

- 공지일정: 2025.02.07.(금) / 기한 내 제출완료 팀 중 **8개팀 선발**
- 심사 근거자료: 제출된 최종 산출물 및 주차별 활동보고서
  - ※ 심사위원의 평점이 가장 높은 팀 선정
- 심사기준
  - ※ 분석력(30점), 타당성(30점), 독창성(20점), 성실/협동성(20점)

## 4. PT심사 및 시상 (1)

- **심사일정: 2025.02.13.(목) 14:00~ 중앙도서관 5층 A/V 교육실(변동가능)**
  - ※ 최종 발표 심사에 불참시 기권 처리함. (수상포기로 간주)
  - ※ 결과보고서 평가 시 일정 수준 미달의 경우 6팀 미만 시상 가능
  - ※ 발표 순서는 당일 즉석 추첨예정
- **발표력 / 설명력 / 질의응답 등 종합적으로 평가**
  - ※ 심사위원의 평점이 가장 높은 팀 선정
- **심사기준**
  - ※ 설득력(40점), 발표·태도(30점), 활용 가능성(30점)

## 4. PT심사 및 시상 [2]

구분	내용	진행자
14:00 ~ 14:10	소개	황규백 지식정보처장
14:10 ~ 14:30	1팀 발표 및 질의응답	김덕성 학술정보운영팀장
14:30 ~ 14:50	2팀 발표 및 질의응답	
14:50 ~ 15:10	3팀 발표 및 질의응답	
15:10 ~ 15:30	4팀 발표 및 질의응답	
15:10 ~ 15:30	휴식	다과 제공
15:30 ~ 15:50	5팀 발표 및 질의응답	김덕성 학술정보운영팀장
15:50 ~ 16:10	6팀 발표 및 질의응답	
16:10 ~ 16:30	7팀 발표 및 질의응답	
16:30 ~ 16:50	8팀 발표 및 질의응답	
16:50 ~ 17:10	휴식	심사 및 결과 취합
17:10 ~	수상자 발표 및 사진촬영	황규백 지식정보처장





---

# THANK YOU

---

A blue-tinted background image of a park with trees and a fountain, overlaid with a semi-transparent blue rectangle. The text "THANK YOU" is centered in white, bold, sans-serif font, flanked by two horizontal white lines.



# Q&A 및 휴식