Renewable Energy 209 (2023) 516–532

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Battery innovation and the Circular Economy: What are patents revealing?

Philipp Metzger [a], Sandro Mendonça [b,c,d], José A. Silva [e], Bruno Damásio [a,*]

[a] NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Campus de Campolide, 1070-312 Lisboa, Portugal
[b] ISCTE Business School, Business Research Unit (BRU-IUL), Avenida das Forças Armadas, 1649-026 Lisboa, Portugal
[c] UECE/REM – ISEG/ University of Lisbon, Rua do Quelhas 6, 1200-781 Lisboa, Portugal
[d] SPRU, University of Sussex, Falmer, Brighton BN1 9RH, UK
[e] Instituto Dom Luiz, Faculdade de Ciências da Universidade de Lisboa, Campo Grande Edifício C1, Piso 1, 1749-016 Lisboa, Portugal

ARTICLE INFO

Keywords:
Secondary batteries
Patents
Technometrics
Text mining
Circular economy

ABSTRACT

This analysis of over 90,000 secondary battery innovations (measured by international patent families) provides a comprehensive account of the long-run progress of a knowledge base with a key role in the transition to a transformative, closed-loop, Circular Economy. Innovation accelerated globally from 2000 to 2019, a sustained dynamic mostly originating in Asia. Patterns of less toxicity and more diversity in technological trajectories are detected and found to bear evidence of pro-circularity. We find a number of emergent technological trajectories, such as solid-state, lithium–sulfur, redox-flow and sodium-ion batteries, each one with a different potential to push ahead the circularity pathway, and which allow for the detection of country clusters. Through a methodology that can be of interest for further research, we examine the extent to which batteries have circular characteristics.

1. Introduction

Since the early days of the first Industrial Revolution in the late 18th century, global energy consumption has been on the rise [1]. Two centuries later, by the time the informational revolution was taking hold [2,3], the pressure was on to reduce CO2 emissions derived from the coal and oil paradigms that preceded it. New socio-technical compacts, from the Rio "Earth Summit" of 1992 to the Paris Agreement of 2015, have been fostering a holistic reform of social organization and of the energy sector in particular. To structure this process of change there is a growing need for new solutions in terms of power generation, distribution, storage, and upkeep. In this context, the Circular Economy framework has been proposed to reconcile economic and sustainable development [4,5].

The importance of batteries has been growing as a solution in a very dynamic puzzle. As a set of technologies at the intersection of the clean-digital transition, their role is expected to grow further in the coming decades [6]. A report about electricity storage developments published by the International Energy Agency (IEA) in association with the European Patent Office (EPO), asserts that "the level of deployment and the range of applicability of batteries [...] expands dramatically" in the foreseeable future [7, p. 28]. In particular, battery technologies will move beyond consumer appliances and into industrial-size types of equipment: "Charging batteries in electric vehicles will become the largest single source of electricity demand, accounting for around 5% of global demand by 2050" [7, p. 29]. Furthermore, "the use of batteries in stationary energy storage applications is [already] growing exponentially" [7, p. 32].

Identifying and monitoring the rate and direction of battery innovation as a condition for a low-carbon future is thus analytically worthwhile and strategically urgent. A growing body of empirical work has recently approached the battery industry from an innovation studies perspective (see [8–11]). Such studies stress how batteries represent a shift away from carbon-intensive technologies based on non-renewables (see also [12]) and symbiotic with post-industrial products, infrastructures and macro-societal models (see [13,14]). Indeed, this emerging patent-based literature has so far mostly dealt with the analysis of one or few batteries defined from a conventional electrochemical innovation perspective. In this paper, we stretch this line of work by providing a broad and long-run appreciation of secondary battery innovation while considering more explicitly how their technological content facilitates a deep transition toward circularity characteristics. In fact, batteries not only contribute to limiting CO2 emissions from fossil fuels, they also have systemically transformative effects. Whereas primary batteries are one-off assets, secondary batteries are rechargeable, i.e., these technologies are therefore intrinsically more pro-circular (vis-a-vis primary ones) since they have a longer and more flexible working life-cycle (the energy services extracted per kilogram of employed material are overwhelmingly superior). Thus,

* Corresponding author.
E-mail address: bdamasio@novaims.pt (B. Damásio).

https://doi.org/10.1016/j.renene.2023.03.132
Received 13 September 2022; Received in revised form 21 March 2023; Accepted 30 March 2023
Available online 8 April 2023

the contribution of (secondary) batteries to closing loops and building a Circular Economy is paramount [15,16]. On the one hand, the progressive replacement of single-use batteries for rechargeable ones reduces materials consumption. On the other hand, more efficient and effective storage capabilities facilitate the progressive mainstreaming of carbonless power while opening the scope for new business models and inducing investment in new infrastructure.

Batteries are, indeed, unfinished business. The introduction of the lithium-ion battery represented a world-changing discontinuity, since its affordability and flexibility plus energy density and reliability enabled a wave of new products and equipment, from smartphones to wearable devices, from smart sensors to electric vehicles (see, e.g., [17]). Furthermore, continuous change and structural variation mean that other transformative impacts are possible. Newer generations of batteries that have characteristics such as rechargeability, higher energy-intensity, longer lifespans, that take up more environmentaly-friendly elements from nature and that reduce/avoid the use of environmentally hazardous materials (like present in the conventional nickel-cadmium and lead–acid technologies) are understood here as facilitating circularity. Moreover, the diversity of technical development pathways also in itself matters from a circular directionality point of view since it dilute the pressure on the narrow pools of scarce minerals needed to engineer batteries and their components. Innovations that represent departures from the technological conventions, for instance by highlighting reuse and repair features, do enhance sustainability in more meaningful ways as they are exemplary of headway heuristics of the shifting knowledge base toward "deep transition" and a "circular economy" (see [18–24]). If batteries are all too often assumed as being part of green solutions, we stress that considering their own circularity is a crucial dimension as "whole-of-system" approaches are developed. Our study provides a way to inquire how relevant batteries are for the Circular Economy approach.

For the present work, we built a new dataset containing 92,700 secondary battery patents (consolidated in terms of *international patent families*, or IPFs) from 2000 to 2019. The raw data was extracted from PATSTAT Online (edition: Autumn 2021), the web interface of the PATSTAT database maintained by the EPO containing a vast collection of data extracted from worldwide patent documents and which is usable for purposes of statistical analysis (see [25]). In the past decades, patents emerged as crucial data for evaluating technical progress [26], including for tackling pressing global challenges (see [27]). Albeit a gush of recent work using patents in connection with energy storage for particular technologies (e.g. [28–31]), patents remain under-exploited for conducting integrative mapping exercises of battery development, i.e. across types, geographies and long stretches of time (some exceptions being [7,14,32]). This paper provides a systematic analysis of patent big data (large period, global scope, all battery domains), but is also distinguished from extant contributions by providing an appraisal of patent textual content from which novel insights regarding "circularity" are derived (for background see [33–36]). In doing so, this paper extends battery patent analysis to the circularity realm by providing a first account of how "circular" these trends have been. In particular, we propose textual patent data as a suitable means for appraising the degree of circularity in new battery advances. For the identification of inventions with circular characteristics, we propose a novel, albeit simple, approach that draws on conventional definitions of Circular Economy (with the emphasis on re-use, repair, recycle, recover. etc.; see, e.g., [37]) in the textual content of patent documents.

We find that global battery patenting activity grew significantly in the 2000–2019 period. This stylized fact means that the comparative advantages of secondary approaches (rechargeable, redeployable, reusable batteries) have been continuously on the rise driven by innovation, making a direct contribution to socio-technical circularity. We also confirm that the majority of battery patents originate mostly from Far East manufacturers, but also show that several Asian and European countries exhibit high battery patent per capita intensities. Four battery technologies (redox-flow, solid-state, sodium-ion, and lithium–sulfur batteries) display increased patenting dynamics from 2000–2009 to 2010–2019, a pattern that can serve to cluster countries in terms of performance on emerging battery types (from which inferences can be made regarding the potential to contribute to circularity in the future). We find that several battery-related technologies and applications, such as energy storage systems, battery management systems, wireless power transmission, electric vehicle charging, and uncrewed aerial vehicles (i.e., drones), grew in relevance both in absolute terms and relative to general battery patenting activity. These results complete and bolster current knowledge regarding the pathways of battery innovation that have been surfacing of late and attracting policy attention [7]. The connections of battery innovation with pro-circular transformations may be non-linear (for instance, batteries are of course intensive in exhaustible mineral resources), but overall we find evidence of trajectories of technical change that are less-toxicity intensive, more diverse in the materials employed and more exploratory in the direction of technologies with greater pro-circular potential. We observe non-trivial activity in the overlap of batteries and the circularity realms, especially after 2010, mostly related to reuse and repair features. In this way, our contribution adds to the still small, but expanding, stock of patent-based scholarly work and grey literature on battery evolution.

This paper is organized as follows: Section 2 refers to battery technology and the theoretical light in which we study them. Section 3 describes the method and empirical materials. In Section 4 the results are presented. These outcomes are discussed in Section 5. Section 6 concludes the article. Detailed descriptions of the data selection process and the methods deployed for this analysis are provided in Appendix A.

2. Batteries in innovation studies

We approach batteries not simply as a stand-alone "device" but as a technological system that is based on a multi-domain, evolving knowledge base. This section sets forth how we understand our subject matter, namely, innovation and the battery technology itself.

2.1. The empirical study of industrial innovation

Innovation is the process through which ideas and knowledge are converted into useful applications. This means that innovation is a multi-phased process, open to feedback at every stage, molded in an ongoing fashion by a variety of players and institutional settings [38,39]. Indeed, progress is seldom uni-linear. As it it well known when evolutionary processes are concerned, the sustained dynamics of change is characterized by openness, multiple learning paths and structural unfolding of diverse exploration avenues [40,41]. In the neo-Schumpeterian tradition, technology is seen as a body of useful knowledge that can, at an analytical level, be statistically measured [42] and has, at a substantive level, systemic properties that can be related to transformative transformations, such as the transition to the Circular Economy [34]. Indeed, In the face of climate neutrality targets, "being innovative in order to be circular" is emphasized as a policy pathway for sustainable industrial development ([43], p. 303).

As innovation started to be regarded as an empirical phenomenon of significant importance, its measurement became an increasingly topical agenda. Quantification of an intrinsically qualitative process is, nevertheless, a difficult and delicate task. Any approach is a partial approach since innovation is a multifaceted phenomenon. But empirical research is analytically desirable in order to understand technological change over time, along space, and across challenges [27]. Plus, empirical innovation studies are instrumental in assisting managerial strategy and public policy [44], especially when critical technologies or radical innovation is at stake [45].

2.2. Secondary batteries

Secondary batteries are able to receive energy in the form of electricity, store it, and at a later time (and with a certain loss due to the energy conversion processes taking place) release it again, feeding electricity back to the grid or powering a given application. Secondary batteries are rechargeable, unlike primary batteries which can only discharge once and then need to be discarded. In the context of the ongoing energy transition (a move away from dispatchable sources such as coal-fired power plants and towards alternatives such as wind and solar, in which input is not controllable), batteries and other means of energy storage constitute a regulating bridge that conjoins the temporal gap between supply and demand while balancing the system as a whole. Moreover, accelerated electrification in the transporting sector, especially in individual mobility, creates a focusing device calling out for more batteries and longer lifespans. What is more, now in the stationary domain, the emphasis on resilience and energy autonomy has only reinforced the role of batteries as backup power, in a combination with inherently variable sources like solar and wind [46,47], see also [48]. As with any other critical technology, batteries have systemic and non-linear impacts [49,50].

When referring to batteries, one has to differentiate between the terms “battery”, “module”, and “cell”. While an entire battery pack potentially consists of multiple modules that are “wired in series and/or (less often) parallel” a module itself consists of multiple cells that “are connected in series or parallel” [51, p. 345]. For simplicity’s sake, secondary batteries, meaning battery packs in their entirety, will hereafter be simply referred to as “batteries”.

2.3. How batteries differ

There is a plethora of battery technologies that differ in several aspects, namely the type of electrodes and electrolytes, their format, applications and in some cases even the working principle is different. This subsection does not attempt to exhaust the full range of existing technologies, but rather to briefly describe the main varieties (the groups of technologies) that are prominent in our analysis.

Lithium-ion (Li-ion) battery is a rechargeable battery that charges and discharges energy through the movement of lithium ions between the negative electrode (anode) and the positive electrode (cathode) [52]. The transport of ions between electrodes occurs through an electrolyte, and a separator is placed between the two electrodes to avoid direct contact between them [53]. Although there are several types of Li-ion batteries, the core material of which is mining-intensive, the use of transition metals such as cobalt and nickel also pose serious environmental, social, and even geopolitical issues that motivate the quest to replace them [54,55].

Solid-state batteries (SsB) are batteries in which the liquid electrolyte is replaced by a solid-state one. Although there a several examples of non-lithium SsB, most of the research is done in the context of lithium-ion technologies. One of the major advantages of solid-state Li-ion technologies, when compared to conventional ones, is that they avoid possible leaks of the liquid electrolyte. Another problem that can be avoided with solid-state electrolytes is the formation of dendrites of lithium which can cause the battery to explode [56]. The main drawback of solid-state electrolytes is that at cool and average temperatures solid oxides have a high resistance to ionic conductivity, making them unsuitable to be used at low and room temperatures. Also, the stress created at the electrode–electrolyte interface at room temperature tends to reduce the battery lifespan [56]. Thus, although SsB theoretically have a higher life expectancy [53], presently they cannot attain the durability of conventional Li-ion batteries [57].

Lead–acid batteries (Pb-acid) batteries were the first rechargeable batteries ever produced. The original Pb-acid battery was composed of two lead electrodes immersed in a sulfuric acid electrolyte [58]. Although there have been significant advances since, such as the Valve Regulated Lead Acid (VRLA) battery [58], the working principle of Pb-acid remains the same. Pb-acid batteries use inexpensive materials, are easy to produce and the technology has a high maturity level, which makes this technology cost-competitive. Pb-acid batteries are widely used as motor starter batteries in combustion engine vehicles, they are also used on off-grid energy systems [59]. The main drawbacks of Pb-acid technologies are their height, short lifecycle, and the use of lead which is toxic and constitutes an environmental problem. On the other hand, recycling for Pb-acid batteries is well established and very high lead recycling rates are achieved [59].

Lithium–sulfur (Li–S) batteries hold the promise to achieve very high energy densities (i.e., beyond 500 Wh/kg), which makes them particularly suited for mobile applications [60]. Also, the use of sulfur as cathode material, which is very abundant and environmentally friendly, makes this type of battery quite attractive [61]. Still, the development of Li–S technologies faces some significant hurdles. First, both sulfur and the discharge product (Li2S) are electronic/ionic insulating thereby hindering charge transport. Second, very large volume changes (up to 80%) during charge/discharge cycling accelerate cathode degradation. Third, lithium polysulfide intermediates dissolve in the electrolyte and shuttle between the cathode and the anode reducing the charge transfer efficiency (Coulomb efficiency) and cycling stability [61,62].

Unlike conventional electrochemical batteries where energy is stored in electrodes, in redox flow batteries (RFBs) energy is stored in the electrolytes. In the RFBs the charge/discharge processes are based on reversible electrochemical reactions of two redox couples that are dissolved in electrolytes. RFBs have two parts that are connected through pumps: the battery stack, where electrochemical reactions occur, and the external tanks, where the electrolytes are stored. The battery stack includes two sets of electrodes, bipolar plates, and current collectors that close a membrane between two electrodes. The membrane conducts the charge carriers and avoid the mix of the two electrolytes [63]. Since the total energy stored is determined by the electrolyte concentration and volume, and the power is determined by the current density and electrode area, the RFBs energy can be sized independently from its power, allowing it to adjust the energy stored by increasing the volume of the electrolytes. This flexibility makes RFBs particularly suited for grid-storage applications. Also, these batteries have a long lifespan, high energy efficiency, and allow low cost for large-scale energy storage [64]. Vanadium redox flow battery is so far the most successful of RFBs because, besides the advantages already mentioned, these batteries benefit from the use of abundant and environmentally friendly electrolytes. The major drawbacks of these batteries are their limited energy density and operating voltage [65].

Sodium-ion (Na-ion) batteries have been proposed as an alternative to Li-ion batteries. Like Lithium, Sodium belongs to the group of alkaline metals, which means that its chemical behavior is in several aspects very similar to lithium, notably its reactivity with water. Due to this similarity, Na-ion and Li-ion batteries are considered sister systems [66], and Na-ion technologies tend to mimic Li-ion chemistry which as favored them in terms of a faster development [67]. One of the main advantages of Na-ion batteries is the fact that sodium is much more abundant (the fourth most abundant element on Earth’s crust) and thus less expensive than lithium [68]. Conversely, the chemical reactivity of sodium with water is higher than that of lithium, which inhibits the use of metallic sodium in the anode. Research in this area is very active and there is not a defined chemistry for the sodium-ion battery, as a lot of different electrodes and electrolytes are being tested [67].

3. Batteries and patents data

The empirical materials for our study are addressed in this section. Intellectual property data on inventions can be, and have been, used to analyze battery development. Whilst they remain partial and imperfect indicators, they remain useful but somewhat underutilized.

3.1. Patents as an innovation indicator

Patents are intellectual property rights on inventions. A patent describes claims to useful ideals and assigns rights to new knowledge. As legal documents patents represent a trade-off. They ascribe ownership but also reveal as wealth of information related to actors, places, dates, etc. In particular, patents disclose data on geographic locations associated with inventors, descriptions and classifications of the respective inventions, and timestamps related to filling and publication dates. This allows for the aggregation of patent counts alongside geographic, temporal, and technological dimensions and makes them a suitable material for a myriad of analytical purposes, from competitiveness studies to sustainability research [26].

Patents are, thus, viewed as resource for capturing the notion of technical change. Patents grant formal protection for an idea that is (1) novel, (2) showing an inventive step, and (3) capable of industrial application [69]. Typically, interested parties (inventors, owners, intellectual property lawyers, patent offices, etc.) apply for formal protection before the ideas are operationally tested and before getting feedback from their commercial roll-out. Surely not all inventions are patented, and the value of other developments or improvements can be appropriated by other means which in turn can be detected and measured (a case in point being trademarks and the digital economy, see [70–73]). Hence, despite only yielding partial and imperfect evidence of innovation, patents are irreplaceable in the toolbox of innovation economists and business analysts [26]. When making a case for patents as a proxy for measuring innovation, Zvi Griliches classically explained that patents “are available; they are by definition related to inventiveness, and they are based on what appears to be an objective and only slowly changing standard” [74, p. 1661]. They also have well-known limitations: there are different propensities to patent across technology areas, their economic value widely varies, service innovations are not captured, etc. More recently, new methodologies have stretched the empirical usefulness of patents [27]. For instance, patents have been repurposed to unveil new insights with regard to pressing global challenges such as environmental progress, human well-being and climate change adaptation (see, e.g., [75,76]).

Recently, patents have been increasingly mobilized to track developments in green innovation, including in strategic emerging sectors like clean technology and renewable energy [12,77]. It is well known that data beyond patent number is of interest: for instance, recent methodological developments have been achieved to extract further information from patents by using patent citation and also internal patent document content [27]. Although it can be seen as a fundamental direction in a broader pro-sustainability transformation, the literature that can be found drawing on battery patents is still emergent. The following subsection briefly reviews it.

3.2. Extant battery patent analysis

A number of energy-related patent-based empirical works have underscored how understanding technological potential can inform eco-innovation promotion and climate change mitigation strategies, including public policy and corporate/start-up development efforts [31]. Recently, a few of these studies have begun to examine the dynamics of innovation in the “world-changing” field of secondary batteries [17]. These have covered especially the lithium-ion variety, which is the dominant solution for today’s informational lifestyle (mobile phones, tablets, laptops; see [8,78]; see also [79,80]).

The scholarly research stream on battery patents is growing. Some research focused on patent counts for just one type of technology for a limited number of countries, namely lithium-ion for the leading countries in the field (e.g. [8,31]. Other studies have moved forward with the empirical strategy, for instance, by proposing a citation network analysis combining knowledge extracted from patent data with results from interviews conducted with lithium-ion battery experts [81]. Stephan et al. [29] examined lithium-ion battery patents from a sectoral diversity perspective and emphasized how the distance from prior knowledge affects certain features of subsequent knowledge (see also [30]). Kittner et al. [82] and Ziegler and Trancik [83] employed the patent proxy in their efforts to model the forces driving the prices of lithium-ion batteries, and found that cumulative patent filings is the best predictor of real prices scaled by energy capacity. Work on alternative chemical alternatives to lithium-ion has been even rare (see [10,13]).

Our contribution complements the still scant, but growing scholarly work on battery evolution. It also extends the existing grey literature on this matter. Specifically, it aims to confirm and consolidate the findings presented in the IEA and EPO report [7] and it can be thus understood as a continuation of their basic methodological approach, enriched by some reasonable additions, which allow for a more granular perspective on some aspects. However, our work also seeks to provide a more encompassing picture of a very vibrant area, including by drilling down for content and uncovering within-text patterns.

The IEA and EPO report presents patent trends related to batteries and electricity storage. In contrast, our own study is more focused (looks at battery technology only) but has a longer time span. The research gaps that we identified and which the current study aims to fill are how patent counts are distributed across continents, how scaling them by the sizes of the respective labor forces affects the outcome of the analysis, what their distribution across another technological classification scheme looks like, how countries can be characterized based on their position in technology space, and what information can be extracted from patent abstracts. What is more, we are able to build bring new perspective with regard to circular directionalities.

3.3. Data acquisition procedures and empirical categories

The raw bulk data used for this study were accessed via subscription at PATSTAT, the online worldwide reference patent repository harboured by EPO. The source is organized according to the International Patent Classification (IPC) scheme. The IPC provides a hierarchical classification scheme that categorizes patents according to different technological areas.

Our extraction strategy for deriving our data subset is described in the detail in the Appendix, and the queries (Transact-SQL) and code (Python) needed to replicate this study are also made available. On the basis of substantive knowledge of the technology (namely the reference EIA and EPO report, but also the recent scholarly battery patent literature) the search was conducted iteratively, with time and care so as to arrive to a robust final dataset. It is on this final dataset that we compute occurrence counts, including when we run content searches for an array of strings on all English titles and abstracts.

This study builds on battery patents that can roughly be characterized in the following way: (1) inventions related to the casing, wrapping, or covering, i.e., non-active parts of batteries; (2) developments in battery electrode manufacturing; (3) innovations related to the manufacturing process of secondary cells; and (4) advances related to charging of batteries. Patents belonging to these four fields were identified using the international patent classification system (IPC). The IPC provides a hierarchical classification scheme that categorizes patents according to different technological areas. While several specific analytical options and constraints are discussed in the analytical section of this paper, the complete details regarding data acquisition and processing are supplied in Appendix A.

In this study, we use the concept of international patent families (IPF). A relevant patent application is a formal request made by one or several applicants at any given patent office of their choice for a unique invention. These could be the European Patent Office (EPO), the United States Patent and Trademark Office (USPTO), or any other national or regional patent office. The IEA and EPO report uses IPFs for aggregating and counting patent applications. They claim that an IPF "is a reliable proxy for inventive activity because it provides a degree of control for patent quality by only representing inventions for which the inventor considers the value sufficient to seek protection internationally" ([7, p. 4]).

The term *patent family* refers to the whole set of patent applications covering the same invention [84]. By counting patent families instead of individual applications, double-counting of inventions is avoided. By restricting the scope of the search protocol to only patent families that contain an international patent application, at least one application to a regional patent office, or applications to at least two distinct national patent offices, one obtains IPFs. One benefit of this restriction is that only patents of higher expected value are assessed, resulting in a more homogeneous dataset with better comparability between elements. In this study we use the same criteria to identify IPFs that the IEA and EPO report used. The regional patent offices are the African Intellectual Property Organization, the African Regional Intellectual Property Organization, the Eurasian Patent Organization, the EPO and the Patent Office of the Cooperation Council for the Arab States of the Gulf.

A drawback of IPFs is that several different definitions are used in patent studies. Moreover, as Schmoch and Gehrke [85] discussed, three limitations regarding the IPF concept itself should be considered: First, the propensity to patent in foreign territories differs between countries of origin, meaning that, for example, an applicant from a European country might be more inclined to seek protection in another European country than an applicant from China might be inclined to seek protection in the US. This can be problematic because both situations would imply that the respective patent is filed in two countries, thus making their patent family an international patent family. Second, patent numbers for some countries in specific technologies, such as Japan in microelectronics, may be overestimated. Third, there can be some turbulence in the evidence since IPFs with seemingly two members at the stage of applications can be reduced to one member, later on, something that may happen with Chinese inventors (regarding the Chinese case, we further refer to Frietsch and Kroll [86]). Schmoch and Gehrke [85] discuss several other concepts that exist parallel to IPFs, highlighting their advantages and limitations.

To ensure comparability with the recent IEA and EPO report, we have kept IPFs as our frame; therefore, all depicted counts refer to IPFs. However, there are some discrepancies between their study and our own; this is something that we are not able to fully account for but works as a stimulus for future research which serves as further attempts to validate the findings of a prior analysis. The comparison between these two studies is not direct because our numbers depict "Lithium-ion" and "Other lithium" separately, because the IEA and EPO report uses another classification system (the Cooperative Patent Classification (CPC)), and because we decided to include charging technologies. Notwithstanding, it is reassuring to note that both studies detect a step-jump around the year 2010 and that the counts are very correlated (ours and their counts yield a Pearson correlation coefficient of 0.9940 (rounded to the fourth decimal place); see Appendix C).

4. Results

In this section aggregate data is used to highlight the major patterns concerning battery progress. Desegregated data is then examined to show how patents reveal more specific information, regarding different, technologies, and connections to circularity.

4.1. Basic stylized facts

The global aggregate yearly volume of battery IPFs increased almost every year during the time frame assessed in this study. There were slight decreases only for two pairs of adjacent years: from 2001 to 2002 and from 2014 to 2015. The whole time period's average yearly growth rate in battery IPFs is 14.3% so between 2000 and 2019 the total IPF output increased more than 11-fold. This dynamic is displayed in Fig. 1.

Asian countries dominate the battery scene: the Asian continent's mean annual battery IPF output is approximately four times higher than Europe's and North America's (a factor of 3.57 and 4.10, respectively). Furthermore, the number of IPFs from Asia increased by 15.96% on average every year during the 2000–2019 period. The average increase for Europe and North America was 13.46% and 10.80%, respectively (see Fig. 2; log-scaled y-axis).

Breaking down battery IPF counts by inventors' countries of origin, the dominance of Asia becomes even more apparent. Fig. 3 shows the eight countries with the highest total battery IPF output over the whole timespan. By 2019 the three top countries in terms of battery IPF output were from the far east: Japan, South Korea, and China. These were followed by the US, Germany, France, Taiwan, and the UK. Japan, the undisputed leader in battery IPF counts during the whole time frame, has been displaying a vibrant rate in the dynamics of inventive output since 2016. China is catching up fast with South Korea, which has held second place in battery IPF output since 2011 when it surpassed the US (for the Chinese case see [87]). Germany also displays growth in battery IPF output. These results echo those of the IEA and EPO report ([7], Figs. 6.2 and 6.3).

By scaling the numbers shown in the previous plot by each country and year's labor force count, one obtains battery IPF intensities [88]. This measure gives perspective on performance, allowing for the assessment of a country's innovative output relative to the size of its working population. Fig. 4 shows the eight countries with the highest scaled total battery IPF output over the whole period and it can be seen that in contrast to Fig. 3, some small European countries are stepping up: Austria, Finland, Switzerland, and Sweden are part of the top eight. It is also worth noting that, in this light, South Korea overtook Japan in 2014, establishing itself as the global leader in terms of battery patent intensities.

4.2. Battery technologies

By assigning battery technology sub-areas to patent families a decomposition of the dataset into 19 battery cell technologies was obtained (detailed description in the Appendix A.2). Fig. 5 presents the developments of IPF counts in the eight major technological categories, selected on the basis of their total IPF count in the entire time frame of 2000–2019. The depicted battery IPF fractional counts are rounded to the closest integer and the eight technologies with the highest total battery IPF count over the given time frame are displayed in descending order.

While the number of IPFs related to lead–acid batteries (i.e. arguably the least circular of the technological options) has been relatively stable over the depicted 20 years, which resulted in its overall share in battery IPFs decreasing steadily over this time period, and while rechargeable alkaline batteries exhibit a slight downwards trend, lithium-ion batteries and other lithium-based battery technologies have soared drastically. Less relevant today than lithium-ion batteries, but with considerably higher counts than other smaller battery technologies, are the four remaining categories presented in Fig. 5: patenting activity related to lithium–sulfur, solid-state, sodium-ion, and redox-flow batteries have seen a notable increase in IPF counts in 2010–2019. In 2019 solid-state batteries reached an all-time maximum.

As previously mentioned, solid-state batteries are a specific configuration mostly implemented in the framework of lithium-ion solutions. In that sense, one might assert that the emergent redox-flow, lithium–sulfur, and sodium-ion technologies provide a substantial contribution to technological heterogeneity and can lead to higher diversification of the materials used in battery manufacture thus avoiding the over-exploitation of scarce resources available in nature such as those already extensively used in the dominant lithium-ion technologies (like lithium, nikel and cobalt). In this sense, the increase in technology

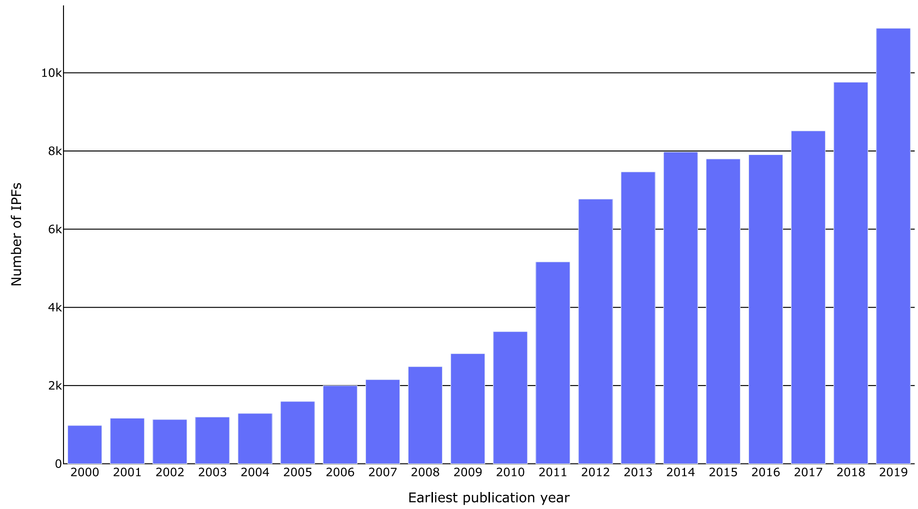

Fig. 1. Total number of battery IPFs, 2000–2019.

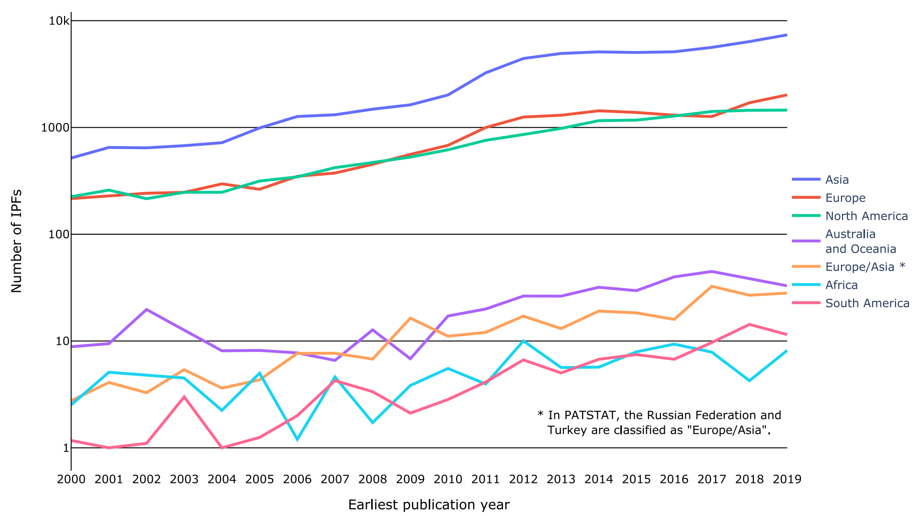

Fig. 2. Battery IPFs by inventors' continents of origin, 2000–2019. Note: The y-axis is log-scaled and all values are incremented by 1. It is clear that the number of battery IPFs from Asia (blue) is considerably higher than that of any other continent.

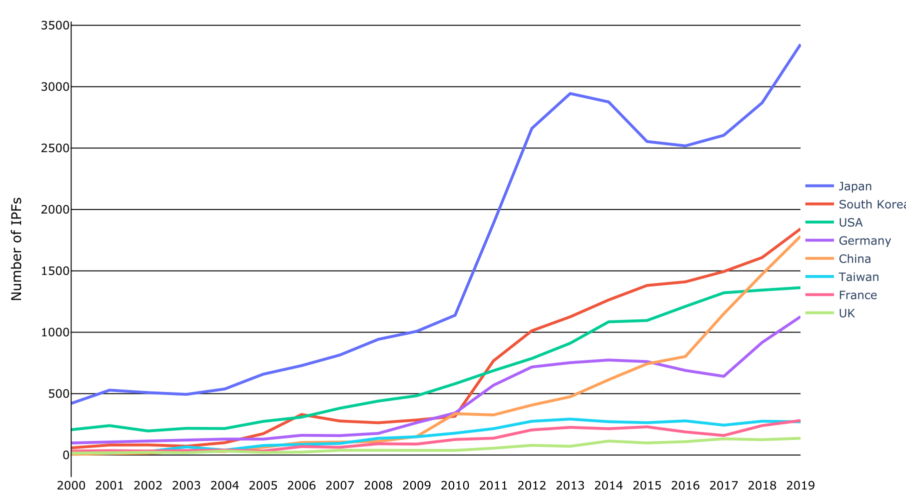

Fig. 3. Battery IPFs by inventors' countries of origin, 2000–2019. Note: The eight countries with the highest total battery IPF counts over the given timeframe are displayed. Japan (blue) has the highest battery IPF output in the given timeframe, whilst other countries' IPF counts (especially South Korea's (red) and China's (orange)) have been surging in the recent decade.

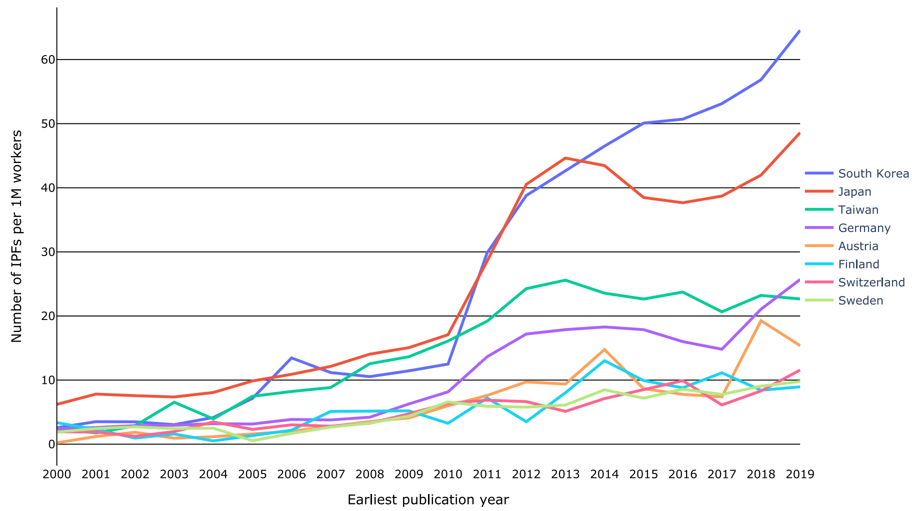

Fig. 4. Battery IPFs per 1M workers by inventors' countries of origin, 2000–2019. Note: The eight countries with the highest total battery IPF intensities over the given timeframe are displayed. In this perspective, South Korea (blue) overtook Japan (red) in 2014.

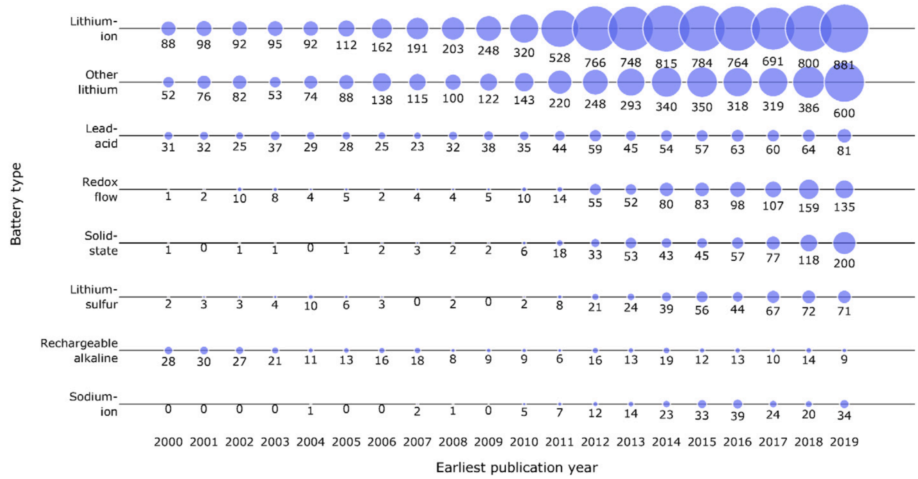

Fig. 5. Global battery patenting activity for the major battery types, 2000–2019. Sorted in descending order by total IPF count.

diversity promoted by innovation has the potential to promote the overall circularity of battery development.

The observation that the recent decade displayed increased patenting activity in these four emerging technologies motivates the way the next part of the analysis is set up: The following subsection describes the results obtained by clustering countries based on their position in a technology space computed using their technology distribution of the years of 2010–2019 (6).

4.3. Country clusters

The most suitable technology realm for clustering was found to be spanned by the countries' distribution values over the four emerging technologies lithium–sulfur, solid-state, sodium-ion, and redox-flow, which display increased patenting activity after 2010, alongside the older lead–acid technology. In attempting to cluster 36 countries using data from 2010 to 2019, k-means was found to be the algorithm with a better R^2 value for all relevant numbers of clusters (for details on this metric see A.5).

Setting the numbers of clusters to two, we obtained a clear separation of the dataset between countries with a high focus on lead–acid batteries (81.91% of IPFs are related to lead–acid batteries in this cluster) and countries with comparatively high shares of IPFs related to the four emerging technologies and consequently a relatively low share of lead–acid related IPFs (19.55%).

Setting the number of clusters to three in order to achieve a more granular separation we observe the following pattern. While countries from cluster 1 are more focused on lead–acid batteries, clusters 2 and 3 exhibit a higher patenting activity related to the four emerging technologies of redox-flow and solid-state batteries (cluster 2) and lithium–sulfur and sodium-ion batteries (cluster 3).

In comparing these results with a two-cluster scenario, one finds that the lead–acid focused cluster from the previous stage is still fairly intact, while the "emerging technologies" cluster has been separated into two. This division results in one country cluster displaying a stronger focus on redox-flow and solid-state batteries and another exhibiting a higher relative focus on sodium-ion and lithium–sulfur-related IPFs. Fig. 6 shows the distribution profiles of the three-clusters solution generated with the k-means variable "random_state" set to zero. The variable "random_state" determines the centroid initialization of k-means and results in deterministic runs of the algorithm when a value is assigned to it.

While the approximate shape of the clustering profile depicted in Fig. 6 is fairly insensitive to alterations or non-assignation of "random_state", the affiliation of the countries to their clusters varied enough to motivate running k-means a higher number of times (with

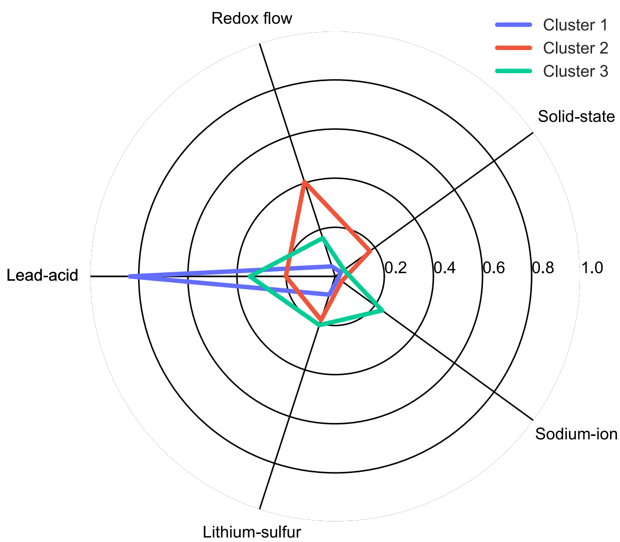

Fig. 6. Cluster of inventors' countries of origin,2010–2019.

the variable "random_state" undefined) to compute each country's cluster affiliation distribution for assessing which cluster each country belongs to in the majority of events. Running k-means 10,000 times resulted in the following most probable cluster affiliations:

- Cluster 1 (16 countries): India, Turkey, Russia, Bulgaria, New Zealand, Luxembourg, Poland, Sweden, Mexico, Malta, North Korea, Serbia, Greece, Hungary, Kazakhstan, Israel.
- Cluster 2 (13 countries): USA, Germany, Taiwan, Austria, Netherlands, Thailand, Switzerland, South Korea, Japan, Belgium, Italy, Australia, Hong Kong.
- Cluster 3 (7 countries): Canada, Spain, Ukraine, UK, France, Norway, China.

Inside each cluster, countries are ordered by (1) their probability p to be in this cluster, and (2) their total IPF count in the five categories. Each country's name is colored according to the following schema, indicating its probability p to belong to the respective cluster:

$p = 1$ $p \in [0.99,\ 1)$ $p \in [0.9,\ 0.99)$ $p \in [0.5,\ 0.9)$

A value of $p = 1$ indicates that a country was assigned to this cluster during each of the 10,000 runs, meaning that its cluster affiliation appears to be quite insensitive to the algorithm's centroid initialization.

In terms of circularity, in line with what was mentioned in the previous section, we can assert that due to their higher technological diversity countries in clusters 2 and 3 have the potential to provide a higher contribution to a more Circular Economy than cluster 1, which is mainly focused on lead–acid technologies. In comparing cluster 2 and 3, it stands out that cluster 2, while having a strong emphasis on solid-state batteries (which as mentioned is essentially a particular type of lithium-ion battery), is mainly focused on two emerging technologies (redox-flow and lithium–sulfur). In contrast, cluster 3 reveals robust innovation activity in three emerging technologies outside the lithium-ion technologies framework (i.e., redox-flow, lithium–sulfur, and sodium-ion), suggesting that countries driving cluster 3 could have a higher potential to contribute to circularity in the future since it is more diversified in its exploration of future alternatives.

4.4. Patent title and abstract mining

The content material of patents is relevant evidence that can be mined, processed, and sorted to leverage classic patent analysis [36, 89]. The top 50 trigrams in terms of their intensity increase between 2000 and 2019 are displayed in Fig. 7. The terms are displayed in descending order of total increase over the given 20-year time period.

The method that was implemented to analyze patent wordage was as follows. Both patent abstracts and titles were searched for meaningful phrases. Besides simply counting occurrences of n-grams for each year (analysis not shown), the approach we refer to as *n-gram intensities*. Counts are scaled by the respective year's number of abstracts (results are similar for titles) and the color gradients represent intra-row intensities. The resulting unit of measure for n-gram intensities is occurrences per 1,000 abstracts; and all depicted n-gram intensities are rounded to the closest integer. Thus, each cell displayed in Fig. 7 is the respective occurrence count thus corrected by the size of the corpus. It should be noted that some patent families do not have a non-NaN English abstract, that is, the number of abstracts associated with a given year can be lower than the number of IPFs associated with that year. For purposes of sensitivity analysis, unigrams (single words), bigrams (strings containing two words), and trigrams (arrays with three words) were extracted and processed. The resulting n-gram counts and n-gram intensities were sorted in three different ways, which are described in detail in Appendix A.6 (Appendix A). The results that we found most meaningful and thus selected for presentation in this paper were indeed the top 50 increasing trigrams extracted from battery patent abstracts. An appreciation of the results is provided considering all the different angles that were implemented (but not shown here).

Trigram counts display several expectable trends like the surge of "lithium secondary battery" and "lithium ion battery". The occurrence counts for these two trigrams increased from 46 to 844 and from 15 to 685, respectively, between 2000 and 2019 and the trigram intensities of "lithium ion battery" indicate a robust upward dynamic not only in absolute terms but also relative to battery patenting activity. The increase of the term "energy storage system", which is also confirmed by its intensity's trajectory, hints at an upsurge in the importance of increasingly complex systems for managing energy storage. This is buttressed by the term "battery management system", also occurring in both counts' (not shown here) and intensities' top 50 trigrams (Fig. 7). As already established by Fig. 5, solid-state batteries have been growing in relevance, especially in the past decade. The increasing counts and intensities for the terms "solid electrolyte layer" and "solid state battery" after 2010 confirm this. Notable trigrams in the subfields of battery charging and electric vehicles are "wireless power transmission" and "electric vehicle charging", which have both increased considerably in both counts and intensities. The surge in relevance for redox-flow batteries (see Fig. 5) is also confirmed by both counts and intensities ("redox-flow battery"). The trigrams "plurality battery cell" (results from "plurality of battery cells" due to stop word removal and lemmatization) and "battery module plurality" (both present in counts and intensities) hint at a substantial increase in innovative output related to compositions of cells and modules inside battery packs. An interesting appearance in the top 50 trigram intensities is the term "unmanned aerial vehicle", exhibiting 4, 13, 16, and 8 occurrences per 1000 abstracts in 2016, 2017, 2018, and 2019, respectively as it indicates an increased field of application related to the deployment of battery technology in drones.

The connection to circularity is not straightforward at first sight but can be elaborated upon. The relevance of innovation in batteries for appraising the transition to a Circular Economy can be further discussed by analyzing these text materials. Indeed, 15 of 50 trigrams with more significant growth in the period 2000–2019 have references to secondary, rechargeable or storage. That is, the technological paradigm is not about primary (less reusable, less enduring, less re-deployable approaches). Moreover, these same (pro-circularity) descriptors appear

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Electrode active material	91	70	64	102	106	100	123	129	119	127	126	192	185	173	167	179	163	155	143	181
Active material layer	17	8	26	23	29	61	58	67	54	67	55	71	62	74	63	50	43	60	64	80
Energy storage device	5	35	24	33	26	13	23	33	44	33	44	36	33	49	46	36	47	64	60	59
Lithium ion battery	15	22	36	26	12	31	30	39	40	43	58	62	70	56	57	67	79	67	73	62
Electrode current collector	7	6	9	10	12	10	12	21	23	37	28	26	30	23	20	22	27	27	43	49
Lithium secondary battery	47	78	64	55	87	82	98	65	50	63	66	66	67	65	62	53	63	44	56	77
Plurality battery cell	3	1	3	1	5	6	5	13	10	10	9	26	26	31	22	27	27	27	32	32
Power storage device	11	3	6	8	2	5	2	14	28	13	54	32	47	42	45	40	40	62	30	36
Current collector electrode	5	7	5	3	6	7	6	13	15	16	17	25	17	15	17	13	17	16	22	29
Secondary battery electrode	30	23	19	21	43	52	58	51	38	43	43	56	62	57	48	52	54	52	45	54
Ion secondary battery	10	21	13	19	25	37	35	34	30	29	48	51	49	55	54	61	50	40	31	32
Power supply device	11	3	22	6	16	18	18	13	34	21	24	27	34	35	36	29	29	30	29	32
Energy storage system	0	6	0	7	2	7	3	12	14	11	14	18	12	16	15	14	18	20	21	20
Electrode mixture layer	0	0	0	3	0	0	2	7	10	14	8	11	13	12	11	14	19	12	18	20
Lithium ion secondary	12	25	16	25	26	40	37	34	31	30	49	52	54	56	56	65	48	41	30	31
Solid state battery	0	0	0	1	0	0	1	1	2	1	1	2	4	8	6	4	4	6	10	18
Battery management system	0	3	3	1	2	2	6	14	16	7	7	12	13	15	12	19	14	16	17	17
Cathode active material	6	9	20	26	7	18	30	23	18	23	27	15	22	23	25	24	21	23	22	22
Layer electrode active	3	1	2	3	2	3	3	8	5	7	9	12	11	12	11	14	13	12	12	18
Energy storage unit	1	1	19	1	3	4	10	7	10	17	13	16	12	10	6	12	12	12	21	16
Power supply system	7	17	22	15	21	13	16	24	19	21	24	29	28	22	21	19	21	22	23	22
Material layer electrode	2	0	1	3	2	7	8	9	4	11	10	15	12	12	10	11	9	12	9	16
Solid electrolyte layer	5	2	1	0	2	0	2	1	8	4	6	9	7	12	5	9	9	13	13	19
Wireless power transmission	0	0	0	0	0	0	0	0	1	5	7	6	17	14	23	22	20	26	15	13
Redox flow battery	0	0	7	10	2	1	0	0	0	1	1	4	7	4	6	12	11	12	12	13
Collector electrode active	1	2	4	1	2	4	2	7	5	5	8	13	7	12	14	9	9	9	8	14
Electrical energy storage	1	3	9	13	14	2	8	18	6	10	15	6	13	16	9	10	10	12	13	14
Power transmission device	0	0	0	2	0	0	0	0	1	11	8	8	8	12	15	22	12	13	14	12
Transition metal oxide	3	6	3	8	2	6	10	8	5	10	11	10	8	9	11	17	14	9	9	15
Power storage element	1	0	0	0	0	0	0	0	5	4	3	1	7	8	12	14	20	19	18	13
Electrolyte secondary battery	26	23	36	30	50	47	56	55	73	46	33	42	45	36	55	55	53	34	27	37
Active material particle	4	5	11	5	11	12	21	15	13	13	11	11	14	12	12	19	21	25	21	15
Electric vehicle charging	0	0	0	0	0	0	1	0	2	3	7	9	14	15	8	7	5	6	7	11
Battery cell electrode	0	1	4	1	1	1	1	4	1	4	4	5	4	8	5	6	7	6	7	11
Battery module plurality	1	3	0	2	0	3	9	1	6	5	5	10	9	10	8	9	8	8	12	12
Control unit configured	0	0	0	0	0	1	1	1	2	2	3	3	5	5	5	6	6	7	6	10
Non-aqueous electrolyte secondary	25	33	40	39	52	48	56	54	73	48	36	42	45	39	56	55	54	34	26	35
State secondary battery	0	3	2	1	2	2	2	2	4	3	5	4	5	5	4	3	7	11	5	10
Secondary battery lithium	6	13	4	8	6	8	5	6	8	9	12	9	10	9	11	12	14	8	13	16
Anode active material	6	2	4	13	11	13	30	28	26	20	23	16	14	15	27	14	9	14	17	16
Current collector layer	0	0	1	2	2	1	2	0	0	0	1	1	1	5	2	1	2	4	10	9
Unmanned aerial vehicle	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	4	13	16	8
Plurality battery module	1	4	1	6	1	2	4	6	1	4	3	10	7	7	6	4	8	5	11	9
Electrode active substance	1	2	9	0	14	10	3	1	1	2	4	4	7	6	11	7	8	4	7	9
Non-aqueous electrolyte solution	3	8	4	0	9	8	6	6	9	5	9	8	9	8	9	6	7	4	7	11
Solid state secondary	0	0	0	0	0	0	1	1	0	0	1	2	1	2	1	1	6	8	3	8
Present electrode active	0	2	2	2	1	1	1	2	1	2	2	7	5	3	7	7	9	8	7	8
Power receiving device	0	0	0	0	2	3	0	3	1	12	6	8	11	16	22	14	10	10	12	8
Solid electrolyte material	0	0	0	0	0	0	0	0	2	0	2	3	4	5	3	1	2	4	4	8
Power storage system	0	0	1	0	0	1	0	4	0	1	1	4	7	7	10	6	9	4	8	8

Fig. 7. Trigram occurrence intensities in battery patent abstracts.

on average 4.15 times in the top 10 trigrams over the period thus providing suggestive evidence on the pro-circularity of battery innovation trends.

Another sign of transformative innovation emerges from this content analysis. The frequent appearance of references to "hydrogen absorbing" "alloy absorbing", "nickel hydroxide", and "hydrogen storage" at the beginning of the time series (mostly in the years 2000, 2001, and 2002) might be attributed to the innovation effort to find alternatives to nickel-cadmium battery types by replacing the highly toxic cadmium by substitutes based on nickel-metal hydride. In other words, in the early part of the first decade there is evidence on breaking new ground towards cleaner combinations, less toxic materials, and more earth/ocean-friendly solutions.

The trigram analysis overall confirms the prominence of lithium-ion technologies and the nature of the most relevant alternative technological paths. But it also hints at the non-linearity of progress towards safer and more sustainable forms of energy storage. Two undercurrents of technical change are particularly telling in this respect, namely the rising importance of non-aqueous electrolytes and the growing interest in solid-state batteries (both mainly associated to lithium-ion batteries). These trends have a rather complex relationship with the Circular Economy. Non-aqueous electrolytes tend to be made of more toxic materials than aqueous ones [90]. And, as of today, solid-state batteries have shorter lifecycles than conventional lithium-ion batteries. So, at first glance, both trends are going against circularity principles. However, both approaches allow for the increase of the energy density of batteries, a feature that is crucial to improve the performance of electric cars, making them more appealing to users, thus accelerating the transition away from fossil fuel-powered cars to electric ones, thus improving circularity at a systemic level. In other words, it may well be that some micro-heuristics (going for non-aqueous electrolytes and solid-state batteries), which in themselves may be less circular, can have pro-circular effects at a macro-systemic level of analysis. Hence, technology analysis and patent indicators are only a partial and subsidiary approximation to the broader meaning of battery innovation and its links with the evolving socio-technic system.

4.5. The circular dimension of battery innovation

Patents signal the *rate* of progress, but it is clear that they also disclose evidence about the *direction* of change. In fact, the qualitative information encoded in the patent documents is a rich complement to the more conventional kinds of data traditionally used in patent-based studies (date, inventors, technologies, etc.). Our analysis deepens the text-driven approach so far carried out by assessing the extent to which circularity concerns were embedded in the technologies being pushed forward. This is implemented by detecting mentions to content strings that can be associated to the Circular Economy, an exercise that to the best of our knowledge was not tried out in this way.

We review the key characteristics that make up the Circular Economy approach from first principles. A way to start is by the classic three "Rs" of circularity: Reduce, Reuse, and Recycle. Moving beyond sketchy slogans, albeit retaining this "3R" starting point, knowledge on circularity is today underpinned by a variety of work that has explored the concept at length (see, e.g., [4,15,34,91,92]). This literature releases words that can be seen as candidates for circular indicators if they appear in patents.

Our first step was to identify wordage that could point to circularity. These relevant keywords were used to drill down our dataset (starting with the "3Rs" as a starter, see below). Some obvious enough words were tested as candidates, but gave no results ("circular", "circularity"). The keywords were made robust by the consolidation of variations, for instance, "circular" and "circularity", "reuse" and "re-using", "recycle" and "recycling", "lifecycle" and "life cycle", "durable" and

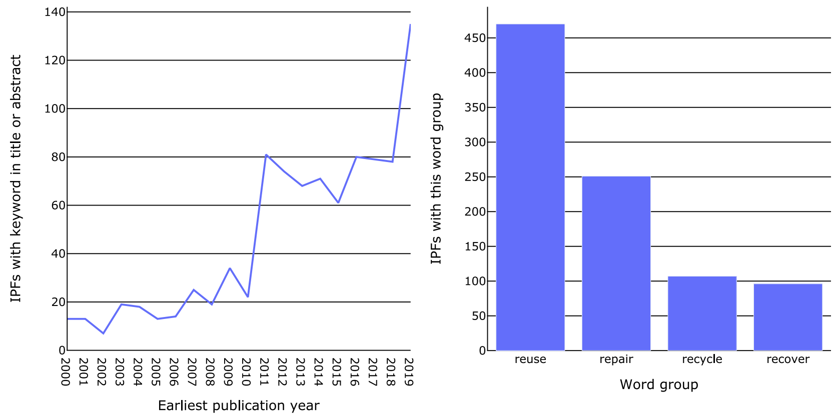

Fig. 8. Occurrence counts of circularity terms in battery patent titles and abstracts. Note: This shows the development of occurrences of IPFs with circular keywords in their titles or abstracts on the left and total occurrence counts of the separate word groups on the right.

"durability", "metabolism" and "metabolic", "upcycle" and "up-cycle", etc. Thus, from the literature we were able to pick the following jargon:

- Specific keywords: "reduce", "reuse", "recycle", "recover", "symbiosis", "urban mining", "waste" and "e-waste", "durable" or "durability", "metabolism" and "metabolic", "crade-to-cradle", "closed loop", "decoupling", "lifecycle", "downcycling", "end-of-life", "upcycle", "extended producer responsability", "technical nutrients".
- General keywords: "circular", "renew", "redesign", "repair".

Our second step was to appraise the returns of the string searches critically. This step is a safeguard against false positives that could surface. While some words gave no results ("circular"), others produced many hits. For instance, the word "reduce" appeared very often raising suspicions of being too undifferentiated. Our technique was to run trigrams to assess the context around the keywords (stopwords were eliminated for this purpose). After an inspection of the arrays (to check if the target words were coincident with the circular concept), we settled for the following key terms taken as indicators of circularity in battery patents (consolidated as word groups with their variations): "reuse" ("re use", "reuse", "re using", "reusing"); "repair" ("repair", "repairing"); "recycle" ("recycle", "recycling"); "recover" ("recover", "recovering", "retrieve", "retrieving").

Our third step was to identify all patent documents in which one or more of these keywords appeared in their title or abstract. We find that in our total of 92,700 IPFs there are 924 observations (1%) for which we are able to ascribe circular characteristics. As Fig. 8 shows, batteries with circular characteristics have trended upwards in absolute numbers (but not in proportion to the total, a dismal finding from this approach).

Results may suggest that batteries have been developed, built, and managed in ways that have improved but still fall short of what would be expected from a full circularity concept, as we have operationalized it and which admittedly may be imperfect. Notwithstanding, the text-as-data approach we have implemented may still be revealing as the majority of the circular IPFs that were found tend to emphasize "Reuse" and "Repair" terminology. Circular concerns are still not very relevant in the battery innovation landscape, but patent analysis could still be developed in the future so as to monitor progress. Such an understanding may lead to both policy and analytical implications, namely, battery design and engineering heuristics could be nudged to more circular set-ups and patent-based research methodologies could be improved.

5. Discussion

Examining Fig. 1, one could infer that the stop-and-go moment between 2011 and 2012 may result from the global financial crisis and the subsequent recession. Assessing Figs. 1, 3, and 4 jointly, one can identify a clear difference in annual battery patenting activity between the two decades assessed in this study (2000–2009 and 2010–2019), both on a global level and for several countries. Combining this knowledge with Fig. 2, it is shown that Asia drives the major part of the increase in battery patenting activity.

The observation obtained from Fig. 2 that the Asian continent has by far the highest battery IPF output worldwide should be accompanied by the remark that the countries classified as "Asia" in PATSTAT account for approximately 60% of the world's labor force. Additionally, when computing each continents' battery IPF intensities, one observes that Asia falls behind both Europe and North America. For interested readers, IPF intensities for each continent are displayed in Fig. B.10 in Appendix B.

Concerning the country-wise patent dynamics presented in Figs. 3 and 4, it is worthwhile mentioning that comprehensive analyses undertaken before defining the final dataset resulted in the observation that most battery patent applications from China in the considered time frame of 2000–2019 are only filed nationally. Given the IPF constraint deployed for this study and the IEA and EPO report [7], these solely nationally filed applications are not considered in either one. In fact, in the current study's dataset, IPFs make up only 19.4% of all battery patent families. It is reasonable to define the data for the current study as such (the same for the recent analysis undertaken by IEA and EPO) because it can be expected that patents filed in only one country are of considerably lesser "value" than international patent families. Including them would thus result in a rather inhomogeneous dataset. Nonetheless, it is worth noting that if the IPF restriction was to be discarded and one-country patent families were to be considered, China (which in fact is the world's largest producer and market) would take the first place in battery patent counts in the majority of years of the recent decade. As a resulting thought, it would be worthwhile investigating the battery patenting dynamics of China in detail within the context of future research to shed light on why China's battery patenting behavior is so nationally-focused and what implications this has for technology analyses in this field.

This study found robust country clusters as they advance along emergent battery innovation pathways. This outcome means there is country variation in terms of technological capabilities and strategies; but also differentials in the pro-circularity pathways ahead. We are thus witnessing specialization and heterogeneous technological trajectories regarding this dimension of the energy transition. As we remarked, these different profiles and choices may be non-neutral in terms of circularity potential. By interpreting the clustering solution presented in Section 4.3, the three resulting clusters could be characterized as follows:

- Cluster 1 – Lead–acid based:
  Many of these countries' battery innovation results are made up of lead–acid battery patents. Their share of battery patents related to the four analyzed emerging technologies is close to zero, except for their lithium–sulfur component, which accounts for approximately 8% of their IPF output in 2010–2019. This "insurgent" cluster contains countries like India, Turkey, and Russia that are considerably industrialized but are not known for their innovative impact on cutting-edge clean technology. This may be a relatively circularity-poor cluster.
- Cluster 2 – Redox advantage:
  Relative to the other two clusters, these countries are putting an increased focus on the two emerging technologies of redox-flow and solid-state batteries. Their patent output related to lead–acid batteries is the lowest of the three clusters and their sodium-ion-related IPF share is close to zero. This cluster contains high-tech industrial nations like the US, Germany, and Taiwan, countries that are known to have explicitly expressed their ambitions in the field of battery technology. The somewhat less exploratory outline of this cluster does not make it the most potentially pro-circular.
- Cluster 3 – Sodium-ion driven:
  These countries focus on sodium-ion and lead–acid batteries, which account for about 35% and 24%, respectively. They have almost no innovative output in solid-state, have a relevant share of redox-flow, and exhibit a greater share in lithium–sulfur batteries compared to the other two clusters. This cluster comprises countries like Canada, China, and the UK. The bet on three promising non lithium-ion technologies may suggest that there is a high pro-circularity potential to be realized.

Interestingly, the wordage material available in the patent documents helps us to build a more detailed and comprehensive picture of battery development. Trigram analysis indicates that batteries are mutating into more complex compacts of technology, able to serve new needs (such as more flexible charging and more mobile applications). There are also some suggestions of pro-circularity as rechargeability and less toxicity seem key organizing principles of battery innovation from the outset of our time frame.

Empirical observations point to a process of technological diversification that offers promising prospects for the Circular Economy. That is to say, lithium-ion does show up as in the data as the hegemonic solution in the battery solution space. However, there are signs of early stages of development in alternatives like emergent redox-flow, lithium–sulfur, and sodium-ion technologies. Batteries based on different materials contribute to alleviate the pressure on finite resources exerted by the dominant conventional lithium-ion by promoting a more balanced exploitation of the Earth's raw materials thereby minimizing impacts on endowments and habitats. Thus, conserving geodiversity is important to the effective management of nature's resources and ensuring the sustainability of environmental conditions [93]. Moreover, multiple learning paths involving a variety of blossoming knowledge options are also valuable from the point of view of long-term economic evolutionary adaptation [94]. That is, as and stressed by much of the literature the economics of technical change, in dynamic processes of change the co-existence of alternatives (that are more in number, distinct in kind, more balanced in terms of portfolio) is relevant for research governance and an insurance against lock-in, constituting potential avenues for future progress in face of irreversibilities and technological uncertainty [40,41].

However, the road towards circularity is not without hurdles. To pave the way to a truly Circular Economy it is essential to consider the place of technologies and organizational arrangements, as well as their interdependencies and complementarities [28,34,91]. Hence, we have to consider the sources of battery innovation, and the rate and direction of technical change, but also assume that storage is part of an evolving socio-technical system (i.e. batteries are no "silver bullet" that kills all storage problems). To develop a whole-of-system approach it is necessary to consider the material elements involved in batteries (how scarce they are, how much quantity is needed, if they are toxic, etc.) and to go beyond the "end-of-pipe" mentality so as to encompass their recyclability (the conditions of the incorporation of recycled materials and the after-life of batteries in the recycling chain). For this transition to take effect also in battery development, non-technological innovation have also to be deployed.

In terms of the overall limitations of this study, it is clear that patents are only a pale indicator of the transition towards a Circular Economy. The patent data, the ITF construct and the source they have all well-known idiosyncrasies which we can only triangulate against by doing a variety of empirical strategies. Content analysis and the effectiveness of extracting circularity markers in patents, taken as a corpus of textual resources, provide extra leverage but have also their own limitations. Patents nevertheless allow for a better empirical appraisal of systemic transformation if only imperfectly so. Certainly, patent evidence does not speak for itself, but as the technological systems advance, they could become even more informative and, as such, be retained in the methodological toolbox.

6. Conclusions

The main findings of this research can be understood as follows. First, we undertook a comprehensive analysis of secondary battery technologies for two decades using global patent data. As such, this study complements other recent work patent-based analysis of innovation in the energy storage sector. We witness a robust upward trend in patenting activity during 2000–2019. The majority of battery patents are found to originate in Asia while high battery patent intensities are revealed in the performance of several Asian and European countries. Overall, a considerable increase in annual battery patenting activity is observed from 2000–2009 to 2010–2019.

Second, we also found that four battery technologies – redox-flow, solid-state, sodium-ion, and lithium–sulfur batteries – have displayed vibrant growth in recent years. Lithium-ion and other lithium-based battery technologies have also surged, whilst lead–acid and rechargeable alkaline batteries' share in battery patenting activity has decreased over the overall time frame. Through patent counts and content analysis we observe patterns of less-toxicity and signs of technological diversification which are conducive to more pro-circularity conditions in the evolving battery knowledge space.

Third, we find that three country clusters emerge over the four emerging battery types and the already established lead–acid technology. The first group contains lead–acid-focused countries, another with a higher focus on redox-flow and solid-state batteries, and a third group that contains countries with higher sodium-ion and lithium–sulfur-related patenting shares. The case can be made that these clusters differ in their degree of pro-circularity potential.

Fourth, through a text mining approach we observed that several developments are defining the knowledge frontier. Namely, we find that technologies and applications such as energy storage systems, battery management systems, wireless power transmission, electric vehicle charging, and uncrewed aerial vehicles (i.e., drones) are growing in relevance both in absolute terms and relative to general battery patenting activity. These developments show that batteries are empowering new ranges of applications, and becoming more effective solutions for the transformative turn in the techno-economic paradigm.

Fifth, the link between battery innovation and economic circularity may be illusive. Although it remains hard to grasp through patent-based methodologies, there are changes that can be associated with progress towards cleaner, less-toxic, more reusable, and more usage-adaptable battery solutions. We find that batteries with circular characteristics have risen in absolute numbers, especially after 2010. The dynamics, however, was not faster than the average thus remaining low in terms of proportion. Evidence on circularity in battery innovations seems so far to be more heavily tilted towards re-use and repair features, and less so towards recycling and recovery of materials. As such, we find some signs of pro-circularity in battery innovation, although not always in a straightforward manner and still not having a priority standing as an heuristics driving research efforts.

All in all, the intersections between storage and circularity via patenting evidence have only been scratched on the surface, more work along these lines is surely promising. Notwithstanding, our results have strategic implications at various levels. To start with, technological cosmopolitanism is a global common good and the best efforts in the realm of international relations should be channeled towards ensuring a free flow of knowledge between the new and old world innovation players; in particular, as with other emergent technologies major developments in batteries are already "post-western", and this new reality should be embraced and managed, not resisted or blocked. Then, given technological uncertainty and critical material dependency/scarcity a portfolio approach should be nurtured at the science and industrial policy level; specifically, structural diversity, open designs, and non-lithium alternatives should be regarded as favorable in to advance energy transition towards sustainability. Also, as different countries specialize in different battery segments, technologists and managers could be made more aware that while batteries promote a cleaner world, they remain heavy on environmental pressures in terms of toxic chemicals and demanding in terms of mineral requirements; that is to say, researchers and entrepreneurs should more explicitly target circularity-friendly set-ups as they navigate the battery knowledge space. In sum, the continuous exploration of new circular opportunities needs a holistic set of strategies at a variety of levels so as to manage drivers' innovation and barriers to battery scale-up. The next decade of battery development could, and should, be oriented by more explicitly circular guideposts.

Understanding the technological development of "clean tech" through data like patents is always an arduous task. Our approach consisted of a systematic appraisal of data and highlights robust results that can be further inquired in the future. In the case of batteries, patent data are thus found to indicate patterns of progress that are both interesting, from an analytical perspective, and useful, from a policy perspective. Batteries are a crucial component of a moving circular target as society adapts to the climate crisis. Techno-economic change requires continuous work on the indicator front as well.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Bruno Damásio and Philipp Metzger acknowledge the financial support provided by Fundação para a Ciência e a Tecnologia, Portugal (FCT) under the project UIDB/04152/2020 — Centro de Investigação em Gestão de Informação (MagIC). Sandro Mendonça acknowledges support by FCT, Portugal, by the Business Research Unit (BRU-IUL) and by UECE-REM (Research Unit on Complexity and Economics). BRU-IUL also benefited from grants UID/GES/00315/2013, UIDB/00315/2020; UIDB/05069/2020; PTDC/EGE-ECO/30690/2017 and is part of the project PTDC/EGE-ECO/30690/2017. José Silva acknowledges FCT support under the project UIDB/50019/2020–IDL.

Appendix A. Data and methods

A.1. The raw data

This study's foundation is the PATSTAT database [25] provided by the European Patent Office, more precisely the Autumn 2021 edition of PATSTAT Online. Transact-SQL or T-SQL is the language used for querying it. The query designed for selecting and downloading the data used for this study is defined in the text file "PATSTAT_Online_query.txt", which is included in the GitHub repository associated with this work, which can be found by following this link: https://github.com/ph1001/battery_patents.git.

The patents that were downloaded from PATSTAT and that make up the raw dataset for this study were all patent applications (including ungranted) that are part of patent families whose intra-family value for the feature "earliest publication year" lies in the time frame of 1999–2019 (the timeframe was later reduced to 2000–2019) and which contain at least one IPC entry matching one of the following codes: **H01M…** (processes or means, e.g., batteries, for the direct conversion of chemical energy into electrical energy), **H02J 3/32** (circuit arrangements for AC mains or AC distribution networks using batteries with converting means), **H02J 7…** (circuit arrangements for charging or depolarizing batteries or for supplying loads from batteries), or **B60L 53…** (methods of charging batteries, specially adapted for electric vehicles; charging stations or onboard charging equipment therefor; exchange of energy storage elements in electric vehicles).

PATSTAT Online has the restriction that all SQL queries must begin with a "SELECT" statement. This fact makes analyses of a higher complexity impossible to achieve inside PATSTAT Online itself. Consequently, data must be queried, downloaded, and then processed in a different environment. The programming language used for all steps after querying the database and downloading the data was Python [95] (Version 3.9.7), more specifically the web application Jupyter Notebook [96] (Version 6.4.3), the data processing libraries pandas [97] (Version 1.3.3) and Numpy [98] (Version 1.20.3), the visualization tools Plotly [99] (Version 5.1.0) and Seaborn [100] (Version 0.11.2), the text mining suite Natural Language Toolkit (NLTK) [101] (Version 3.6.5), and the analytics toolboxes Scikit-learn [102] (Version 0.24.2) and SciPy [103] (Version 1.7.1).

Ancillary sources were used. The labor force counts used for scaling were downloaded from the World Bank's website [104] and for the specific case of Taiwan from the website of "National Statistics: Republic of China (Taiwan)" [105].

A.2. Preprocessing and data reduction

Preprocessing and data reduction steps undertaken to obtain the final dataset from the raw data downloaded from PATSTAT are defined in the Jupyter Notebook "01_create_dataset.ipynb", which is included in the GitHub repository linked above. The following paragraphs contain a summary of these preprocessing steps.

First, the raw data downloaded from PATSTAT Online was loaded and checked for its integrity. Then each patent family's earliest intra-family values for the features "earliest publication date" and "earliest publication year" were determined and added as new columns to every row of the dataset (i.e., they were harmonized on patent family level). Like this, patent families can easily be assigned to their respective year later during the analyses. Next, all patent families were classified and tagged as either "IPF", "singleton", or "neither". The resulting tags are stored in the newly created column "tag". Next, more tags for further data selection were created. This process took place in five steps as described below:

- First, every patent family was scanned for the IPC codes related to non-active battery parts, electrodes, or secondary cells (IPC codes H01M 2…, H01M 50…, H01M 4…, and H01M 10…). Patent families containing any of these codes were added in their entirety, except if they contained any of the IPC codes H01M 6…, H01M 8…, H01M 12…, H01M 14…, or H01M 16…, which are related to primary cells, fuel cells, hybrid cells, electrochemical current or voltage generators not provided for in groups H01M 6/00-H01M 12/00, and structural combinations of different types of electrochemical generators, which were hereby explicitly excluded from the analysis. The patent families passing this stage were tagged as "non-active parts, electrodes, secondary cells".

- In a second step, every patent family was scanned for the IPC codes related to “circuit arrangements for ac mains or ac distribution networks using batteries with converting means” (H02J 3/32), “circuit arrangements for charging or depolarizing batteries or for supplying loads from batteries” (H02J 7...), “methods of charging batteries, specially adapted for electric vehicles” (B60L 53...), or “secondary cells; methods for charging or discharging” (H01M 10/44). Patent families that contained any of these codes were added in their entirety, except if they contained any of the IPC codes listed for exception in the above step or any of the codes B60L 53/54, B60L 53/55, or B60L 53/56 that refer to charging stations using fuel cells, capacitors, or mechanical storage means, respectively. Patent families that passed this stage were tagged as “charging”.
- As a third step, to identify affiliations of the resulting patent families to a set of technological categories, each patent family’s titles and abstracts were scanned using individual sets of regular expressions for each technology. These regular expressions are defined in the Jupyter notebook “01_create_dataset.ipynb”. Titles and abstracts of all languages were considered and a patent family was selected in its entirety if any substring of its titles or abstracts matched any of the respective regular expressions. Note that—to decrease the risk of false positives—before scanning abstracts for these regular expressions, they were cut off at the beginning of any appearance of the string “independent claims are also included for”. The selected patent families were assigned the value 1 in the newly created columns with the column name “is x”, with $x \in$ {Lead–acid, Lithium–air, Lithium-ion, Lithium–sulfur, Other Lithium, Magnesium-ion, nickel–cadmium, nickel–iron, nickel–zinc, nickel–metal hydride, Rechargeable alkaline, Sodium–sulfur, Sodium-ion, Solid-state, Aluminium-ion, Calcium(-ion), Organic radical} being the name of the respective technology. Please note that due to the considerable overlap of the concept of solid-state batteries with other technologies, especially lithium-ion batteries, all patent families that were classified as patents related to solid-state batteries were untagged in any other category in which they acquired tags through the process described here. To be very clear: This especially means that the lithium-ion battery category does not contain any patent families tagged as solid-state battery inventions.
- The fourth step’s purpose was to add patent data related to redox-flow and nickel–hydrogen batteries to the dataset. For this purpose, a combination of IPC classes queries and text queries was deployed. The reason for this separate step is that redox-flow and nickel–hydrogen batteries are closely related to fuel cells. Consequently, patents associated with them are often included in IPC classes that were excluded by the above steps. Analogous to the above steps, the IPC classes qualifying for potential inclusion were H01M 2..., H01M 50..., H01M 4..., H01M 8..., and H01M 10... and the IPC classes demanding exclusion were H01M 6..., H01M 12..., H01M 14..., and H01M 16.... Analogous to the above step, these patent families’ titles and abstracts were then scanned using one set of regular expressions for redox-flow and another for nickel–hydrogen batteries. These regular expressions can be reviewed in the Jupyter notebook “01_create_dataset.ipynb”. All patent families that passed this stage were assigned the value 1 in the newly created columns with the names “is redox-flow” or “is nickel–hydrogen”, respectively.
- As the last step, another additional column was computed: The dataset column “technologies one hot sum” contains the sum across each row’s “is <technology name>” values. This sum is needed in the rare cases where technology classifications overlap. The share of patent families with more than one technology associated with them was 0.61% in the final dataset. The counts resulting from these overlapping technologies were not counted multiple times but, using the respective “technologies one hot sum” value, distributed as equal fractions across the overlapping classes.

The tags created in the above steps were used for selecting the appropriate data for each analysis. All patent families not having the “IPF” tag were filtered out before all analyses. They were kept in the unfiltered dataset only for completeness, having potential future analyses with a broader scope in mind. The data selection method applied before each analysis that is based on the labels whose creation was described above is presented in Fig. A.9:

A.3. Counting patents

As already mentioned in the Introduction, the methodological setup of this study roughly follows the framework defined in the IEA and EPO report [7]. This means that all dates in this study refer to the earliest publication date within the respective IPF, and the geographic distributions were calculated based on the geographic information assigned to the respective inventors in PATSTAT. Each inventor was assigned an equal fraction of the respective count where multiple inventors were indicated. We believe there is a limitation to this approach, which is described as follows: For identifying the inventors, their PATSTAT name attribute “psn_name” is used. The harmonization of this feature, which PATSTAT carried out, is not complete. For example, pairs of entries like “KERUEL BERNARD” and “BERNARD KERUEL” exist, which in reality correspond to the same inventor, but are consequently treated as two different individuals. This shifts the fractions of countries of origin in these entries’ patent families in favor of the country of the unharmonized name.

The code used for counting patents by countries is contained in the Jupyter Notebook “02_counts_technologies_clustering.ipynb”, which is part of the GitHub repository linked at the beginning of this section.

A.4. Methods: Battery technologies

Unlike the IEA and EPO report [7], in the current study fractional counting also applied when breaking down counts by technological categories. Whenever an IPF was classified as belonging to more than one category, each technology was assigned an equal fraction of the respective count. This situation only happened in a tiny minority of the cases since only 0.61% of all IPFs were assigned to more than one technology. The code used for counting patents by technologies is contained in the Jupyter Notebook “02_counts_technologies_clustering.ipynb”, which is part of the GitHub repository linked at the beginning of this section.

A.5. Methods: Clustering

The metric R^2 applied for comparing the performance of several clustering algorithms using varying numbers of clusters can be characterized as follows:

$$R^2 = \frac{SSB}{SST} = \frac{SST - SSW}{SST} = 1 - \frac{SSW}{SST} \in [0, 1] \tag{A.1}$$

where

$$SSB = \sum_{i=1}^{p} n_i (\overline{X}_i - \overline{X})^2 = sum\ of\ squared\ differences\ between\ groups \tag{A.2}$$

and

$$SSW = \sum_{i=1}^{p} \sum_{j=1}^{n_i} (X_{ij} - \overline{X}_i)^2 = sum\ of\ squared\ differences\ within\ groups \tag{A.3}$$

and

$$SST = \sum_{i=1}^{p} \sum_{j=1}^{n_i} (X_{ij} - \overline{X})^2 = total\ sum\ of\ squared\ differences \tag{A.4}$$

All patent data with the intra-patent-family earliest publication year in 1999-2019 containing at least one IPC class entry defined in the SQL query.

↓ Tag data

Still the same data but with newly created tags columns “tag" (values: “IPF," “singleton," and “neither"), “non-active parts, electrodes, or secondary cells," “charging," and “is <technology name>" (values: 0 and 1).

↓ Filter using tags

IPFs that were tagged as at least one of “non-active parts, electrodes, or secondary cells," “charging," “is redox-flow," or “is nickel–hydrogen."

↓ Discard year 1999

Battery IPFs with intra-family earliest publication year lying inside 2000-2019. Possibility to disaggregate further by using “is <technology name>" columns.

Fig. A.9. Flow chart depicting the data selection process for this study. The entire raw dataset was labeled using newly created columns. Before each analysis, the final dataset was acquired by filtering, using labels and timestamp columns.

with

$p = number\ of\ clusters,$

$n_i = number\ of\ elements\ in\ cluster\ i,$

$\overline{X_i} = centroid\ of\ cluster\ i,$

$\overline{X} = center\ of\ whole\ dataset,\ and$

$X_{ij} = jth\ element\ of\ cluster\ i.$

Relations (A.1) are true if and only if $SST = SSW + SSB$, which is the case because:

$$SST = \sum_{i=1}^{p}\sum_{j=1}^{n_i}(X_{ij} - \overline{X})^2 = \sum_{i=1}^{p}\sum_{j=1}^{n_i}(X_{ij} \underbrace{-\overline{X_i} + \overline{X_i}}_{=0} - \overline{X})^2$$

$$= \sum_{i=1}^{p}\sum_{j=1}^{n_i}(X_{ij} - \overline{X_i})^2 + \sum_{i=1}^{p}\sum_{j=1}^{n_i}(\overline{X_i} - \overline{X})^2 + 2\sum_{i=1}^{p}\sum_{j=1}^{n_i}(X_{ij} - \overline{X_i})(\overline{X_i} - \overline{X})$$

$$= \sum_{i=1}^{p}\sum_{j=1}^{n_i}(X_{ij} - \overline{X_i})^2 + \sum_{i=1}^{p}\sum_{j=1}^{n_i}(\overline{X_i} - \overline{X})^2 + 2\sum_{i=1}^{p}(\overline{X_i} - \overline{X}) \overbrace{\underbrace{\sum_{j=1}^{n_i}(X_{ij} - \overline{X_i})}_{=0}}^{(A.5)}$$

$$= \sum_{i=1}^{p}\sum_{j=1}^{n_i}(X_{ij} - \overline{X_i})^2 + \sum_{i=1}^{p}\sum_{j=1}^{n_i}(\overline{X_i} - \overline{X})^2$$

$$= \overbrace{\sum_{i=1}^{p}\sum_{j=1}^{n_i}(X_{ij} - \overline{X_i})^2}^{(A.3)} + \overbrace{\sum_{i=1}^{p} n_i(\overline{X_i} - \overline{X})^2}^{(A.2)} = SSW + SSB$$

with

$$\sum_{j=1}^{n_i}(X_{ij} - \overline{X_i}) = \sum_{j=1}^{n_i} X_{ij} - \sum_{j=1}^{n_i}\overline{X_i} = \frac{n_i}{n_i}\sum_{j=1}^{n_i} X_{ij} - n_i\overline{X_i} = n_i\overline{X_i} - n_i\overline{X_i} = 0 \quad (A.5)$$

A higher R^2 value indicates a better clustering solution, given a non-varying dataset and a fixed number of clusters. Clustering algorithms that were compared are k-means and hierarchical agglomerative clustering using complete, average, single, and Ward linkage. The numbers of tested clusters ranged from two to nine.

The decision to use only the five dimensions “lead–acid”, “redox-flow”, “solid-state”, “sodium-ion”, and “lithium-sulfur” resulted from extensive testing of other configurations, especially those that included “Lithium-ion”, “Other lithium”, or a joint category of “Lithium-ion and other lithium”. These tests were not found to be satisfying since it was observed that the lithium-related IPFs were overshadowing the other categories due to their sheer amount, resulting in clustering solutions that lacked the clear interpretability of the solution presented in this work. Lithium-air batteries, another battery technology that has received increased attention in recent years [13], was considered a candidate feature for this analysis but was discarded due to its still very low yearly IPF counts. The code used for clustering countries based on their technology distribution is contained in the Jupyter Notebook “02_counts_technologies_clustering.ipynb”, which is part of the GitHub repository linked at the beginning of this section.

A.6. Methods: Title and abstract mining

Unigrams, bigrams, and trigrams were extracted from cleaned abstracts and titles from which meaningless words and phrases had been removed and in which certain synonyms and anomalies had been treated. The n-gram counts method simply counts occurrences and displays them as annual sums. In contrast, the n-gram intensities method does the same with the difference that its resulting values are scaled using each years’ numbers of abstracts or titles, respectively. Three ways for presenting the identified n-grams were designed for this study:

- Method 1a: Sorted in descending order of increase over the given timeframe of 2000–2019 with the measure used for sorting being $m_1 = count_{last} - count_{first}$.
- Method 1b: Sorted in ascending order of increase over the given timeframe of 2000–2019 with the measure used for sorting being

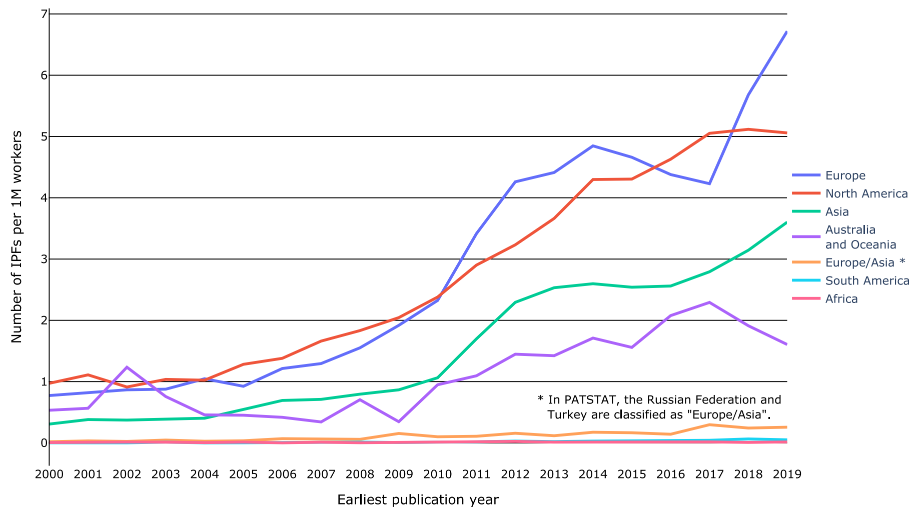

Fig. B.10. Battery IPFs per 1M workers by inventors' continents of origin, 2000–2019. In terms of battery IPF intensities, Europe and North America outperform Asia.

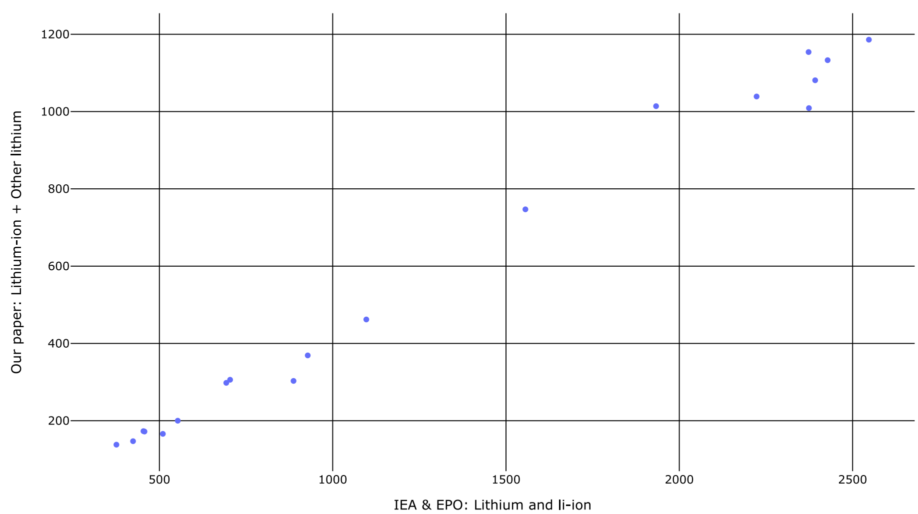

Fig. C.11. Linear relationship between the sum of the "Lithium-ion" and "Other lithium" series from this study and the "Lithium and li-ion" series from the IEA and EPO report.

m_1. This method's purpose is to show n-grams that exhibit a negative increase, i.e., have decreased over the given time period.
- Method 2: Sorted in descending order with the measure used for sorting being $m_2 = \sum abs(year - to - year\ difference_{i,i+1})$. This method's purpose is to show n-grams whose count or intensity changed the most (in absolute terms) between all adjacent years.

The results displayed in the tables that are presented in this study were obtained using method 1a, patent abstracts, and trigrams. The code for computing these results is contained in the Jupyter Notebook "03_title_ and _abstract_mining.ipynb", which is part of the GitHub repository linked at the beginning of this section. The results obtained by using the methods and data combinations not presented in this paper can best be viewed by opening the HTML file "03_title_ and _abstract_mining.html", which is also available in the same folder. The combinations for which results were computed can be characterized by the Cartesian product $c = \{n = 1,\ n = 2,\ n = 3\} \times \{n - gram\ counts,\ n - gram\ intensities\} \times \{method\ 1a,\ method\ 1b,\ method\ 2\} \times \{titles,\ abstracts\}$.

Appendix B. Battery IPF intensities for each continent

Fig. B.10 presents the development of the number of battery IPFs per 1M workers (battery IPF intensities) for each continent. In terms of battery IPF intensities, Europe and North America outperform Asia. Asia contributed approximately 60% to the global labor force in the 2000–2019 timeframe (Europe and North America contributed approximately 9% and 8%, respectively). This imbalance explains why Asia's battery patenting activity is lower in the perspective of this representation.

Appendix C. Comparison with the IEA and EPO report

There is a discrepancy between our study and the [7] report in terms of data volume. As remarked above, the difference, however, is not easy to pin down. The comparisons are not direct since, for instance, our study presents "Lithium-ion" and "Other lithium" separately while the authors of the IEA and EPO report display a joint "Lithium and li-ion" series in their Figure 4.6. We conclude that we can replicate the trends but not the levels (higher in the IEA and EPO report). To double-check the correlation between our and IEA and EPO's lithium variable we plot "Lithium-ion" + "Other lithium" from this study and "Lithium and li-ion" from [7] against each other. This indeed yields a very linear relationship as shown in Fig. C.11

References

[1] V. Smil, Energy and Civilization: A History, MIT Press, 2018.

[2] C. Freeman, F. Louçã, F. Louçã, F. Louça, F. Louçâ, As time goes by: from the industrial revolutions to the information revolution, Oxford University Press, 2001.

[3] F. Louçã, B. Damásio, C. Vasconcelos, S. Mendonça, The fundamental things apply: Twenty years after, the impact of As Time Goes By, mimeo, 2023.

[4] A. Stephan, T.S. Schmidt, C.R. Bening, V.H. Hoffmann, The sectoral configuration of technological innovation systems: Patterns of knowledge development and diffusion in the lithium-ion battery technology in Japan, Res. Policy 46 (4) (2017) 709–723.

[5] I.E. Nikolaou, K.P. Tsagarakis, An introduction to circular economy and sustainability: Some existing lessons and future directions, Sustain. Prod. Consum. 28 (2021) 600–609.

[6] A. Yildizbasi, Blockchain and renewable energy: Integration challenges in circular economy era, Renew. Energy 176 (2021) 183–197.

[7] IEA, EPO, Innovation in batteries and electricity storage, a global analysis based on patent data, 2020.

[8] L.J. Aaldering, J. Leker, C.H. Song, Analysis of technological knowledge stock and prediction of its future development potential: The case of lithium-ion batteries, J. Clean. Prod. 223 (2019) 301–311.

[9] M. Alochet, J.P. MacDuffie, C. Midler, Mirroring in production? Early evidence from the scale-up of battery electric vehicles (BEVs), Ind. Corp. Chang. (2022).

[10] A. Block, C.H. Song, Exploring the characteristics of technological knowledge interaction dynamics in the field of solid-state batteries: A patent-based approach, J. Clean. Prod. 353 (2022) 131689.

[11] J.P. Murmann, B.A. Schuler, Exploring the structure of internal combustion engine and battery electric vehicles: implications for the architecture of the automotive industry, Oxford Academic, Industrial and Corporate Change Association, 2022, pp. 1–26.

[12] L. Jiang, F. Zou, Y. Qiao, Y. Huang, Patent analysis for generating the technology landscape and competition situation of renewable energy, J. Clean. Prod. 378 (2022) 134264.

[13] L.J. Aaldering, C.H. Song, Tracing the technological development trajectory in post-lithium-ion battery technologies: A patent-based approach, J. Clean. Prod. 241 (2019) 118343, http://dx.doi.org/10.1016/j.jclepro.2019.118343, URL https://www.sciencedirect.com/science/article/pii/S0959652619332135.

[14] J.A. Silva, T. Guilhemre, S. Mendonça, Reconfiguring the battery innovation landscape, Foresight STI Gov. 17 (1) (2023) 34–50.

[15] A. De Jesus, P. Antunes, R. Santos, S. Mendonça, Eco-innovation pathways to a circular economy: Envisioning priorities through a delphi approach, J. Clean. Prod. 228 (2019) 1494–1513.

[16] A. De Jesus, M. Lammi, T. Domenech, F. Vanhuyse, S. Mendonça, Eco-innovation diversity in a circular economy: Towards circular innovation studies, Sustainability 13 (19) (2021) 10974.

[17] M. Aykol, P. Herring, A. Anapolsky, Machine learning for continuous innovation in battery technologies, Nat. Rev. Mater. 5 (10) (2020) 725–727.

[18] A. De Jesus, P. Antunes, R. Santos, S. Mendonça, Eco-innovation in the transition to a circular economy: An analytical literature review, J. Clean. Prod. 172 (2018) 2999–3018.

[19] K.M. Winslow, S.J. Laux, T.G. Townsend, A review on the growing concern and potential management strategies of waste lithium-ion batteries, Resour. Conserv. Recy. 129 (2018) 263–277.

[20] P. Zhu, D. Gastol, J. Marshall, R. Sommerville, V. Goodship, E. Kendrick, A review of current collectors for lithium-ion batteries, J. Power Sources 485 (2021) 229321.

[21] D.L. Thompson, J.M. Hartley, S.M. Lambert, M. Shiref, G.D. Harper, E. Kendrick, P. Anderson, K.S. Ryder, L. Gaines, A.P. Abbott, The importance of design in lithium ion battery recycling–a critical review, Green Chem. 22 (22) (2020) 7585–7603.

[22] S.S. Sharma, A. Manthiram, Towards more environmentally and socially responsible batteries, Energy Environ. Sci. 13 (11) (2020) 4087–4097.

[23] Z. Li, Y.-K. Kuo, A.R. Mahmud, A.A. Nassani, M. Haffar, I. Muda, Integration of renewable energy, environmental policy stringency, and climate technologies in realizing environmental sustainability: Evidence from OECD countries, Renew. Energy 196 (2022) 1376–1384.

[24] S. Wang, E. Zhao, H.K. Razzaq, Dynamic role of renewable energy efficiency, natural resources, and climate technologies in realizing environmental sustainability: Implications for China, Renew. Energy 198 (2022) 1095–1104.

[25] G. De Rassenfosse, H. Dernis, G. Boedt, An introduction to the patstat database with example queries, Aust. Econ. Rev. 47 (2014) http://dx.doi.org/10.1111/1467-8462.12073.

[26] S. Mendonça, U. Schmoch, P. Neuhäusler, Interplay of patents and trademarks as tools in economic competition, in: Springer Handbook of Science and Technology Indicators, Springer Handbooks. Springer, Cham, 2019, http://dx.doi.org/10.1007/978-3-030-02511-3_42.

[27] S. Mendonça, H. Confraria, M.M. Godinho, Appropriating the returns of patent statistics: Take-up and development in the wake of Zvi Griliches, SSRN (2021) http://dx.doi.org/10.2139/ssrn.3971764.

[28] J.A. Silva, S. Oliveira, G. Távora, S. Mendonça, et al., The role of innovation in the future PV and storage markets, in: Proceedings 31st European Photovoltaic Solar Energy Conference and Exhibition, 2015, pp. 3175–3178.

[29] A. Stephan, C.R. Bening, T.S. Schmidt, M. Schwarz, V.H. Hoffmann, The role of inter-sectoral knowledge spillovers in technological innovations: The case of lithium-ion batteries, Technol. Forecast. Soc. Change 148 (2019) 119718, http://dx.doi.org/10.1016/j.techfore.2019.119718, URL https://www.sciencedirect.com/science/article/pii/S0040162517306583.

[30] A. Stephan, L.D. Anadon, V.H. Hoffmann, How has external knowledge contributed to lithium-ion batteries for the energy transition? Iscience 24 (1) (2021) 101995.

[31] M. Baumann, T. Domnik, M. Haase, C. Wulf, P. Emmerich, C. Rösch, P. Zapp, T. Naegler, M. Weil, Comparative patent analysis for the identification of global research trends for the case of battery storage, hydrogen and bioenergy, Technol. Forecast. Soc. Change 165 (2021) 120505.

[32] G. Távora, J. Silva, S. Mendonça, Evolução tecnológica no armazenamento de energia: uma perspectiva a partir das patentes de baterias, in: CIES2020-XVII Congresso IbÉRico E XIII Congresso Ibero-Americano de Energia Solar, LNEG-Laboratório Nacional de Energia e Geologia, 2020, pp. 1323–1329.

[33] H.-A.H. Dang, U. Serajuddin, Tracking the sustainable development goals: Emerging measurement challenges and further reflections, World Dev. 127 (2020) 104570.

[34] A. De Jesus, S. Mendonça, Lost in transition? Drivers and barriers in the eco-innovation road to the circular economy, Ecol. Econom. 145 (2018) 75–89.

[35] P. Morales, M. Flikkema, C. Castaldi, A.-P. de Man, The effectiveness of appropriation mechanisms for sustainable innovations from small and medium-sized enterprises, J. Clean. Prod. (2022) 133921.

[36] N.M. Denter, L.J. Aaldering, H. Caferoglu, Forecasting future bigrams and promising patents: introducing text-based link prediction, Foresight (2022).

[37] A. De Jesus, M. Lammi, T. Domenech, F. Vanhuyse, S. Mendonça, Eco-innovation diversity in a circular economy: Towards circular innovation studies, Sustainability 13 (19) (2021) 10974.

[38] J. Caraça, B.-A. ke Lundvall, S. Mendonça, The changing role of science in the innovation process: From queen to cinderella? Technol. Forecast. Soc. Change 76 (6) (2009) 861–867, http://dx.doi.org/10.1016/j.techfore.2008.08.003, URL https://www.sciencedirect.com/science/article/pii/S0040162508001455.

[39] B. Ribeiro, P. Shapira, Private and public values of innovation: A patent analysis of synthetic biology, Res. Policy 49 (1) (2020) 103875.

[40] A. Nicita, U. Pagano, The Evolution of Economic Diversity, Vol. 2, Psychology Press, 2001.

[41] A. Stirling, A general framework for analysing diversity in science, technology and society, J. R. Soc. Interface 4 (15) (2007) 707–719.

[42] F. Castellaci, S. Grodal, S. Mendonca, M. Wibe, Advances and challenges in innovation studies, J. Econ. Issues 39 (1) (2005) 91–121, http://dx.doi.org/10.1080/00213624.2005.11506782, arXiv:https://doi.org/10.1080/00213624.2005.11506782.

[43] M. Mazzanti, E. Zecca, 18. Industry, innovations and transition to the green and circular economy, Handb. Ind. Dev. (2023) 302.

[44] A.B. Santos, M.L. Bogers, M.T. Norn, S. Mendonça, Public policy for open innovation: Opening up to a new domain for research and practice, Technol. Forecast. Soc. Change 169 (2021) 120821, http://dx.doi.org/10.1016/j.techfore.2021.120821, URL https://www.sciencedirect.com/science/article/pii/S0040162521002535.

[45] V. Tiberius, H. Schwarzer, S. Roig-Dobón, Radical innovations: Between established knowledge and future research opportunities, J. Innov. Knowl. 6 (3) (2021) 145–153.

[46] I.E. Kosmadakis, C. Elmasides, G. Koulinas, K.P. Tsagarakis, Energy unit cost assessment of six photovoltaic-battery configurations, Renew. Energy 173 (2021) 24–41.

[47] M.S. Ziegler, Evaluating and improving technologies for energy storage and backup power, Joule 5 (8) (2021) 1925–1927, http://dx.doi.org/10.1016/j.joule.2021.07.015, URL https://www.sciencedirect.com/science/article/pii/S2542435121003536.

[48] P.A. Østergaard, N. Duic, Y. Noorollahi, S. Kalogirou, Renewable energy for sustainable development, Renew. Energy (2022).

[49] M. Marx, J.S. Gans, D.H. Hsu, Dynamic commercialization strategies for disruptive technologies: Evidence from the speech recognition industry, Manage. Sci. 60 (12) (2014) 3103–3123, http://dx.doi.org/10.1287/mnsc.2014.2035, arXiv:https://doi.org/10.1287/mnsc.2014.2035.

[50] I.E. Kosmadakis, C. Elmasides, D. Eleftheriou, K.P. Tsagarakis, A techno-economic analysis of a pv-battery system in Greece, Energies 12 (7) (2019) 1357.

[51] A. Vezzini, 15 - Lithium-ion battery management, in: G. Pistoia (Ed.), Lithium-Ion Batteries, Elsevier, Amsterdam, 2014, pp. 345–360, http://dx.doi.org/10.1016/B978-0-444-59513-3.00015-7, URL https://www.sciencedirect.com/science/article/pii/B9780444595133000157.

[52] P. Nzereogu, A. Omah, F. Ezema, E. Iwuoha, A. Nwanya, Anode materials for lithium-ion batteries: A review, Appl. Surf. Sci. Adv. 9 (2022) 100233.

[53] C. Li, Z.-y. Wang, Z.-j. He, Y.-j. Li, J. Mao, K.-h. Dai, C. Yan, J.-c. Zheng, An advance review of solid-state battery: Challenges, progress and prospects, Sustain. Mater. Technol. 29 (2021) e00297.

[54] C. Banza Lubaba Nkulu, L. Casas, V. Haufroid, T. De Putter, N.D. Saenen, T. Kayembe-Kitenge, P. Musa Obadia, D. Kyanika Wa Mukoma, J.-M. Lunda Ilunga, T.S. Nawrot, et al., Sustainability of artisanal mining of cobalt in DR Congo, Nat. Sustain. 1 (9) (2018) 495–504.

[55] W. Fu, D. Kim, F. Wang, G. Yushin, Stabilizing cathodes and interphases for next-generation li-ion batteries, J. Power Sources 561 (2023) 232738.

[56] J.G. Kim, B. Son, S. Mukherjee, N. Schuppert, A. Bates, O. Kwon, M.J. Choi, H.Y. Chung, S. Park, A review of lithium and non-lithium based solid state batteries, J. Power Sources 282 (2015) 299–322.

[57] A. Block, C.H. Song, Exploring the characteristics of technological knowledge interaction dynamics in the field of solid-state batteries: A patent-based approach, J. Clean. Prod. 353 (2022) 131689.

[58] J. Garche, P. Moseley, E. Karden, Lead–acid batteries for hybrid electric vehicles and battery electric vehicles, in: Advances in Battery Technologies for Electric Vehicles, Elsevier, 2015, pp. 75–101.

[59] G.J. May, A. Davidson, B. Monahov, Lead batteries for utility energy storage: A review, J. Energy Storage 15 (2018) 145–157.

[60] T. Li, X. Bai, U. Gulzar, Y.-J. Bai, C. Capiglia, W. Deng, X. Zhou, Z. Liu, Z. Feng, R. Proietti Zaccaria, A comprehensive understanding of lithium–sulfur battery technology, Adv. Funct. Mater. 29 (32) (2019) 1901730.

[61] M. Zhao, B.-Q. Li, X.-Q. Zhang, J.-Q. Huang, Q. Zhang, A perspective toward practical lithium–sulfur batteries, ACS Cent. Sci. 6 (7) (2020) 1095–1104.

[62] A. Manthiram, S.-H. Chung, C. Zu, Lithium–sulfur batteries: progress and prospects, Adv. Mater. 27 (12) (2015) 1980–2006.

[63] H. Zhang, X. Li, J. Zhang, Redox Flow Batteries: Fundamentals and Applications, CRC Press, 2017.

[64] G. Tomazic, M. Skyllas-Kazacos, Redox flow batteries, in: Electrochemical Energy Storage for Renewable Sources and Grid Balancing, Elsevier, 2015, pp. 309–336.

[65] C.-N. Sun, M. Mench, T. Zawodzinski, High performance redox flow batteries: an analysis of the upper performance limits of flow batteries using non-aqueous solvents, Electrochim. Acta 237 (2017) 199–206.

[66] K. Kubota, M. Dahbi, T. Hosaka, S. Kumakura, S. Komaba, Towards K-ion and Na-ion batteries as ''beyond Li-ion'', Chem. Rec. 18 (4) (2018) 459–479.

[67] J.-M. Tarascon, Na-ion versus Li-ion batteries: complementarity rather than competitiveness, Joule 4 (8) (2020) 1616–1620.

[68] M.D. Slater, D. Kim, E. Lee, C.S. Johnson, Sodium-ion batteries, Adv. Funct. Mater. 23 (8) (2013) 947–958.

[69] OECD, OECD Patent Statistics Manual, OECD PUBLICATIONS, 2, rue André-Pascal, 75775 PARIS CEDEX 16, 2009.

[70] S. Mendonça, T.S. Pereira, M.M. Godinho, Trademarks as an indicator of innovation and industrial change, Res. Policy 33 (9) (2004) 1385–1404.

[71] C. Castaldi, All the great things you can do with trademark data: Taking stock and looking ahead, Strateg. Organ. 18 (3) (2020) 472–484.

[72] C. Castaldi, S. Mendonça, Regions and trademarks: research opportunities and policy insights from leveraging trademarks in regional innovation studies, Reg. Stud. 56 (2) (2022) 177–189.

[73] G. Tsironis, T. Daglis, K.P. Tsagarakis, Social media and EU companies' engagement in circular economy: A linkedin approach, Sustain. Prod. Consum. 32 (2022) 802–816.

[74] Z. Griliches, Patent statistics as economic indicators: A survey, J. Econ. Lit. 28 (4) (1990) 1661–1707, URL http://www.jstor.org/stable/2727442.

[75] S. Losacker, 'License to green': Regional patent licensing networks and green technology diffusion in China, Technol. Forecast. Soc. Change 175 (2022) 121336.

[76] B.K. Sovacool, P. Newell, S. Carley, J. Fanzo, Equity, technological innovation and sustainable behaviour in a low-carbon future, Nat. Hum. Behav. 6 (3) (2022) 326–337.

[77] WIPO, World intellectual property report 2022 - the direction of innovation, 2022.

[78] G. Zubi, R. Dufo-López, M. Carvalho, G. Pasaoglu, The lithium-ion battery: State of the art and future perspectives, Renew. Sustain. Energy Rev. 89 (2018) 292–308.

[79] C.M. Costa, T.L. Quintanilha, S. Mendonça, Manuel castells and informationalism 1, in: Management, Organizations and Contemporary Social Theory, Routledge, 2019, pp. 256–270.

[80] C. Cardoso, C.M. Costa, S. Mendonça, B. Damásio, The "Network Society" moves in mysterious ways: 25 years in the reception of a core concept, 2023, mimeo.

[81] A. Malhotra, H. Zhang, M. Beuse, T. Schmidt, How do new use environments influence a technology's knowledge trajectory? A patent citation network analysis of lithium-ion battery technology, Res. Policy 50 (9) (2021) 104318, http://dx.doi.org/10.1016/j.respol.2021.104318, URL https://www.sciencedirect.com/science/article/pii/S0048733321001190.

[82] N. Kittner, F. Lill, D.M. Kammen, Energy storage deployment and innovation for the clean energy transition, Nat. Energy 2 (9) (2017) 17125, http://dx.doi.org/10.1038/nenergy.2017.125.

[83] M.S. Ziegler, J.E. Trancik, Re-examining rates of lithium-ion battery technology improvement and cost decline, Energy Environ. Sci. 14 (2021) 1635–1651, http://dx.doi.org/10.1039/D0EE02681F, URL http://dx.doi.org/10.1039/D0EE02681F.

[84] A. Dechezleprêtre, Y. Ménière, M. Mohnen, International patent families: from application strategies to statistical indicators, Scientometrics 111 (2) (2017) 793–828, http://dx.doi.org/10.1007/s11192-017-2311-4.

[85] U. Schmoch, B. Gehrke, China's technological performance as reflected in patents, Scientometrics 127 (1) (2022) 299–317, http://dx.doi.org/10.1007/s11192-021-04193-6.

[86] R. Frietsch, H. Kroll, China's foreign technology market perspectives, Deutsch-Chin. Innov.: Rahm. Chancen Herausford. (2020) 365–375, Metropolis.

[87] D.H. Hsu, P.-H. Hsu, Q. Zhao, Rich on paper? Chinese firms' academic publications, patents, and market value, Res. Policy 50 (9) (2021) 104319, http://dx.doi.org/10.1016/j.respol.2021.104319, URL https://www.sciencedirect.com/science/article/pii/S0048733321001207.

[88] P. Neuhäusler, O. Rothengatter, R. Frietsch, Patent applications - structures, trends and recent developments 2018, in: Studien zum deutschen Innovationssystem 4-2019, Leibniz Information Centre for Economics - ZBW, Berlin, 2019, URL http://hdl.handle.net/10419/194274.

[89] P.-H. Hsu, D. Lee, P. Tambe, D.H. Hsu, Deep learning, text, and patent valuation, SSRN (2020) http://dx.doi.org/10.2139/ssrn.3758388, URL https://ssrn.com/abstract=3758388.

[90] F. Wang, O. Borodin, M.S. Ding, M. Gobet, J. Vatamanu, X. Fan, T. Gao, N. Eidson, Y. Liang, W. Sun, et al., Hybrid aqueous/non-aqueous electrolyte for safe and high-energy Li-ion batteries, Joule 2 (5) (2018) 927–937.

[91] C. Lehmann, F. Cruz-Jesus, T. Oliveira, B. Damásio, Leveraging the circular economy: Investment and innovation as drivers, J. Clean. Prod. (2022) 132146.

[92] M. Alizadeh, A. Kashef, Y. Wang, J. Wang, G.E.O. Kremer, J. Ma, Circular economy conceptualization using text mining analysis, Sustain. Prod. Consum. (2022).

[93] J. Hjort, J.E. Gordon, M. Gray, M.L. Hunter Jr., Why geodiversity matters in valuing nature's stage, Conserv. Biol. 29 (3) (2015) 630–639.

[94] S. Mendonça, The revolution within: ICT and the shifting knowledge base of the world's largest companies, Econ. Innov. New Technol. 15 (8) (2006) 777–799.

[95] G. Van Rossum, F.L. Drake, Python 3 Reference Manual, CreateSpace, Scotts Valley, CA, 2009.

[96] T. Kluyver, B. Ragan-Kelley, F. Pérez, B. Granger, M. Bussonnier, J. Frederic, K. Kelley, J. Hamrick, J. Grout, S. Corlay, P. Ivanov, D. Avila, S. Abdalla, C. Willing, Jupyter notebooks – a publishing format for reproducible computational workflows, in: F. Loizides, B. Schmidt (Eds.), Positioning and Power in Academic Publishing: Players, Agents and Agendas, IOS Press, 2016, pp. 87–90.

[97] W. McKinney, Data structures for statistical computing in python, in: Stéfan van der Walt, Jarrod Millman (Eds.), Proceedings of the 9th Python in Science Conference, 2010, pp. 56–61, http://dx.doi.org/10.25080/Majora-92bf1922-00a.

[98] C.R. Harris, K.J. Millman, S.J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N.J. Smith, R. Kern, M. Picus, S. Hoyer, M.H. van Kerkwijk, M. Brett, A. Haldane, J.F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, T.E. Oliphant, Array programming with NumPy, Nature 585 (7825) (2020) 357–362, http://dx.doi.org/10.1038/s41586-020-2649-2.

[99] Plotly Technologies Inc., Collaborative data science, 2015, URL https://plot.ly.

[100] M. Waskom, M. Gelbart, O. Botvinnik, J. Ostblom, P. Hobson, S. Lukauskas, D.C. Gemperline, T. Augspurger, Y. Halchenko, J. Warmenhoven, J.B. Cole, J. de Ruiter, J. Vanderplas, S. Hoyer, C. Pye, A. Miles, C. Swain, K. Meyer, M. Martin, P. Bachant, E. Quintero, G. Kunter, S. Villalba, Brian, C. Fitzgerald, C. Evans, M.L. Williams, D. O'Kane, T. Yarkoni, T. Brunner, Mwaskom/seaborn: v0.11.2 (august 2021), 2021, http://dx.doi.org/10.5281/zenodo.5205191.

[101] S. Bird, E. Klein, E. Loper, Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit, " O'Reilly Media, Inc.", 2009.

[102] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, Édouard Duchesnay, Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. 12 (85) (2011) 2825–2830, URL http://jmlr.org/papers/v12/pedregosa11a.html.

[103] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, et al., SciPy 1.0: fundamental algorithms for scientific computing in python, Nature Methods 17 (3) (2020) 261–272, http://dx.doi.org/10.1038/s41592-019-0686-2.

[104] The World Bank, Labor force, total, 2022, https://data.worldbank.org/indicator/SL.TLF.TOTL.IN Accessed: 10. Jan. 2022, 12:07.

[105] National Statistics; Republic of China (Taiwan), Labor force, total, 2022, https://eng.stat.gov.tw/ct.asp?xItem=42761&ctNode=1609&mp=5 Accessed: 10. Jan. 2022, 12:49.