Evaluating Interdisciplinary Contributions to Synchrotron Radiation Studies: Scientometric Insights into Novelty and Disruption

Dong Lu[1,2], Wang Zhiqiang[2,3], Tao Zhiyu[2], Li Yizhan[1,2], Zhang Hong[2], Li Zexia[1,2]*
[1]National Science Library of Chinese Academy of Sciences
[2] Department of Information Resources Management, School of Economics and Management, University of the Chinese Academy of Sciences
[3] UNU-MERIT, Maastricht University

Abstract : Research infrastructures consisting of multiple experimental stations and functioning as large-scale scientific facilities are critical to advancing scientific understanding and addressing complex, interdisciplinary challenges. This study explores the relationship between the interdisciplinarity of these stations and the novelty and disruption of their academic outputs, using the US. Department of Energy's Advanced Photon Source (APS) as a case study. We analyzed 29,692 papers produced by 53 experimental stations of the APS from 1992 to 2020, each of which received more than five citations by 2023. The novelty of academic outputs was assessed using SciBERT to measure semantic similarities between titles within the citation network, while disruption was measured with a relative dependency index. Interdisciplinarity was quantified by variety, balance, disparity and interdisciplinary integration index. The results of the ordinary least squares (OLS) regression show a negative correlation between novelty and disciplinary disparity and a significant negative correlation between disruption and interdisciplinarity. These results suggest that interdisciplinary approaches promote the exchange of ideas and holistic understanding between disciplines, but that specialized expertise remains critical for groundbreaking research. This balance encourages further investigation to optimize the interplay between interdisciplinary collaboration and disciplinary specialization to drive scientific progress.

Keywords : Experimental Stations; Interdisciplinarity; Novelty; SciBERT; Disruption

Received: 20.6.2024 Accepted: 9.12.2024 ISSN 2241-1925

Introduction

National research infrastructures are a crucial component of a country's scientific and technological resources. These facilities significantly a nation's basic scientific research and technological innovation capacity, as highlighted in seminal studies (Bu et al., 2021; OECD, 2020). Such infrastructures encompass multiple experimental stations equipped with a range of large-scale scientific instruments, including synchrotron radiation light sources, spallation neutron sources and X-ray free-electron lasers. These instruments are essential for conducting high-impact research in various scientific disciplines. The benefits of these facilities for scientific progress are well documented. Researchers have achieved a variety of groundbreaking and innovative breakthroughs that underscore the central role of these infrastructures in the national framework for scientific and technological innovation (Fecher et al., 2021). Notable facilities such as the Advanced Photon Source (APS) and the European Synchrotron Radiation Light Source (ESRF) have been instrumental in research that has led to the award of three Nobel Prizes (Helliwell & Mitchell, 2015). For instance, pivotal studies such as "Ribosome research", which was awarded the Nobel Prize in Chemistry in 2009, and "Research on G protein-coupled receptors"," which was honored with the Nobel Prize in Chemistry in 2012, both relied heavily on these synchrotron radiation sources.

Given the predominant focus on basic research in these infrastructures, academic publications are the most important scientific output and a vital medium for the dissemination of knowledge (Beck et al., 2019). Therefore, an accurate assessment of the innovative quality of these academic papers is essential to evaluate the scientific impact of such infrastructures (Fabre et al., 2021). Initially, the experimental stations of national scientific user facilities may not have specific scientific objectives; their primary goal is to provide researchers with the most comprehensive experimental tools and methods. Once they are operational, the range of scientific research directions supported by these stations expands.

This study delves into the relationship between the interdisciplinarity of experimental stations and the novelty and disruption of academic outputs. Specifically, it examines how the interdisciplinarity of experimental stations affects the novelty and disruption of innovation in academic outputs, using the 53 beamline experimental stations at Argonne National Laboratory's Advanced Photon Source (APS) as a case study. By focusing on this aspect, the study seeks to illuminate the correlation between the general utility of experimental stations and the specific innovative contributions they enable, thereby enhancing our understanding of the strategic role these facilities play in advancing scientific knowledge.

Related Literature

Evaluation of Interdisciplinary Applications of Scientific Instruments

The experimental stations of the research infrastructure combine a wide range of advanced scientific instruments. These stations are not only a catalyst for bringing together different areas of research, but also serve as indispensable resources that provide both specialized and general-purpose tools and methods that are critical to a wide range of scientific explorations. The spectrum of applications for these instruments is broad, ranging from niche technological applications to broader, more diverse research needs. The generality of scientific instruments has been studied mainly through qualitative analyzes (Adeyemi, 1996; Hentschel, 2015; Rosenberg, 1992; Staley, 2002). In parallel, there is a body of research that uses quantitative methods and scientometric analyzes to complement these narratives (Soderstrom et al., 2022; Zhang et al., 2019).

In particular, Zhang et al. (2019) used the indicator of disciplinary balance together with the citation frequency of academic papers as a metric to assess the interdisciplinarity and scientific impact of large-scale facilities, respectively. Their results underlined a positive correlation between interdisciplinarity and the scientific impact of such platform-type facilities. However, their study did not go into a detailed examination of interdisciplinarity within individual experimental stations. On a related front, Soderstrom et al. (2022) proposed that scientific instruments with research applications that cover a variety of disciplines and are not limited to specific fields should be classified as general-purpose instruments. In addition, they made a quantitative assessment of the generality of the European Synchrotron Radiation Source (ESRF) experimental stations using the Herfindahl-Hirschman Index (HHI) as a metric, but without examining the relationship between instrument generality and scientific impact. This study attempts to apply interdisciplinarity as a metric to assess the generality of research infrastructure experimental stations through a scientometric framework. Furthermore, it seeks to uncover the correlation between the interdisciplinarity of these stations and the innovative impact of their academic outputs.

Interdisciplinarity, as conceptualized through the lens of integrating knowledge across multiple disciplines to make progress in understanding or solving complex problems that extend beyond the purview of a single discipline (Sciences et al., 2005), is typically assessed along three dimensions: variety, balance and disparity (Mutz, 2022). Variety addresses the multiplicity within a system, quantifying the array of distinct element types it comprises. Balance scrutinizes the distribution pattern of the elements in the different categories, focusing on the evenness of distribution across the different types of elements. Disparity, meanwhile, assesses the extent to which items differ from each other and explore the extent of variation between item type (Zhou et al., 2022).

When assessing interdisciplinarity, the indicators are divided into three categories (Leydesdorff & Rafols, 2011; Zhang & Leydesdorff, 2021): (1) Single-dimensional indicators, covering measures of disciplinary variety (such as total categories and disciplinary specialization), balance (such as the Simpson index and the Gini coefficient), and disparity (illustrated by the Jaccard similarity index and cosine similarity). (2) Two-dimensional indicators, such as the Brillouin index, the Shannon entropy and the Herfindahl-Hirschman index, which simultaneously assess disciplinary variety and balance. Conversely, the Concentric Index and the Scholastic Divisional Index assess both disciplinary variety and disparity. (3) Integrated indicators, exemplified by the Disciplinary Integration Index (Porter et al., 2007), the Rao-Stirling Index (Stirling, 2007), the 2Ds Index (Porter & Rafols, 2009), the DIV Index (Leydesdorff et al., 2019), which provide a holistic assessment of interdisciplinarity. Each indicator, which differs in its conceptual basis, focus and methodological approach, serves a unique assessment purpose and highlights the diverse lens through which interdisciplinarity can be explored and quantified.

Quantifying the Novelty of Academic Papers

The assessment of the novelty of academic papers focuses on the introduction of new scientific knowledge, which historically has been assessed through peer review. With the increasing volume of publications, quantitative methods have emerged to identify novel insights, moving beyond traditional peer review processes. Arthur (2010) and Uzzi et al. (2013) highlighted innovation as a novel recombination of existing knowledge and proposed quantitative frameworks for assessing novelty through various methods, including subject classification codes, MESH codes, International Patent Classification codes, and keywords, as detailed by various researchers (Boudreau et al., 2016; Shibayama et al., 2021; Sreenivasan, 2013; Verhoeven et al., 2016; Yan et al., 2020). However, the mere combination of knowledge is no guarantee of innovation, suggesting that more sophisticated tools are needed to accurately measure novelty.

Jeon et al. (2022) identified novel papers as those deviating from mainstream research, utilizing outlier detection algorithms such as the Local Outlier Factor (LOF) to measure novelty, with further developments incorporating both outlier characteristics and citation patterns for a comprehensive assessment of novelty (Zhou et al., 2021). Innovation is also evident in the emergence of new concepts to describe phenomena, with methods focusing on the first occurrence of words or sequences to measure novelty (Arts et al., 2021; Balsmeieri et al., 2018). However, not all new words signify substantial innovation, and focusing only on content outliers may overlook the quality and relevance of research.

The application of new methods to address new problems has been emphasized, with Heffernan and Teufel (2018) introducing a "problem-solution" framework, and Luo et al. (2022) using supervised learning to assess novelty through semantic distances. This approach has been extended by machine learning and deep learning models for word classification and extraction (Kondo et al., 2011; Lu et al., 2021; Zhang et al., 2024) presenting a novel era of LLM-based term analysis. However, these methods face challenges such as the need for manual annotation and the adaptability of the model in different research fields, implying significant time and resource investments.

Citation network analysis has been shown to be effective in measuring novelty (Amplayo et al., 2018; Varga, 2018), although its effectiveness is limited by citation delays. Jeon et al. (2023) used fastText to construct a vector space model in which papers with similar scientific knowledge are close to each other, and used the local outlier factor (LOF) to measure the novelty of the scientific knowledge contained in the papers on a numerical scale. Shibayama et al. (2021) used the word vector embedding method to calculate the title similarity of the publication reference pairs to measure the novelty. Indicator-based evaluations using regression or machine learning models to analyze publication metrics offer broad applicability (Matsumoto et al., 2021; Verhoeven et al., 2016; Zhou et al., 2021).

Quantifying the Disruptiveness of Academic Papers

The concept of disruption traces its roots to economic development theory pioneered by Schumpeter, who posited that the restructuring of production factors can increase productivity and thereby displace established modes of production (Schumpeter, 1961). Extending Schumpeter's theoretical framework, Christensen (2015) posited that disruption arises from the penetration of novel technologies into established value networks, a process characterized primarily by the phenomenon of technological substitution.

Building on this conceptual foundation, Funk and Owen-Smith (2017) pioneered the first quantitative approach to measure the extent of technological substitution. This methodology used a novel analysis of the links between focal patents, cited patents and cited patents to assess the degree of disruptiveness. Based on this framework, Wu et al. (2019) further refined the measurement of disruptiveness specifically in the area of scientific publications and introduced the D-index. Subsequently, numerous researchers have attempted to optimize the D-index from different perspectives to improve its applicability (Bornmann et al., 2020; Bu et al., 2021; Park et al., 2023; Yang et al., 2023). Bornmann et al. (2020) proposed the DI5 index to assess the strength of links between the citing and cited aspects of focal papers (FPs). This index identifies a developmental FP by requiring that the citing papers reference not only one, but at least five of the FP's cited papers. Bu et al. (2021) introduced the dependency index (DEP), defined as the average number of citation links from a paper citing the FP to the references cited by the FP. Bittmann et al. (2022) used disruption indicators (DI1, DI5, DI1n, DI5n, and DEP) to assess the innavative disruption of milestone and non-milestone papers from Physical Review E, finding that

DI5 and DEP performed well in assessing the disruptiveness of published papers.

The popularity of the D-index is evident not only in its application in various disciplines, but also in the recent introduction of a journal-level index variant Jiang and Liu (2023) introduced the Journal Disruption Index (JDI) and presented it as a novel alternative to the conventional journal-level metrics, especially the Journal Impact Factor (JIF), as provided by Clarivate. In addition to metrics focusing on substitution, the academic community has also explored alternative methods to assess the disruptiveness of scientific outputs (Boudreau et al., 2016; Foster et al., 2015; Wang et al., 2023). Zhou et al. (2021) proposed a disruption-oriented adaptation of the PageRank algorithm tailored to assess the disruptiveness of publications by analyzing citation networks. Westerski and Kanagasabai (2019) emphasized the importance of unconventional applications as a source of disruption and advocated the integration of outlier detection algorithms into the evaluation process. Wang et al. (2023) proposed the Entity-based Disruption Index (EDI) to integrate network-based and text-based analytical approaches. Complementary to these approaches, some researchers conceptualize disruption as an emergent property characterized by its discontinuous nature. Zhou et al. (2019) and Zhang et al. (2021) have used a topic landscape analysis to identify emergent topics as indicators of disruptive potential.

To summarize, the current research landscape is primarily focused on assessing the disruptive capacity of scientific contributions through their potential to displace existing paradigms. This has led to the development and refinement of a variety of methodologies, all aimed at measuring the degree of substitution from different angles, thus enriching our understanding of scientific disruptiveness.

Data and Method

Data Processing

In this study, we used the U.S. Advanced Photon Source Research Output Repository (APS) as the cornerstone data source, drawing from a rich pool of scientific outputs from 1985 to 2023. The APS facility, known for its important contributions to the scientific community, was instrumental in creating an extensive corpus of 42,411 scientific publications. This corpus covers a wide range of scientific communication, including but not limited to journal articles, monographs, proceedings of scientific symposia, technical exhibitions, patent documents, and other various publication modalities (data collection date: January 8, 2024). Based on the methodological framework proposed by Bornmann and Tekles (2019), which stipulates the need for a minimum period of three years for citation as well as a minimum number of five citations to assess the innovative nature of an article, we meticulously focused our analysis on the scientific publications attributed to the APS Light Source between 1985 and 2020. This subset amounted to 36,212 publications, of which a significant proportion (34,274) were journal articles, each clearly identified and associated with specific experimental stations within the APS facility.

After checking the DOI numbers of these journal articles labeled with the experimental stations, we performed a search in the Web of Science (WoS) database. The aim of this search was to extract the metadata of articles with at least five citations and classify them into 53 categories based on the names of the experimental stations. A total of 1,051,432 references were cited in 25,669 articles, with a total number of 1,445,663 cited articles. In cases where an article uses multiple experimental stations for research, this is recorded as a separate entry for each station involved. The methodology of data collection and processing is described in detail in Table 1.

Table1. Steps of Data Collection and Processing of Academic Papers of the APS Experimental

Step	Processing Content	Data Sets	Number of Records
1	Search for journal articles with the Experiment Station label	The raw data set	35, 443
2	Publication year ≤ 2020	The raw data set	34, 274
3	Frequency of citations ≥ 5	Papers to be evaluated	25, 669
4	Search for the cited reference of the paper to be evaluated	The reference data set	1, 051, 432
5	Search for the publications cited the paper to be evaluated	The citation data set	1, 445, 663

Method

Dependent Variable

(1) Innovative Novelty Index of Academic Papers

In this study, the novelty of academic papers was evaluated based on the method proposed by Shibayama et al. (2021). We utilized the Digital Object Identifiers (DOIs) of the references cited within the papers under review to retrieve their titles from the Web of Science (WoS) or Crossref database. We also used SciBERT as a library for word embedding. The calculation of the novelty of a document is performed in the following consecutive steps.

Step 1: First, we vectorize the title of the *i*-th reference as v_i calculated as the mean of word embedding of all words included.

Step 2: Second, we compute the distance of each pair of cited documents. The cosine distance between the *i*-th and *j*-th reference ($1 \leq i < j \leq n$, *n* is the number of the cited references) is calculated using the following formula:

(1) $d_{ij} = 1 - \frac{v_i \cdot v_j}{|v_i||v_j|}$

The value of d_{ij} ranges from 0 to 2, where a larger value indicates a larger distance.

Step 3: Finally, we aggregate the distance values across all pairs of cited references. As a measure of the novelty of a focal document, we take the 100-percentile value of the distance values. The calculation formula is as follows:

(2) $NOVEL = R^{-1}$

where $R(d_{ij})$ is the ordinal rank of d_{ij} of all the distances of $n(n-1)/2$ reference pairs.

(2) Innovative Disruption Index of Academic Papers

The relative independence index (PCR) proposed by Bu et al. (2021) is used to assess the innovative disruption of academic papers by measuring how strongly the citation impact of the paper being assessed depends on the citation impact of its references. This indicator relies on the intricate network of scientific citations to measure the disruptive potential of academic publications. In the visual representation of this citation network, typically shown in Figure 3, the article under scrutiny, referred to as a focal paper (FP), is marked with a solid rhombus symbol. Conversely, the references cited by the FP are represented by hollow rhombus, highlighting their role in the foundational underpinning of the FP's scholarship.

To further delimit the citation network, there are three clear symbols that identify different categories of citing publications: (1) solid circles (N_i) represent publications that reference both the FP and one or more of its cited references, indicating a direct engagement with the FP's intellectual framework; (2) hollow circles (N_j) symbolize publications that cite only the FP, suggesting a direct influence or critique of the FP's contributions; and (3) hollow triangles (N_k) denote publications that engage with the references of the FP but do not cite the FP itself, reflecting an indirect contextual relationship. When calculating the relative independence index PCR, N_k is not taken into account and is calculated according to the following formula:

(3) $PCR = \frac{N_i}{N_i + N_j}$

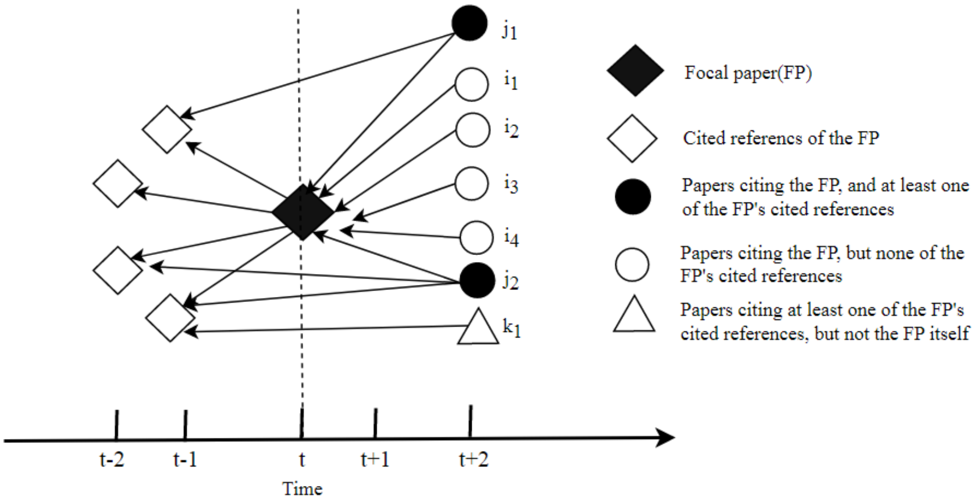

Figure 1. Schematic Diagram of the Relative Independence of a Publication

Using the given formula, the relative independence index PCR of the data set comprising the focal papers was calculated using Python. As shown in Figure 1, the disruption value of the FP was two-third. The index varies between 0 and 1. A PCR value of 1 for the FP means that none of the publications citing the FP cite any publications that the FP itself cites. This shows that the FP has a distinctly independent impact within the citation network. Conversely, a PCR value of 0 for an FP signifies that all publications citing the FP also cite at least one publication that the FP has cited, reflecting a lower independent citation impact for the FP.

Independent Variables

In this study, the experimental station is considered as a scientific research unit whose interdisciplinarity is assessed by evaluating the interdisciplinary scope of its academic outputs. Specifically, we used four independent variables to assess the interdisciplinarity of APS (Advanced Photon Source) experimental stations: the total number of disciplinary categories (N_S) according to Porter and Chubin (1985), disciplinary dispersion via the 1-Gini coefficient (Leydesdorff & Rafols, 2011), disciplinary distance through the 1-cosine similarity (Hamers et al., 1989) and the Rao-Stirling index. These metrics enable the evaluation of the academic output from experimental stations concerning variety (VAR), balance (BAL), and disparity (DIS), either singularly or in an integrated manner. The Web of Science Subject Classification System, which includes 254 categories, was used to categorize the journals in this study. Using the open-source tools developed by Leydesdorff (jcitnetw.exe and mode2div.exe), the interdisciplinary metrics for scientific publications across different trial sites were calculated in the study using Pajek software.

Variety (VAR): Measured by the total count of different subject categories (N_S) covered by a publication, reflecting the disciplinary breadth it encompasses. The formula for VAR is as follows:

(1) $N_S = \sum_{i \in S} S_i$

where i is the number of papers within the discipline *S*, and S_i is the total number of discipline categories to which the paper collection i belongs.

Balance (BAL): Determined by the complement of the Gini coefficient across discipline categories, offering a view into the even distribution of the spread of a publication across different disciplines. The formula for BAL is as follows:

(2) $1 - Gini = 1 - \frac{\sum_{i=1}^{n}(2i-n-1)x_i}{n\sum_{i=1}^{n} x_i}$

where n denotes the count of unique disciplines, and x_i, arranged in ascending order, signifies the quantity of publications within the i-th disciplinary category.

Disparity (DIS): Quantified by 1-cosine similarity. The DIS measures the disciplinary difference between the subject categories of the papers. The formula for DIS is as follows:

(3) $1 - Cos(x,y) = 1 - \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{(\sum_{i=1}^{n} x_i^2)(\sum_{i=1}^{n} y_i^2)}}$

where n represents the number of disciplines involved, x_i denotes the count of references associated with discipline *i*, y_i indicates the count of records belonging to discipline *i*. The DIS value lies between 0 and 1, with values approaching 1 indicating higher disciplinary congruity between the pairs of publications being compared.

Rao-Stirling (RS) Index: Measures interdisciplinarity, distinguishing variety, balance, and disparity as the three components of interdisciplinarity. The formula for RS is as follows:

(4) $RS = \sum_{i \neq j} d_{ij} p_i p_j$

where d_{ij} denotes the distance between subject *i* and subject *j*. $p_i = x_i/X$, $X = \sum x_i$, and x_i denotes the number of cells belongs to subject *i*.

Control Variables

In this study, we identified and integrated seven variables as potential confounding factors that may influence the value of the innovative novelty and disruption index.

(1) Number of references (*NR*): This metric serves as an indicator of the integrative capacity of a paper, with a higher number of citations reflecting a comprehensive engagement with existing knowledge (Jones et al., 2008).

(2) Length of title *(LTI)*: Longer titles may contain more descriptive vocabulary that more accurately describes the content and scope of the research, reflecting greater novelty or depth (Foster, 2018).
(3) Number of authors (*NAU*): Collaboration among multiple authors can lead to a synthesis of different expertise, which could enhance the paper's impact, although this effect may vary depending on the discipline (Sugimoto et al., 2011).
(4) Number of institutions (*NAD*): The extent of cross-institutional collaboration is posited to enrich research quality and its consequent impact, through the amalgamation of disparate research cultures and resources (Jones et al., 2008).
(5) Proportion of research papers (*NRP*): This takes into account the expected difference in scientific impact between research articles and review articles, with the former presumed to make a greater contribution to scientific progress.
(6) Number of years of activity of the experimental station (*NY*): It is assumed that the length of time an experiment station has been in operation is an indicator of cumulative scientific output and overall impact in the academic field.
Table 2 presents abbreviations, descriptions and types for each variable considered in this study.

Table 2. Summary of Variables

Categorization	**Abbreviation**	**Description**	**Type**
Dependent Variables	*NOVEL*	Novelty metric in the focal paper, as shown in Equation (1-2)	Continuous
	PCR	Relative independence index in the focal paper, as shown in Equation (3)	Continuous
Independent Variables	*VAR*	The total number of disciplinary categories the focal paper spans, as shown in Equation (4)	Count
	BAL	The evenness of the distribution of disciplines of the focal paper, as shown in Equation (5).	Continuous
	DIS	The distance among the disciplines of the focal paper, as shown in Equation (6).	Continuous
	RS	The index combining variety, balance, and disparity to measure the paper's interdisciplinarity, as shown in Equation (7).	Continuous
Control Variables	*NR*	Number of references in the focal paper	Count
	LTI	Length of the title of the focal paper	
	NAU	Number of authors contributing to the focal paper	Count
	NAD	Number of institutions associated with the	Count

		authors of the paper	
	NRP	The proportion of papers classified as research-based	Count
	NY	The operating time of the experimental station	Count

Results Discussion

Description Statistics

In this study, the variable values for each experimental station were calculated using the arithmetic mean of the data from their academic outputs, as shown in Table 3. The analysis of the interdisciplinarity of each experimental station reveals that the research infrastructures with multiple experimental stations, such as synchrotron light sources, are significantly capable of facilitating research in a wide range of scientific fields. These include disciplines such as materials science, nanotechnology, condensed matter physics and energy studies. This generality underscores the crucial role of synchrotron light sources in promoting interdisciplinary integration and advancement, and thus advances in various fields of science. Specifically, the scientific contributions from the experimental stations 02-BM, 02-ID, 13-BM, 15-ID, 18-ID and 32-ID span a broad disciplinary spectrum and contain findings from over 60 different research areas. In contrast, the academic contributions from the test stations 14-ID(XSD), 17-BM(IMCA), 27-ID and 29-ID were characterized by a narrower disciplinary breadth, as they each covered no more than 15 research areas.

The study involved a descriptive statistical analysis of the correlation between the above variables, the results of which are shown in Table 4. Given that the variables were not normally distributed according to the Kolmogorov-Smirnov test (K-S test), Spearman's correlation analysis was used. It can be seen that there is a correlation between the novelty of the academic outputs of the experimental stations and five variables, namely disparity, RS, number of references, number of authors and proportion of research papers. Additionally, the relative independence index (PCR) is significantly associated with the disparity index, RS, and the operational duration of the experimental station. Subsequent application of multiple linear regression tests for all variables showed that the variance inflation factor (VIF) for each variable did not exceed 5, with a peak value of 3.448. This result indicates that there are no significant multicollinearity issues within the variables, confirming the feasibility of regression analysis to further interpret these relationships.

Table 3. Statistical Values of Relevant Variables of APS Light Source Experimental Stations

Experimental Stations	**Number of Papers**	**NOVEL**	**PCR**	**VAR**	**BAL**	**DIS**	**RS**	**NR**	**LTI**	**NAU**	**NAD**	**NRP**	**NY**
01-BM	210	0.4286	0.2791	42	0.0130	0.9993	0.2422	32.9190	11.8857	6.5857	2.8095	0.9857	24
01-ID	533	0.4178	0.2616	52	0.0064	0.9994	0.2320	41.1651	12.9812	6.2083	2.9737	0.9662	24
02-BM	310	0.4325	0.3215	74	0.0074	0.9998	0.2075	40.2097	12.4065	6.7903	3.1774	0.9613	26
02-ID	515	0.4290	0.3138	84	0.0060	0.9996	0.2673	44.8019	12.2194	8.3670	3.6194	0.9573	29
03-ID	346	0.4512	0.1649	30	0.0126	0.9989	0.3017	45.5000	11.8844	8.4335	3.8497	0.9798	26
04-ID	471	0.4520	0.2198	28	0.0109	0.9987	0.2775	41.4161	11.0106	8.8174	4.1401	0.9873	21
05-BM	305	0.4459	0.2954	45	0.0103	0.9996	0.2507	48.0852	12.5967	5.8623	2.4295	0.9869	23
05-ID	738	0.4	0.	5	0.	0.	0.	45	11	5.	2.	0.	2

		477	2327	3	0067	9988	2953	.7995	.5976	5108	0556	9878	3
06-BM	37	0.4196	0.20114	21	0.05446	0.99999	0.24003	44.7297	13.9459	7.0000	2.94559	1.00000	5
06-ID	583	0.4406	0.19440	36	0.00075	0.99992	0.26995	41.7959	11.5901	7.7221	3.50433	0.98633	20
07-BM	53	0.3954	0.27996	19	0.03444	0.99998	0.15591	34.7925	12.0377	6.3962	2.52833	0.98111	11
07-ID	188	0.4216	0.29227	35	0.01441	0.99994	0.21443	35.3298	11.1543	9.5000	3.87723	0.97877	24
08-BM	119	0.4607	0.21147	35	0.02221	0.99994	0.22251	42.3866	13.1176	6.7983	2.57798	0.96664	18
08-ID	640	0.4458	0.22294	36	0.01000	0.99983	0.29666	48.2828	11.6016	7.8844	3.07734	0.98891	23
09-BM	307	0.4716	0.25885	32	0.01337	0.99993	0.26670	54.0945	13.5016	9.5179	3.33355	0.98770	21
09-ID	338	0.4341	0.24554	52	0.00086	0.99997	0.24668	45.7367	12.9142	7.4467	3.26663	0.99111	21

10-BM	365	0.4714	0.2762	32	0.0117	0.9993	0.2707	54.0301	13.3726	8.2548	2.9781	0.9973	15
10-ID	526	0.4462	0.3078	51	0.0073	0.9988	0.2363	46.5057	13.4563	7.1540	2.9696	0.9943	22
11-BM	930	0.4827	0.2180	48	0.0063	0.9981	0.2587	46.2828	12.3688	7.0054	2.9817	0.9903	13
11-ID	1757	0.4557	0.2358	59	0.0038	0.9977	0.2674	46.9943	12.6135	7.7245	3.3625	0.9824	22
12-BM	531	0.4493	0.2809	46	0.0083	0.9987	0.2740	48.9096	12.5639	6.8870	2.6629	0.9831	23
12-ID	1070	0.4543	0.2430	50	0.0050	0.9986	0.2853	49.9280	12.1028	7.2794	2.8664	0.9785	23
13-BM	493	0.4385	0.2157	73	0.0063	0.9997	0.2323	49.7404	12.5477	6.0751	3.2982	0.9797	23
13-ID	943	0.4390	0.1989	53	0.0047	0.9993	0.2446	48.2587	12.5907	6.6872	3.8070	0.9767	24
14-BM	746	0.4572	0.207	44	0.008	0.997	0.303	44.4504	13.7078	5.9115	2.383	0.978	27

			2		1	3	5				4	6	
14-IDBIO	262	0.4464	0.1837	32	0.0161	0.9989	0.2701	44.7863	12.6832	7.4618	2.7824	0.9618	27
14-IDXSD	13	0.4150	0.1276	10	0.1268	0.9997	0.2074	34.6154	13.7692	9.6153	3.6154	1.0000	7
15-ID	838	0.4554	0.2491	64	0.0051	0.9985	0.2511	47.7554	12.1575	6.5358	2.7721	0.9737	23
16-BM	403	0.4582	0.2027	29	0.0124	0.9991	0.2841	43.0720	11.1017	8.1787	3.8660	0.9901	16
16-ID	788	0.4486	0.2012	35	0.0075	0.9982	0.2790	41.7830	11.2081	7.2284	3.6053	0.9810	18
17-BMIMCA	84	0.4663	0.3190	14	0.0303	0.9995	0.2531	36.4762	14.8690	16.6786	1.6310	1.0000	17
17-BMXSD	334	0.4841	0.2386	32	0.0140	0.9984	0.2822	50.7096	12.5629	8.3293	3.3683	0.9760	10
17-ID	1046	0.4631	0.2777	36	0.0059	0.9968	0.3053	38.3193	13.6597	14.8050	2.3356	0.9924	23
18-ID	506	0.433	0.2	65	0.0	0.9	0.3	49.7	13.0	6.40	2.9	0.9	23

		6	037		087	991	100	609	889	51	783	427	
19-BM	556	0.4560	0.1935	37	0.0105	0.9980	0.3163	43.4155	13.9658	7.2842	2.6043	0.9928	22
19-ID	1676	0.4626	0.1929	56	0.0056	0.9945	0.3397	47.4230	13.3747	7.8162	2.8759	0.9761	28
20-BM	697	0.4640	0.2650	53	0.0066	0.9988	0.2508	48.0933	12.9096	7.7102	3.1105	0.9957	21
20-ID	361	0.4411	0.2916	50	0.0079	0.9997	0.2248	46.3684	13.3324	7.4931	3.4875	0.9806	27
21-ID	1519	0.4673	0.1826	54	0.0056	0.9947	0.3389	46.4174	13.5583	9.4674	2.6386	0.9888	19
22-BM	565	0.4668	0.1812	39	0.0101	0.9984	0.3289	46.6566	14.1451	8.1027	2.6053	0.9982	16
22-ID	1592	0.4673	0.1720	47	0.0057	0.9959	0.3386	47.1753	13.8411	8.2858	2.5660	0.9950	19
23-ID	1665	0.4633	0.1673	50	0.0058	0.9960	0.3404	48.0643	12.9495	8.4180	2.7471	0.9922	16

24-ID	1928	0.4681	0.1545	45	0.0600	0.9911	0.3517	47.9507	12.7697	7.5233	2.7474	0.9948	16
26-ID	105	0.4361	0.2525	29	0.0241	0.9996	0.2234	45.5619	11.3238	8.9333	3.2381	0.9524	14
27-ID	34	0.4584	0.1090	11	0.0749	0.9999	0.2496	46.5882	10.6471	10.9412	5.0588	1.0000	6
29-ID	21	0.5006	0.1973	14	0.0836	0.9999	0.1862	48.2381	11.0476	13.4286	4.8571	1.0000	7
30-ID	120	0.4421	0.1521	24	0.0275	0.9998	0.2822	43.2167	10.7333	9.3917	4.2417	0.9833	14
31-ID	283	0.4610	0.2180	31	0.0149	0.9992	0.3093	43.8799	13.9187	11.2261	2.6996	0.9894	20
32-ID	423	0.4322	0.2686	77	0.0065	0.9996	0.2303	40.0969	12.7234	7.4421	3.0284	0.9740	21
33-BM	206	0.4498	0.2598	26	0.0172	0.9991	0.2635	38.1942	11.6456	8.3204	3.5049	0.9903	18
33-ID	253	0.4388	0.299	43	0.012	0.999	0.248	40.0514	11.1186	7.0198	3.142	0.980	22

			3		4	4	1				3	2	
34-ID	316	0.4432	0.26224	36	0.01117	0.99990	0.25338	35.9430	11.1677	7.3449	3.43304	0.96844	20
35-ID	44	0.4359	0.24334	15	0.04998	0.99999	0.21881	44.4091	12.4773	9.2955	2.59009	1.00000	6

Table4. Description Statistics and Correlations

Vari	1	2	3	4	5	6	7	8	9	10	11	12
1.N	1											
2.PC	-	1										
3.V	-	0.21	1									
4.B	-	-	-	1								
5.DI	-	0.27	-	0.50	1							
6.RS	0.55	-	0.15	-	-	1						
7.N	0.46	-	0.32	-	-	0.3	1					
8.LT	0.23	-	0.18	-	-	0.2	0.1	1				
9.N	0.36	-	-	0.33	0.03	0.1	-	0.00	1			
10.N	-	-	-	0.14	0.32	-	-	-	0.3	1		
11.N	0.37	-	-	0.30	-	0.1	0.0	0.25	0.4	-	1	
12.N	-	0.31	0.61	-	-0.2	0.1	0.0	0.05	-	-	-	1
Mea	0.44	0.23	41.2	0.01	0.99	0.2	44.	12.5	8.1	3.	0.98	19.
Med	0.44	0.23	39.0	0.01	0.99	0.2	45.	12.5	7.7	2.	0.98	21.
SD	0.01	0.05	16.9	0.02	0.00	0.0	4.8	1.01	2.0	0.	0.01	6.0
Min	0.39	0.10	10.0	0.00	0.99	0.1	32.	10.6	5.5	1.	0.94	5.0
Max	0.50	0.32	84.0	0.12	1.00	0.3	54.	14.8	16.	5.	1.00	29.
VIF			3.07	3.44	3.11	3.3	1.8	3.10	1.6	2.	1.79	3.4
Tole			0.32	0.29	0.32	0.3	0.5	0.32	0.6	0.	0.55	0.2

Regression Analysis

In this study, the use of the ordinary least squares (OLS) regression model is based on the nature of the novelty index NOVEL and the relative disruption index DEP as non-negative continuous variables. To mitigate the effects of potential heteroscedasticity that could bias the analysis, the study uses a robust standard error model to counteract heteroscedasticity and thus increase the reliability of the regression results.

The results in Table 5 indicate that, although a positive correlation ($r = 0.013$) was observed between the interdisciplinarity index (RS) and the novelty index (NOVEL), this relationship is not statistically significant. It is noteworthy that there is a statistically significant negative correlation between the disciplinary diversity (DIS) of the experimental stations and the novelty of their academic outputs. One possible reason for this phenomenon is that research directions may be too broad if a research station's research spans several very different disciplines. This can lead to inadequate allocation of resources and insufficient focus on each research direction, which affects the depth and novelty of the research. In addition, the increasing complexity of data and the challenges associated with its interpretation can affect the clarity of research findings. This complexity can lead to more conservative conclusions, reducing the novelty of the resulting scientific work. In addition, both the number of references and the number of authors are significantly positively correlated with the novelty of the paper. A larger number of authors with diverse research backgrounds and professional fields can increase the visibility and citation probability of the paper. This diversity also contributes to the innovative novelty of the work, as it incorporates a broader range of perspectives and expertise. Moreover, the inclusion of a greater number of references indicates an extensive literature review and a comprehensive engagement with existing research, which can provide a solid foundation for novel contributions.

When regression analysis was performed on the PCR index, a significant negative correlation between the interdisciplinarity index RS and the PCR, with a coefficient (r) of -0.931 and a statistical significance value (p) of less than 0.001. The regression analysis shows that while experimental stations with a research infrastructure generally have a high level of generality, those with a higher concentration of disciplines are likely to produce more innovative academic papers. Conversely, stations with a strong interdisciplinary orientation are crucial for promoting disciplinary integration, although their outputs tend to be less innovative. This trend underscores the central role of disciplinary depth in fostering innovation and suggests that a rich web of specialized fields of knowledge can serve as fertile ground for breakthroughs in research.

Table5. Regression Analysis Results of Variables on Novelty and Disruption of Academic

Variables	NOVEL OLS regression coefficient	PCR OLS regression coefficient
VAR	-0.000(-0.735)	0.000(0.842)
BAL	-0.246(-1.884)	-0.957**(-3.062)
DIS	-3.460*(-2.173)	-0.074(-0.018)
RS	0.013(0.139)	-0.931***(-6.340)
NR	0.002***(4.219)	0.000(0.431)
LTI	-0.001(-0.551)	-0.008(-1.450)
NAU	0.004**(3.147)	0.008***(4.795)
NAD	0.001(0.423)	-0.035***(-3.962)
NRP	0.203(1.055)	0.389(0.932)
NY	-0.000(-0.526)	0.002(1.452)
Constant	3.609*(2.151)	0.257(0.062)
R^2	0.649	0.732
AdjustmentR^2	0.566	0.669

Note: significance level ***p<0.001, **p<0.01, *p<0.05, t value in parentheses.

Robustness Check

The authors conducted two robustness checks. First, using the same dataset, we used Google's Universal Sentence Encoder as a word embedding library, replacing SciBERT to calculate novelty, while we used the relative dependency index DEP instead of the independence index PCR. The index DEP is calculated using the following formula:

(1) $DEP = \frac{N_{j\times \text{cited}}}{N_i + N_j}$

where $N_{j\times cited}$ indicates the number of times the citing document cites the FP's cited reference, N_i indicates the number of publications that cite both the FP and its cited references, and N_j denotes the number of publications that cite only the FP. The DEP is an inverse indicator that is converted into a positive indicator after forward processing. The calculation formula is as follows:

(2) $X = \frac{X - X_{min}}{X_{max} - X_{min}}$

where X_{max} represents the maximum value of all values in the variable's dataset. X_{min} denotes the minimum of all values in the variable dataset. X indicates a value in the original data.

Second, we employed the interdisciplinary diversity index DIV* (Zhang & Leydesdorff, 2021), as an alternative measure in the regression analysis, in place of the Rao-Stirling index. The DIV* index as is computed using the follow formula:

$$(3)\ DIV^* = n_c * [1 - G(c)] * \sum_{\substack{i=1, \\ j=1, \\ i \neq j}}^{\substack{j=n_c \\ i=n_c}} (d_{ij})$$

where n_c indicates the variety of disciplinary categories covered by the paper *c*, $G(c)$ is the Gini coefficient of the *c;* and d_{ij} is the disparity between the disciplinary categories of the *i*-th and *j*-th papers cited in the evaluated paper.

The result presented in Table 6 reveals a significant negative correlation between the disciplinary disparity index of the experimental stations and the novelty calculated with Google's Universal Sentence Encoder. In addition, there is a negative correlation between the interdisciplinarity index DIV* and the relative dependence index DEP. The results of the correlation analysis were generally consistent with the results of the novelty calculation using the SciBERT model and the relative independence index PCR as the dependent variable.

When employing the interdisciplinary index DIV* as the independent variable, the result of the correlation analysis remains consistent with that obtained through the use of the interdisciplinary index DIV*. In particular, the regression analysis shows a negative correlation between the disciplinary disparity and the novelty as well as a positive correlation between the interdisciplinarity index RS and novelty. However, neither of these correlations reached statistical significance. There is also a negative correlation between the interdisciplinary index DIV* of the experimental station and the relative independence index PCR.

Table6. Robustness Test Results

Variables	NOVELG (OLS)	DEP (OLS)	Variables	NOVEL (OLS)	PCR (OLS)
VAR	0.000 (0.945)	-0.008 (-1.712)	*VAR*	-0.000 (-0.244)	-0.000 (-0.202)
BAL	-0.196 (-1.721)	-4.535 (-1.527)	*BAL*	-0.219 (-1.581)	-0.842** (-2.576)
DIS	-4.587** (-2.930)	-28.038 (-0.708)	*DIS*	-1.434 (-0.617)	-16.212 (-1.900)
RS	0.140 (1.608)	-3.135* (-1.999)	*DIV**	0.007 (0.960)	-0.107*** (-4.724)
NR	0.001 (1.589)	0.004 (0.560)	*NR*	0.002*** (4.430)	-0.000 (-0.216)
NAU	0.007* (2.176)	0.052 (0.634)	*NAU*	-0.002 (-0.639)	-0.007 (-0.957)
NAD	0.001 (1.106)	-0.015 (-0.764)	*NAD*	0.004*** (3.397)	0.004 (1.933)
NRP	-0.006 (-1.493)	0.090 (0.726)	*NRP*	0.000 (0.094)	-0.023 (-1.808)

NY	-0.174 (-0.785)	0.019 (0.005)	*NY*	0.207 (1.063)	0.268 (0.711)
Constant	5.664*** (3.504)	29.062 (0.724)	Constant	1.584 (0.661)	16.366 (1.917)
R^2	0.708	0.243	R^2	0.654	0.702
Adjustment R^2	0.639	0.04	Adjustmen tR^2	0.572	0.632

Note: significance level ***p<0.001, **p<0.01, *p<0.05, t value in parentheses

Conclusion

The study uses a bibliometric analysis to explore the interdisciplinarity of research infrastructure experimental stations, a critical aspect for understanding the dynamics of scientific research. Using the novelty and disruption index, the study evaluates the interaction between the interdisciplinarity of the stations and the novelty and disruption of their scientific outputs. The results underline the essential interdisciplinary character of these stations and their role in promoting basic research in various fields.

The regression analysis shows that a negative correlation between the disciplinary disparity of the experimental station and the novelty of its academic outputs. The technological and methodological requirements of different scientific disciplines for synchrotron light sources are very different, which increases the complexity of the operation and maintenance of such facilities. Focusing on specific scientific disciplines allows the synchrotron facility to specialize and refine its techniques and methods, helping to achieve breakthrough and novel results in these areas. By optimizing the allocation of resources and improving the quality of technical support, the facility can more effectively promote research quality in specific disciplines, increasing the impact and originality of research results. Moreover, the trend that experimental stations focusing on specific disciplines often produce publications with a higher innovative disruption is significant. This indicates a robust positive correlation between specialization and the capacity to conduct groundbreaking research. Specialization likely allows researchers to delve deeper into their area of expertise, push the boundaries of knowledge and make high-impact discoveries. Conversely, experimental stations that cover a broader interdisciplinary spectrum play a crucial role in the scientific ecosystem by enabling a wide range of scientific inquiries. Although these stations tend to produce fewer groundbreaking publications, their contributions should not be underestimated. Interdisciplinary research stations are central to the integration of different perspectives and methods, which can lead to the development of new research paradigms and advancements at the interfaces of traditional disciplines. These facilities are particularly suited to answering complex scientific questions that require a holistic approach that utilizes the strengths and insights of different disciplines.

This dichotomy points to a complex interplay between the breadth and depth of disciplinary focus within experimental stations and their impact on scientific progress. It suggests that while interdisciplinary approaches are critical for the cross-pollination of ideas and the advancement of holistic understanding across fields, the concentration of expertise within specific disciplines remains a key driver of novel research findings. This insight invites further exploration into how experimental stations can optimally balance interdisciplinary collaboration with disciplinary specialization to maximize their contribution to the frontiers of scientific knowledge. Given these insights findings, the strategic development and design of research infrastructure experimental stations should focus on configurations that optimize academic output, advancements, and impact. Such a strategic focus is essential to foster scientific progress and maximize academic impact, thereby enhancing the utility of research infrastructures as key assets in the national and international research landscape. This study not only contributes to our understanding of the structural dynamics of research infrastructures, but also helps to formulate more informed decisions regarding their future development and optimization.

This study meticulously evaluates the interdisciplinarity of experimental stations within research infrastructures using bibliometric indices like the scientific publications' novelty and relative innovative disruption. While these metrics provide valuable insights, it is important to acknowledge the inherent limitations of this analytical approach. Firstly, the complexity of academic advancement cannot be fully captured solely by measures of publication quality and impact. Other significant bibliometric indicators, such as citation frequency, number of references, and the citation time window of the assessed publications, also play a critical role in defining the quality of academic outputs. These factors should be comprehensively examined in future studies to provide a more nuanced understanding of the academic impact and advancement at experimental stations. In addition, the robustness of the novelty and relative disruption indices used in this study, which are central to assessing the innovative quality of academic outputs from multi-station research infrastructures, requires further validation. This validation should be conducted in other research infrastructures to ensure its general applicability and reliability.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Adeyemi, H. (1996). Academic Research Infrastructure: Acquisition of Instrumentation to Provide Scientific and Data Visualization Capability at Clark Atlanta University [Awarded Grant].

Amplayo, R. K., Hong, S., & Song, M. (2018). Network-based approach to detect novelty of scholarly literature [Article]. *INFORMATION SCIENCES*, *422*, 542-557. https://doi.org/10.1016/j.ins.2017.09.037

APS. *APS Publications*. https://www.aps.anl.gov/Science/Publications

Arthur, W. B. (2010). *The Nature of Technology: What It Is and How It Evolves*. Penguin Adult. https://books.google.com.tw/books?id=phkgssCTlpgC

Arts, S., Hou, J., & Gomez, J. C. (2021). Natural language processing to identify the creation and impact of new technologies in patent text: Code, data, and new measures [Article]. *Research Policy*, *50*(2), Article 104144. https://doi.org/10.1016/j.respol.2020.104144

Balsmeieri, B., Assaf, M., Chesebro, T., Fierro, G., Johnson, K., Johnson, S., . . . Fleming, L. (2018). Machine learning and natural language processing on the patent corpus: Data, tools, and new measures [Article]. *JOURNAL OF ECONOMICS & MANAGEMENT STRATEGY*, *27*(3), 535-553. https://doi.org/10.1111/jems.12259

Beck, S., Mahdad, M., Beukel, K., & Poetz, M. (2019). The Value of Scientific Knowledge Dissemination for Scientists-A Value Capture Perspective [Article]. *Publications*, *7*(3), Article 54. https://doi.org/10.3390/publications7030054

Bittmann, F., Tekles, A., & Bornmann, L. (2022). Applied usage and performance of statistical matching in bibliometrics: The comparison of milestone and regular papers with multiple measurements of disruptiveness as an empirical example [Article]. *Quantitative Science Studies*, *2*(4), 1246-1270. https://doi.org/10.1162/qss_a_00158

Bornmann, L., Devarakonda, S., Tekles, A., & Chacko, G. (2020). Are disruption index indicators convergently valid? The comparison of several indicator variants with assessments by peers [Article]. *Quantitative Science Studies*, *1*(3), 1242-1259. https://doi.org/10.1162/qss_a_00068

Bornmann, L., & Tekles, A. (2019). Disruption index depends on length of citation window [Article]. *Profesional De La Informacion*, *28*(2), Article e280207. https://doi.org/10.3145/epi.2019.mar.07

Boudreau, K. J., Guinan, E. C., Lakhani, K. R., & Riedl, C. (2016). Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science [Article]. *Management Science*, *62*(10), 2765-2783. https://doi.org/10.1287/mnsc.2015.2285

Bu, Y., Waltman, L., & Huang, Y. (2021). A multidimensional framework for characterizing the citation impact of scientific publications [Article]. *Quantitative Science Studies*, *2*(1), 155-183. https://doi.org/10.1162/qss_a_00109

Christensen, C. M. (2015). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business Review Press. https://books.google.com.hk/books?id=lURBCgAAQBAJ

Fabre, R., Egret, D., Schopfel, J., & Azeroual, O. (2021). Evaluating the scientific impact of research infrastructures: The role of current research information systems [Article]. *Quantitative Science Studies*, *2*(1), 42-64. https://doi.org/10.1162/qss_a_00111

Fecher, B., Kahn, R., Sokolovska, N., Voelker, T., & Nebe, P. (2021). Making a Research Infrastructure: Conditions and Strategies to Transform a Service into an Infrastructure [Article]. *Science and Public Policy*, *48*(4), 499-507. https://doi.org/10.1093/scipol/scab026

Foster, D. (2018). Long, short, and efficient titles for research articles.

Foster, J. G., Rzhetsky, A., & Evans, J. A. (2015). Tradition and Innovation in Scientists' Research Strategies [Review]. *American Sociological Review*, *80*(5), 875-908. https://doi.org/10.1177/0003122415601618

Funk, R. J., & Owen-Smith, J. (2017). A Dynamic Network Measure of Technological Change [Article]. *Management Science*, *63*(3), 791-817. https://doi.org/10.1287/mnsc.2015.2366

Hamers, L., Hemeryck, Y., Herweyers, G., Janssen, M., Keters, H., Rousseau, R., & Vanhoutte, A. (1989). Similarity Measures in Scientometric Research - The Jaccard Index Versus Salton Cosine Formula [Article]. *Information Processing & Management*, *25*(3), 315-318. https://doi.org/10.1016/0306-4573(89)90048-4

Heffernan, K., & Teufel, S. (2018). Identifying problems and solutions in scientific text [Article]. *Scientometrics*, *116*(2), 1367-1382. https://doi.org/10.1007/s11192-018-2718-6

Helliwell, J. R., & Mitchell, E. P. (2015). Synchrotron radiation macromolecular crystallography: science and spin-offs [Review]. *Iucrj*, *2*, 283-291. https://doi.org/10.1107/s205225251402795x

Hentschel, K. (2015). A periodization of research technologies and of the emergency of genericity [Article]. *Studies in History and Philosophy of Modern Physics*, *52*, 223-233. https://doi.org/10.1016/j.shpsb.2015.07.009

Jeon, D., Ahn, J. M., Kim, J., & Lee, C. (2022). A doc2vec and local outlier factor approach to measuring the novelty of patents [Article]. *TECHNOLOGICAL*

FORECASTING AND SOCIAL CHANGE, *174*, Article 121294. https://doi.org/10.1016/j.techfore.2021.121294

Jeon, D., Lee, J., Ahn, J. M., & Lee, C. (2023). Measuring the novelty of scientific publications: A fastText and local outlier factor approach [Article]. *Journal of Informetrics*, *17*(4), Article 101450. https://doi.org/10.1016/j.joi.2023.101450

Jiang, Y., & Liu, X. (2023). A construction and empirical research of the journal disruption index based on open citation data [Article]. *Scientometrics*, *128*(7), 3935-3958. https://doi.org/10.1007/s11192-023-04737-y

Jones, B. F., Wuchty, S., & Uzzi, B. (2008). Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science [Article]. *Science*, *322*(5905), 1259-1262. https://doi.org/10.1126/science.1158357

Kondo, T., Nanba, H., Takezawa, T., & Okumura, M. (2011, 2011//). Technical Trend Analysis by Analyzing Research Papers' Titles. Human Language Technology. Challenges for Computer Science and Linguistics, Berlin, Heidelberg.

Leydesdorff, L., & Rafols, I. (2011). Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations [Article]. *Journal of Informetrics*, *5*(1), 87-100. https://doi.org/10.1016/j.joi.2010.09.002

Leydesdorff, L., Wagner, C. S., & Bornmann, L. (2019). Interdisciplinarity as diversity in citation patterns among journals: Rao-Stirling diversity, relative variety, and the Gini coefficient [Article]. *Journal of Informetrics*, *13*(1), 255-269. https://doi.org/10.1016/j.joi.2018.12.006

Lu, W., Huang, S., Yang, J., Bu, Y., Cheng, Q., & Huang, Y. (2021). Detecting research topic trends by author-defined keyword frequency [Article]. *Information Processing & Management*, *58*(4), Article 102594. https://doi.org/10.1016/j.ipm.2021.102594

Luo, Z., Lu, W., He, J., & Wang, Y. (2022). Combination of research questions and methods: A new measurement of scientific novelty [Article]. *Journal of Informetrics*, *16*(2), Article 101282. https://doi.org/10.1016/j.joi.2022.101282

Matsumoto, K., Shibayama, S., Kang, B., & Igami, M. (2021). Introducing a novelty indicator for scientific research: validating the knowledge-based combinatorial approach [Article]. *Scientometrics*, *126*(8), 6891-6915. https://doi.org/10.1007/s11192-021-04049-z

Mutz, R. (2022). Diversity and interdisciplinarity: Should variety, balance and disparity be combined as a product or better as a sum? An information-theoretical and statistical estimation approach [Article]. *Scientometrics*, *127*(12), 7397-7414. https://doi.org/10.1007/s11192-022-04336-3

OECD. (2020). *OPTIMISING THE OPERATION AND USE OF NATIONAL RESEARCH INFRASTRUCTURES.*

Park, M., Leahey, E., & Funk, R. J. (2023). Papers and patents are becoming less disruptive over time [Article]. *Nature*, *613*(7942), 138-+. https://doi.org/10.1038/s41586-022-05543-x

Porter, A. L., & Chubin, D. E. (1985). An Indicator of Cross-Disciplinary Research [Article]. *Scientometrics*, *8*(3-4), 161-176. https://doi.org/10.1007/bf02016934

Porter, A. L., Cohen, A. S., Roessner, J. D., & Perreault, M. (2007). Measuring researcher interdisciplinarity [Article]. *Scientometrics*, *72*(1), 117-147. https://doi.org/10.1007/s11192-007-1700-5

Porter, A. L., & Rafols, I. (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time [Article]. *Scientometrics*, *81*(3), 719-745. https://doi.org/10.1007/s11192-008-2197-2

Rosenberg, N. (1992). Scientific Instrumentation and University-Research [Article]. *Research Policy*, *21*(4), 381-390. https://doi.org/10.1016/0048-7333(92)90035-3

Schumpeter, J. A. (1961). *The Theory of Economic Development: An Inquiry Into Profits, Capital, Credit, Interest, and the Business Cycle*. Oxford University Press. https://books.google.com.hk/books?id=SyvsLQEACAAJ

Sciences, N. A. o., Engineering, N. A. o., & Medicine, I. o. (2005). *Facilitating Interdisciplinary Research*. The National Academies Press. https://doi.org/doi:10.17226/11153

Shibayama, S., Yin, D., & Matsumoto, K. (2021). Measuring novelty in science with word embedding [Article]. *PLOS ONE*, *16*(7), Article e0254034. https://doi.org/10.1371/journal.pone.0254034

Soderstrom, K. R., Astrom, F., & Hallonsten, O. (2022). Generic instruments in a synchrotron radiation facility [Article]. *Quantitative Science Studies*, *3*(2), 420-442. https://doi.org/10.1162/qss_a_00190

Sreenivasan, S. (2013). Quantitative analysis of the evolution of novelty in cinema through crowdsourced keywords [Article]. *SCIENTIFIC REPORTS*, *3*, Article 2758. https://doi.org/10.1038/srep02758

Staley, R. (2002). Instrumentation between science, state and industry, vol 22 [Book Review]. *Social Studies of Science*, *32*(3), 469-476, Article Unsp [0306-3127(200206)32:3;469-476;027786]. https://doi.org/10.1177/0306312702032003005

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society [Review]. *Journal of the Royal Society Interface*, *4*(15), 707-719. https://doi.org/10.1098/rsif.2007.0213

Sugimoto, C. R., Ni, C., Russell, T. G., & Bychowski, B. (2011). Academic Genealogy as an Indicator of Interdisciplinarity: An Examination of Dissertation Networks in Library and Information Science [Article]. *Journal of the American Society for Information Science and Technology*, *62*(9), 1808-1828. https://doi.org/10.1002/asi.21568

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical Combinations and Scientific Impact [Article]. *Science*, *342*(6157), 468-472. https://doi.org/10.1126/science.1240474

Varga, A. (2018). Novelty and Foreseeing Research Trends: The Case of Astrophysics and Astronomy [Article]. *ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES*, *236*(1), Article 21. https://doi.org/10.3847/1538-4365/aab765

Verhoeven, D., Bakker, J., & Veugelers, R. (2016). Measuring technological novelty with patent-based indicators [Article]. *Research Policy*, *45*(3), 707-723. https://doi.org/10.1016/j.respol.2015.11.010

Wang, S., Ma, Y., Mao, J., Bai, Y., Liang, Z., & Li, G. (2023). Quantifying scientific breakthroughs by a novel disruption indicator based on knowledge entities [Article]. *Journal of the Association for Information Science and Technology*, *74*(2), 150-167. https://doi.org/10.1002/asi.24719

Westerski, A., & Kanagasabai, R. (2019). In search of disruptive ideas: outlier detection techniques in crowdsourcing innovation platforms. *International Journal of Web Based Communities*, *15*(4), 344-367. https://doi.org/10.1504/IJWBC.2019.103185

Wu, L., Wang, D., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology [Article]. *Nature*, *566*(7744), 378-+. https://doi.org/10.1038/s41586-019-0941-9

Yan, Y., Tian, S., & Zhang, J. (2020). The impact of a paper's new combinations and new components on its citation [Article]. *Scientometrics*, *122*(2), 895-913. https://doi.org/10.1007/s11192-019-03314-6

Yang, A. J., Deng, S., Wang, H., Zhang, Y., & Yang, W. (2023). Disruptive coefficient and 2-step disruptive coefficient: Novel measures for identifying vital nodes in complex networks [Article]. *Journal of Informetrics*, *17*(3), Article 101411. https://doi.org/10.1016/j.joi.2023.101411

Zhang, H., Zhang, C., & Wang, Y. (2024). Revealing the technology development of natural language processing: A Scientific entity-centric perspective [Article]. *Information Processing & Management*, *61*(1), Article 103574. https://doi.org/10.1016/j.ipm.2023.103574

Zhang, L., & Leydesdorff, L. (2021). The Scientometric Measurement of Interdisciplinarity and Diversity in the Research Portfolios of Chinese

Universities [Article]. *Journal of Data and Information Science*, *6*(4), 13-35. https://doi.org/10.2478/jdis-2021-0027

Zhang, L., Wang, D., & Zhang, L. (2019). Research of The Influence of Interdisciplinarity and Team Cooperation on the Scientific Effects Based on Large-scale Scientific Facilities. *Journal of Management World*, *35*(12), 199-212. https://doi.org/10.19744/j.cnki.11-1235/f.2019.0173

Zhang, Y., Cai, X., Fry, C. V., Wu, M., & Wagner, C. S. (2021). Topic evolution, disruption and resilience in early COVID-19 research [Article]. *Scientometrics*, *126*(5), 4225-4253. https://doi.org/10.1007/s11192-021-03946-7

Zhou, H., Guns, R., & Engels, T. C. E. (2022). Are social sciences becoming more interdisciplinary? Evidence from publications 1960-2014 [Article]. *Journal of the Association for Information Science and Technology*, *73*(9), 1201-1221. https://doi.org/10.1002/asi.24627

Zhou, Y., Dong, F., Liu, Y., & Ran, L. (2021). A deep learning framework to early identify emerging technologies in large-scale outlier patents: an empirical study of CNC machine tool [Article]. *Scientometrics*, *126*(2), 969-994. https://doi.org/10.1007/s11192-020-03797-8

Zhou, Y., Lin, H., Liu, Y., & Ding, W. (2019). A novel method to identify emerging technologies using a semi-supervised topic clustering model: a case of 3D printing industry [Article]. *Scientometrics*, *120*(1), 167-185. https://doi.org/10.1007/s11192-019-03126-8