Heliyon 8 (2022) e10738

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Hydrothermal biomass processing for green energy transition: insights derived from principal component analysis of international patents

Silviu Florin Acaru [a,*], Rosnah Abdullah [a,b], Daphne Teck Ching Lai [c,d], Ren Chong Lim [a,**]

[a] Centre for Advanced Material and Energy Sciences (CAMES), Universiti Brunei Darussalam, Jalan Tungku Link, BE1410 Brunei Darussalam
[b] Faculty of Science (FOS), Universiti Brunei Darussalam, Jalan Tungku Link, BE1410 Brunei Darussalam
[c] School of Digital Science, Universiti Brunei Darussalam, Jalan Tungku Link, BE1410 Brunei Darussalam
[d] Institute of Applied Data Analytics (IADA), Universiti Brunei Darussalam, Jalan Tungku Link, BE1410 Brunei Darussalam

HIGHLIGHTS

- Innovation intensities for different green innovation techniques were measured.
- Policymakers have decisive role in advancing techniques at dominant design phase.
- Emerging technologies aim attaining synthetic natural gas and C5 sugars.
- Exploratory techniques focus on sewage sludge and connectivity to wastewater plants.
- Trending techniques point towards achieving a circular economy.

ARTICLE INFO

Keywords:
Innovation intensity
Natural language processing
Principal component analysis
Hydrothermal biomass processing
Patent data analytics
Green energy transition

ABSTRACT

As efforts to achieve Net Zero are intensifying, there is a strong need to identify the technological positioning of green process innovations that can support the green energy transition. A veritable contender to support these efforts is the hydrothermal biomass processing technology. This process innovation comprises diverse techniques that can convert biomass substrates into valuable low-carbon fuels. Coordination across all available conversion approaches is encouraged to propel the application of those that consider the environmental and sustainability impacts. We assessed the innovation intensity for different techniques under this green process innovation through applying natural language processing and deployment of principal component analysis on patent data. We positioned our techniques within four distinctive groups (*intense*, *dormant*, *emerging*, and *exploratory*). In this way, we tracked which hydrothermal technique currently dominates international applications and which ones are gaining traction in the future.

1. Introduction

The impact of climate change on weather and air quality are widely experienced by communities across the globe. It is well understood that the release of greenhouse gases plays an important role in these adversities. Redressing to sustainable resource utilization as soon as possible was formally agreed upon by 196 state representatives in the 2015 Paris Agreement. In June 2021, the leaders of the Group of Seven reiterated their commitment to the Paris Agreement and led a technology-driven transition away from fossil fuels toward Net Zero (G7, 2021). At the COP26 in Glasgow, the commitment to reach ~1.5 °C was reportedly 'kept alive' and still considered a long-term goal. The collection of such efforts to diversify the global economy from consumption of fossil fuels toward renewable energy is widely known as the green energy transition.

Renewable energy sources are primarily categorized into hydroelectric, wind, solar, and biomass. Meeting the Paris agreement's demand requires a nine-fold increase, likely to be completed by wind and solar energy (Baruch-Mordo et al., 2019). The production expansion of these two renewable energy sources worries climate activists because it can devastate the environment by changing landscapes and threatening endangered species (Kiesecker et al., 2019). Currently, less than half of the power production in Europe is from renewable sources. Biomass accounts

* Corresponding author.
** Corresponding author.
E-mail addresses: s.f.acaru@outlook.com (S.F. Acaru), renchong.lim@ubd.edu.bn (R.C. Lim).

https://doi.org/10.1016/j.heliyon.2022.e10738
Received 15 December 2021; Received in revised form 15 July 2022; Accepted 16 September 2022

for 60% of the renewable energy source (Camia et al., 2021). The predictions are that renewable energy in Europe will reach 70 % by 2030 (Goldman and Group, 2022). In the US, renewable energy production is lower than the energy produced from fossil fuels and gas, approximately 12%. However, the energy generated from biomass is significant, supplying the American market with slightly less than half of its renewable energy consumption (EIA, 2021).

The benefit of biomass-derived energy over the others stands out through its ability to use the waste from various industries, thus contributing to a circular economy model. However, most biomass substrates are heterogeneous and have an increased resistance to decomposition. Hydrothermal biomass processing can transform biomass substrates into valuable low-carbon fuels through liquefaction, gasification, or carbonization. These methods can operate with different feedstock materials, catalysts additions, operating parameters, reaction medium and conditions giving rise to a broad set of techniques (Mathanker et al., 2021; Okolie et al., 2019; Sharma et al., 2020). The main incentive for researching hydrothermal biomass processing conversion is its ability to process mixed dry and wet feedstocks simultaneously. Water is used as a reactant and has a positive impact on the environment. At high pressures and temperatures, water reaches a subcritical or supercritical state, depending on the process conditions. In this way, it substitutes the toxic and environmentally harmful chemicals utilized in other conversion processes.

Despite the benefits, the acceleration of the green energy transition through this form of renewable energy is anticipated to affect other economic sectors dependent on biomass as a raw material. Large biomass feedstocks, which are required to produce energy, strain the supply chain by increasing resource demand, challenging the economic viability (Gielen et al., 2019). Investigations on the indirect environmental impact were attempted, with inconclusive results (Ingrao et al., 2019). For this reason, hydrothermal processing technologies that can use multiple or local feedstock sources would be sustainable.

Clarity concerning the operation methods, the environmental impact, and raw material utilization requires an in-depth analysis of multiple sources of innovation and technology status corroborated by a wide range of stakeholders. Therefore, the question we aim to address in this study is: *which hydrothermal biomass technique is gaining traction that could facilitate the process of the green energy transition?*

A well-known theory implies that technology follows a cyclic movement divided into four phases: an era of ferment, a dominant design, an era of incremental change, and technological discontinuity. The dominant design is the peak of a technique. It gains its supremacy over the other phases through direct investment guided by policymakers and regulators (Anderson and Tushman, 1990; Kalthaus, 2020). Innovation in hydrothermal biomass can be found in various formats and can be used to disclose the status of the technology. Indicators of innovation includes patents, among others, such as R&D expenditure, human resources, collaborative networks, surveys, or new product releases. Patent statistics have been increasingly used to extract innovation patterns and conduct technology forecasting studies (Dziallas and Blind, 2019; Wang and Zhao, 2021). Therefore, patents represent a meaningful way to gain insights on technological change and provide an output of innovation measurements.

Nevertheless, not every patent document can be considered as input into deriving this measurement form. There are many reasons behind a patent filing, ranging from copyright protection, investment attraction, an increase in company valuation, or an expressive hint of launching a new product to the market. The cost associated with the application under multiple jurisdictions can be correlated with the indication of high interest and potentially significant investment in the field. Thereby, patents filed under the Patent Cooperation Treaty (PCT) have an increased chance of commercialization. An analysis of patent filings covering international patent families for low-carbon energy showed that the green energy overtook the fossil fuels in the early 2000s. A rapid expansion was recorded until 2010, with a further accelerated trend until 2013, followed by a slight drop and steady growth until 2019 (IEA, 2021). The PCT applications in renewable technologies followed the same trend as international patent families for low-carbon energy. Reported patent publications hit the highest number in 2012, at 4500 filings from 830 in 2002, and stabilized to an average of 3000 published patents per year until 2019 (Nurton, 2020).

Our study will focus on highly specific patent data filed under the PCT related to hydrothermal and biomass. We deploy a methodology based on traditional methods such as NLP algorithms to collate relevant inputs from the patent claims section. Claims are the source of legal protection against infringements and allow the owner the exploitation for a limited period of time. In this way, we access the most valuable technical information. We then apply the principal component analysis (PCA) algorithm to measure the innovation intensity for different techniques. We articulate the innovation intensity of different techniques by using the International Patent Classification (IPC) system, a hierarchical arrangement centered on language-independent symbols (WIPO, 2019). We transform the loading distance of the first two principal components (PCs) from the PCA into magnitude of component loadings. Implementing this algorithm aimed to gain insights into the data, which can help funnel the techniques capable of answering the research question. We provide background information concerning patent analytics in Section 2, then we describe the methodology in Section 3 and present the results and discussion in Section 4.

2. Background for patent data analytics

The available tools vary, ranging from commercial subscription services (e.g., Azure, Hana, Sisense, Neural Designer) to open source projects (Python and R libraries) (Saini et al., 2020). Technologies can be explored, and the trends can be predicted through various natural language processing (NLP) techniques and algorithm combinations. Depending on the scope of study, some sections of the patent are processed, while others are disregarded.

Patent bibliometric and semantic data was used as input and methods involving citations (Wang et al., 2020), patent data vectoring (Aharonson and Schilling, 2016), networks (Choi and Hwang, 2014), keyword occurrence frequency (Yoon and Magee, 2018) were used to track knowledge flows across time and regions, reveal a company patent portfolio, conduct technological trend analysis and build topographic maps for uncovering new technological opportunities. Technological trends (Chanchetti et al., 2016) and emerging technology elements (Moehrle and Caferoglu, 2019) were forecasted by using patent information sections such as titles, abstracts and IPC codes. NLP methods such as tokenization, part-of-speech tagging, and syntactic parsing were used to build features for analysis and interpretation from patent claims section (Han et al., 2017).

The latest developments in intellectual property analytics highlight that most research focuses on artificial neural networks, backpropagation learning methods, support vector machines, or conditional random fields (Aristodemou and Tietze, 2018). Suominen et al. (2017) were able to contrast companies' knowledge profiles by combining topic modeling such as Latent Dirichlet Allocation (LDA) with network analysis. Zhu et al. (2002) used PCA to produce conceptual indices to enhance information retrieval and group related terms into conceptual clusters. In another study, the PCA algorithm was used to generate key terms that would represent the input for subsequent processing steps (Zhou et al., 2019). Wu et al. (2016) combined three data mining techniques, namely self-organizing maps (SOM), kernel principal component analysis (KPCA), and support vector machine (SVM), to analyze patents and predict their quality. SOM was used for clustering, and the results were then used as input to KPCA to extract the main patent features from patent documents. Subsequently, SVM was utilized to assess the classification quality of the patents.

Although substantial progress was achieved in this field, the methods described so far do not measure the innovation intensity of different patented techniques. Clustering patents together will highlight the similarity between documents; the use of titles, abstracts, and IPC codes would only provide a brief overview of the techniques. The discovery of emerging elements via semantic analysis does not indicate an industrial transition, which is only used for improvement purposes. Understanding company knowledge profiles is also beneficial, but it does not guarantee that the technologies they patented are intensely researched. Methods such as key term generation and classification quality assessment present only a tiny fraction of the innovation areas and cannot be used in assessing innovation intensity. As a progressive development, the method presented in this article articulates the innovation intensity of different techniques utilizing the magnitude of component loadings derived from the PCA algorithm.

3. Methodology

Our methodology consisted of a step-by-step workflow (outlined in Figure 1) to collate a patent dataset sample from the WIPO's database, PATENTSCOPE. The scope of the dataset was to capture high-impact patents that are more likely to contribute to the energy transition rather than capture a high number of patents but with reduced relevance and significance. Hence, a filter was applied to select patents published under the PCT as applicants who filed PCT applications have the intention to commercialize the technique covered in multiple jurisdictions. Furthermore, filtering for PCT applications ensured that we have only selected one patent from a family of patents protecting a single invention, as multiple patents are often filed to protect commercially valuable patents in different jurisdictions. The search strategy, covering a broad range with maximized precision, included the words *hydrothermal* and *biomass* in the 'Front-Page' field, which is used to select patent filings for the technological area of interest, i.e., hydrothermal biomass processing. The word "processing" was not used as it is redundant given that it is well understood that the word "hydrothermal" is an adverbial term for a process. The resulted search string was "*FP:(hydrothermal and biomass) AND OF:(WO)*" where "*OF:(WO)*" filters for PCT documents. The period was defined between 2006 to 2020, where the initial year coincided with the significant reform of the patent classification system implemented on January 1, 2006.

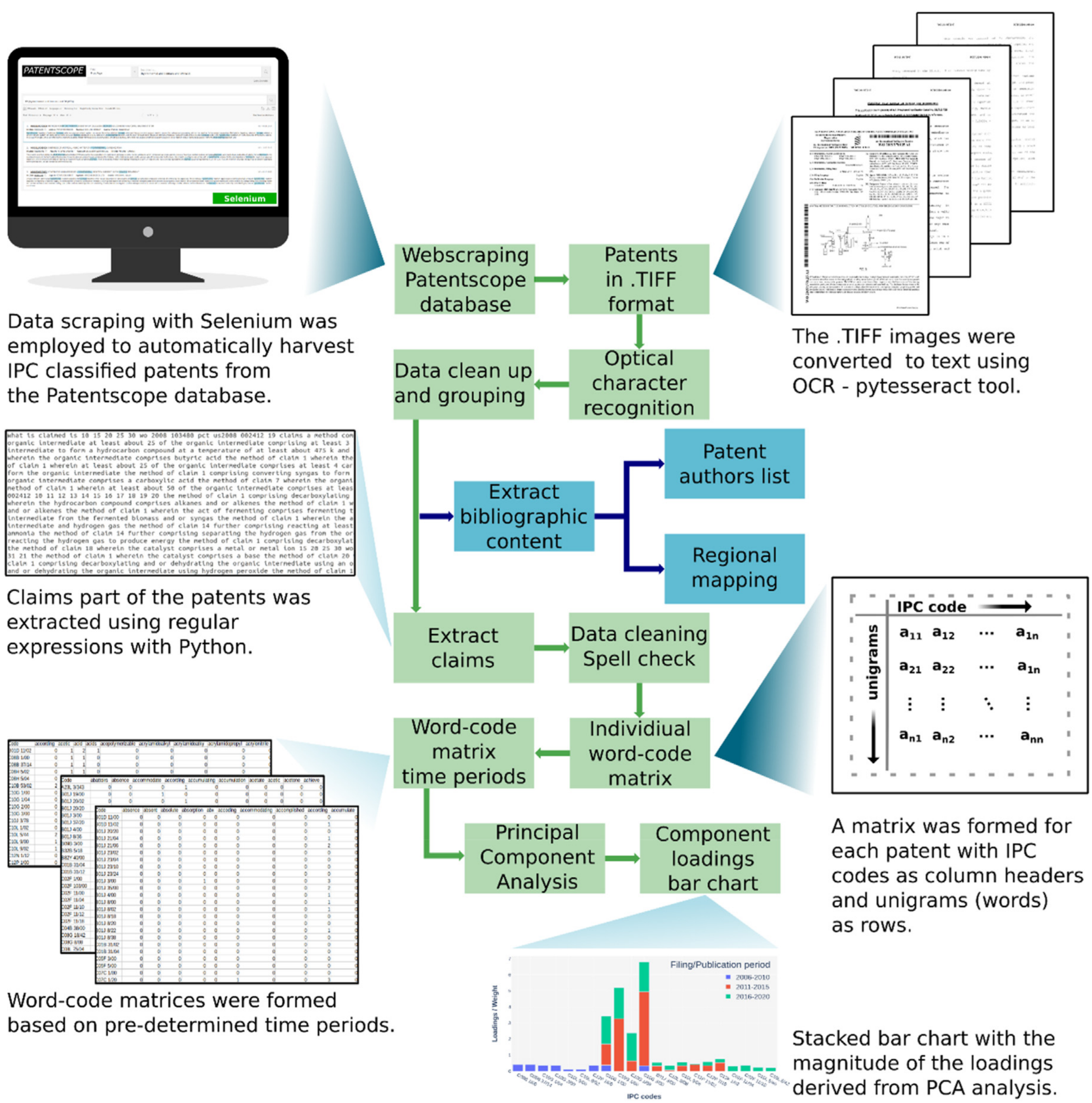

Figure 1. Details the processing workflow used to extract data for NLP analysis. The green arrow path is used for PCA investigation. The blue arrow-sub path reveals the applicants' names and builds the regional mapping.

Firstly, a total of 194 PCT documents were downloaded in ZIP format. The IPC codes were also harvested from the HTML page located under the PCT bibliographic data tab, during this step. The ZIP files contained the TIFF images of the patent documents. The retrieved documents were language checked, and several were discarded due to the use of languages other than English in the main body, indicating that not all PCT documents are fully translated into English. The process resulted in a dataset containing 108 PCT documents. Each of the TIFF images were converted to text via optical character recognition. The bibliographic metadata from the front pages containing the patent applicant, the filing date, and the country of origin was extracted using NLP techniques. To survey the applicant profiles, the metadata was then analyzed. We must mention that some patents had designated applicants from different countries. In this study, only the first applicant was considered when analyzing country distribution. For example, if the first applicant is listed as Shell US while the second applicant is listed as Shell Netherlands, only Shell US was selected.

Secondly, the claims section extraction was performed using regular NLP expressions. The documents had no identifiable section to define the end of the claims. Hence, the text included sentences up to the end of the patent document. These were removed in subsequent text cleaning. Additionally, the text was lower-cased, abbreviations and single characters were removed, and the whole text was spell-checked.

The next step included splitting the patent documents into three time periods based on the patenting trends reported in green energy technologies. The first matrix comprised the period between 2006 and 2010. The start of our selected period coincided with the year that major reform of the patent classification was implemented by WIPO, on January 1, 2006. The end of this subset was in line with the change in the patenting trend reported earlier for low-carbon energy patents (IEA, 2021). The second matrix comprised of patents between 2011 and 2015. We chose these years to contrast with the trends reported in the case of PCT applications for renewable technologies (Nurton, 2020). The end of the data subset allows for capturing patents filed in 2013 with the 18 months processing time for publication from being granted. The last matrix contained patent data from 2016 to 2020. We selected this period to capture the steady growth for low-carbon energy patents. The selected patent date corresponded to the *international filing date*, which is the date at which a patent document was submitted to a patent office. Where this date was not available, the publication date was considered.

For each patent document, a matrix was built using unique single words, called unigrams, as rows and IPC codes as columns. Each element in the matrix represented the word count for the respective claim in the patent document. Such a matrix can offer insights between words and IPC codes representing a given technique. Lastly, for each period all the patent document matrices were merged. The elements from individual matrices were treated as unique entries and were summed up.

In the next step, the PCA algorithm was applied to each of the three matrices to obtain a metric as a measure for innovation intensity. PCA is a mathematical algorithm that reduces the dimensionality of datasets while minimizing information loss. It is a statistical tool to represent data tables into smaller datasets for trends, clusters, and outliers' purposes (Shlens, 2014). The reduction is realized by transforming the matrix with respect to a new system of coordinate axes called principal components obtained using singular value decomposition. Principal components are derived from the covariance matrix, a measure of the original matrix's correlation between the rows and columns. Eigenvalues associated with each principal component are a measure of variance contained within a given principal component. If the matrix is composed of values with different scale units, it is necessary to normalize them (Legendre and Legendre, 1998). In this case, normalization was not applied because we used the same base units derived from counting the unigrams. From the build of our matrix, IPC code for a technique under the technological domain of interest is represented by a loading vector.

The scree plots for each matrix showed that two principal components captured the most variation. Scree plot figures for each matrix are available in the Supplementary material (Figures S5, S6, S7, and S8). As such, the loadings contained two components called component loadings. The loading magnitude on a two-component axis was determined as follows:

$$Loading\ magnitude = \sqrt{Component\_Loading\_1^2 + Component\_Loading\_2^2} \quad (1)$$

The components of the loading vector are a measure of the significance of the corresponding IPC code in the new principal components. As such, we used the loading magnitude as a measure of innovation intensity for the technique represented by the IPC code of the loading vector. The magnitude can be understood as follows. A popular technique will make use of similar words across multiple patent documents. The IPC codes are the fields covering these techniques; hence similar patent documents will be filed under the same classification codes. We obtained principal component axes that capture the maximum amount of variance across the dataset by applying PCA to our matrices. PCA biplots figures for each of the matrices and the whole dataset is available in Supplementary material (Figures S1, S2, S3, and S4).

Since many factors influence innovation, the loadings indirectly result from these inputs. The loading magnitudes for each technique generated from each period matrix were stacked in a bar chart, and the loadings sharing the same IPC code were added on top of each other, resulting in a cumulative weight. This loading magnitude is the innovation intensity measured across all three periods of a given technique under an IPC code. They were plotted on an x-y graph where loading magnitudes are expressed on the y-axis, and the corresponding 139 IPC codes on the x-axis (Figure 3). Loading magnitudes were also calculated for the subsequent PC 3 and PC 4 by extending Eq. (1) (see Supplementary material, Eq. S1). Comparison between the technological landscaping using the first two PCs, and the following two (PC 3 and PC 4) is available in the Supplementary material (Figure S9).

4. Results and discussion

The PCT applications in this domain had a steady pace in the first period, 2006 to 2010, with only 7 patents filed, followed by a rapid ascension, peaking at 68 patents in the 2011–2015 interval, and decreased to 33 in the latest period, 2016 to 2020. In the following section, the causes of these records are justified using the metadata details such as patent applicants and regional mapping. Additionally, this is corroborated with the focus of the discussion, the innovation intensity of different techniques found. Based on the loading magnitudes, four groups of techniques became apparent. We positioned these groups within the technological life cycle of hydrothermal biomass processing.

4.1. Patent applicants and regional mapping

From the patent dataset sample, 38 entities were identified as applicants, which included private companies, universities, and private individuals. Between 2006 and 2020, Shell Oil Company led the sector with a total of 48 patents, amounting to 44.4% of the data. The following highest applicant by patent numbers was ExxonMobil Research and Engineering with 6 patents. Csl Carbon Solutions Ltd had 4 patents in the dataset like Battelle Memorial Institute. The rest of the applicants had either 1 or 2 patent applications. A graphic representation including names of applicants, the number of patents, and the region they filed in is shown in Figure 2. The regional filing of patents helps understand the market dynamics with regards to technology and innovation incentives. We can see that our dataset captured patents filed in multiple jurisdictions, a show of widespread support through policies and investment. 69% of the data was filed in the US, but this number is partly due to Shell Oil Company being considered a US entity only. Shell Oil Company accounted for 44.4 % of the patent data sample. Nonetheless, the chart shows that inventors from different regions are working on the

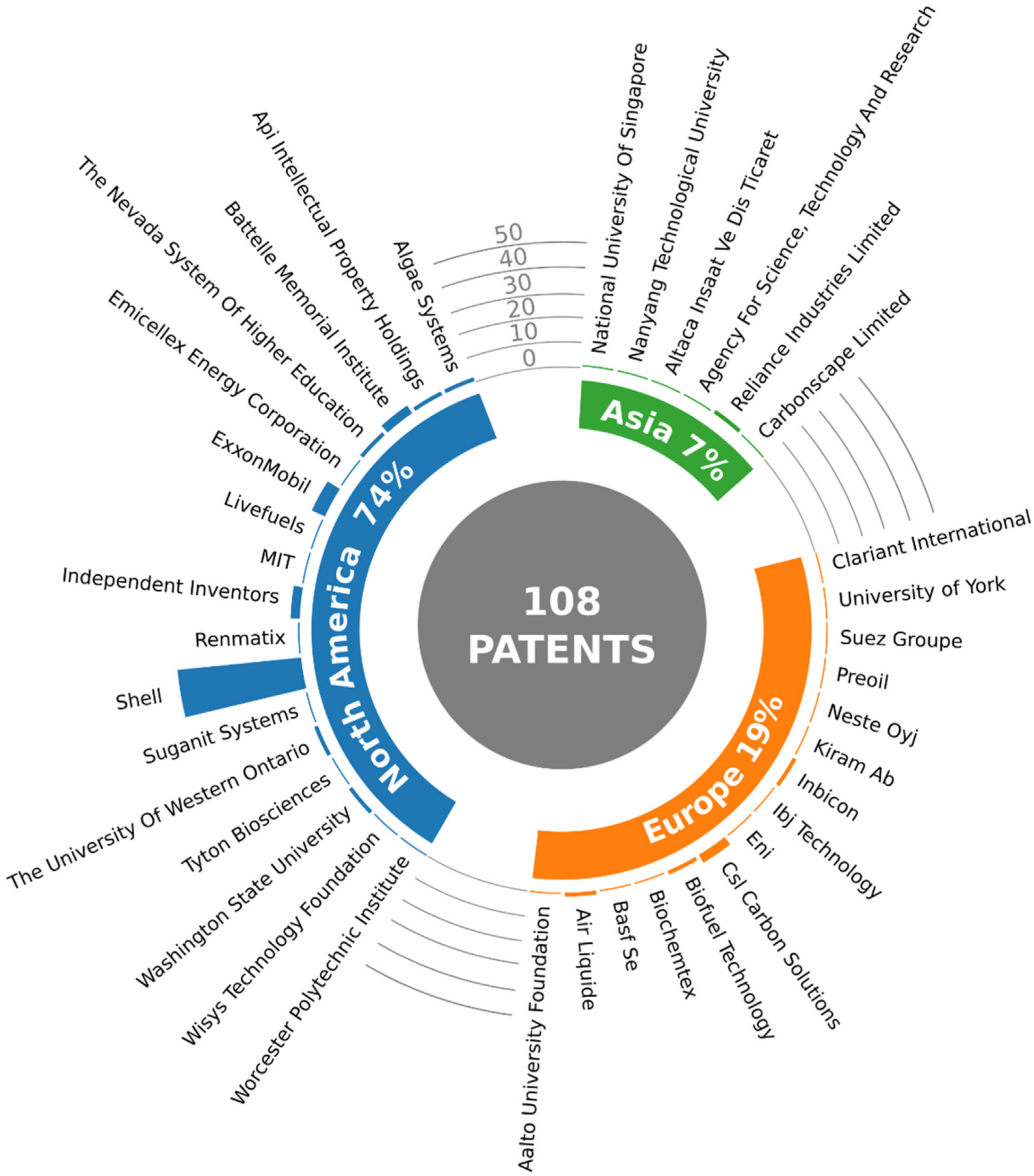

Figure 2. Is a breakdown of the region of filing for the patent dataset. Surrounding the figure are the patent owner names. The bars are a measure of the number of filings per inventor, where 0 to 50 is the guidance scale. The independent inventor names are available in Supplementary material, Table S1.

technology, hence diversifying the techniques. From an NLP standpoint, the dataset diversity is encouraging since it includes a varied lexicon. This adds up to the quality and multiplicity of the data, resulting in a balanced assessment of innovation intensities.

4.2. Assessment of innovation intensity

Loadings having a high weight result from significant efforts being absorbed, meaning innovation intensity is considerable. On the other hand, a low loading magnitude can be due to knowledge spillover and is attributed to the IPC code hierarchy or exploration of new technological avenues. Similarities between loading magnitudes and time group sharing allowed for further categorisation, resulting into four clusters for the different techniques: *dormant*, *intense*, *emerging*, and *exploratory* (Figure 3).

Dormant techniques are represented by loadings in the first period with magnitudes less than 0.5 weight units. These loadings showed no additions in the subsequent time groups. The IPC codes covered by this type of research comprised of improvements on traditional fuels, such as charcoal and fermentation processes (see Table 1), which are considered declining industries.

C08H 5/02 and C08H 5/04 are IPC codes that underwent a reclassification due to the constant development of the technology. This signifies a shift in the approach taken and regular review of the classification system by WIPO to assess the green innovation techniques.

Intense techniques are IPC codes with loading magnitudes of more than 1.5 weight units and shared across multiple periods (see Table 2). These techniques are associated with mature technology which can be due to policies, investment, and knowledge accumulation, a theory we explained in the introduction section (Anderson and Tushman, 1990; Kalthaus, 2020).

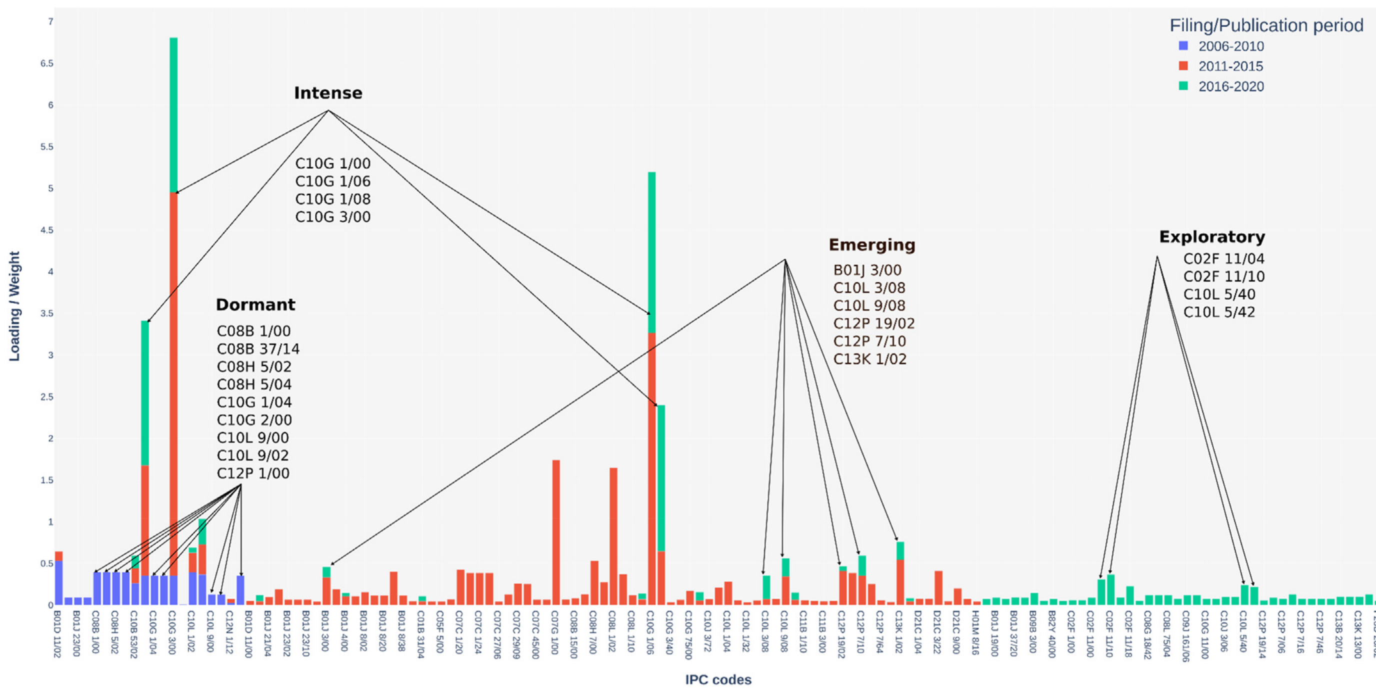

Figure 3. Illustrates the technological landscaping and the assessment of innovation intensity. The stacked bars were generated using PC 1 and PC 2 according to Eq. (1) from each subset of the patent dataset. The weight of the loadings is measured in a dimensionless unit.

Table 1. Highlights the IPC codes and dormant techniques.

Dormant techniques	
IPC code	Summary based on description found in the IPC database
C08B 1/00 C08B 37/14	Production of cellulose and hemicellulose through fermentation processes.
C08H 5/02 C08H 5/04	Macromolecular compounds derived from lignin or lignocellulosic materials. Reclassified due to technical development in the 2010 version.
C10G 1/04 C10G 2/00	Production of hydrocarbon involving steam equipment.
C10L 9/00 C10L 9/02	Improvement of traditional solid fuels properties.
C12P 1/00	Stimulated fermentation with the use of different enzymes.

Table 2. Lists the IPC codes identified under the intense area of research.

Intense techniques	
IPC code	Summary based on description found in the IPC database
C10G 1/00 C10G 1/06 C10G 1/08 C10G 3/00	Separation of hydrocarbons into useful oils by processes such as hydrogenation with moving catalysts.

Analysis of techniques under this group revealed interest in the efficient exploitation of biomass through process improvements and optimization. For example, applicants are interested in exploiting work related to pre-treatment of the biomass in the form of fatty amines by exposing it to a catalyst comprising rare earth oxide, alkali oxide, or alkaline earth oxide (Roberts et al., 2014). Exploration in this area showed that average bio-crude yields for lignocellulosic and macroalgae biomasses in hydrothermal processes without catalysts and water as solvent reached 26.9 wt.% and 17.5 wt.%, respectively.

Due to the hydrothermal pathways of the carbohydrate-rich substrates, the application of catalysts is known to improve bio-crude yields. Applying strong alkaline solutions neutralizes the pH and inhibits the formation of organic acids, thus hindering the repolymerization reactions. After applying alkali catalysts to lignocellulosic substrates, bio-crude yields increase by 50 units–135 wt.% (Haarlemmer et al., 2016; Hu et al., 2020; Zhang et al., 2018).

Another topic included "*the removal of chlorine and phosphorus/removal of metal and its anions from cellulosic biomass prior to catalytic hydrogenation/hydrogenolysis/hydrodeoxygenation*" (Powell and Chheda, 2015). Regardless of the substrate, high bio-crude yields can also be achieved by metal-based catalysts. Their presence activates the H_2 molecules and favors hydrodeoxygenation (De et al., 2015).

Specific techniques covered by IPC codes in Table 2 include the following:

- A system using concentrated thermal energy from focused high beam energy to provide sufficient energy for driving the hydrothermal liquefaction of biomass to bio-crude oil;
- An improved hydrothermal liquefaction system with separations efficiencies higher than conventional processes;
- Partial removal of water from biomass prior to hydrothermal treatment;
- Methods of feeding the reactor, such as a horizontal or an inclined surface;
- Promoting hydrogen gas distribution in the presence of a digestion solvent and slurry catalyst;
- Pretreatment of the biomass in the form of fatty amines by exposing it to a catalyst comprising rare earth oxide, alkali oxide, or alkaline earth oxide;
- Integration of a reactor unit with a catalyst reduction unit;
- Methods of separation of cellulosic fine via centrifugation within the processing loop;
- A step involving subcritical temperature to extract specific components followed by supercritical treatment of the remainder to produce bio-oil from algae biomass;
- System comprising two heating units to facilitate biomass conversion.

Emerging techniques - Stacked loadings of magnitudes up to 1.5 cumulative weight that showed interest in the last two periods defines emerging techniques. Six IPC codes and corresponding techniques were discerned (see Table 3). They have high cumulative weights, but their industrial applications have yet to make it mainstream. Furthermore, adopting these techniques on a larger scale competes with the established, mature techniques.

Analysis of the patents revealed process-related inventions. For example, the inventions protected relate to "*subjecting the biomass to an oxidation process followed by supercritical water gasification or liquefaction*" (Ozdenkci and Koskinen, 2017). Studies have shown that water at supercritical conditions, such as 400 °C and 350 bars, achieves complete deoxygenation, favoring liquefaction at the expense of gasification. The ionic product remains unchanged under these operating conditions compared to lower operating parameters reported in academic literature, such as 350 °C and 250 bars. Under these conditions, the process is still energy efficient, primarily because of the heat capacity and the low compressibility of liquids (Castello et al., 2018).

Techniques under the emerging area also covered topics related to multiple processing steps. Understandably, since the biochemical composition of biomass varies depending on the substrate. As such, it is intuitive to target processes that address attaining specific compounds. Specifics of the techniques covered by IPC codes in Table 3 also include the following:

- Flash cooling a dilute acid hydrolysis reaction of a biomass feedstock;
- A continuous feeding of the hydrothermal reactor;
- A microwave-assisted hydrothermal conversion;
- An introduction of water under pressure in a reactor set up to facilitate the C5 sugar release;
- The use of a slurry catalyst capable of activating molecular hydrogen in a hydrothermal digestion process;
- The generation of fuel gas by hydrothermal carbonization followed by another step of hydrolysate catalysis;
- Production methods of biogas and bio-oil of the extractant process in the hydrothermal reactor;
- Hydrothermal carbonization with water recirculation;
- Hydrothermal carbonization using catalysts;
- The conversion of biomass such as coal or humic substance into cement additive;
- Hydrothermal treatment to convert biomass in the presence of pH buffering agent and supported hydrogenolysis catalyst with various chemical compositions;
- A combined system (hydrothermal liquefaction and catalytic hydrothermal liquefaction) to convert biomass into bio-oils and aqueous effluent for further processing;
- A design of a system that uses solar thermal energy to conduct the hydrothermal process.

Table 3. Shows IPC codes and emerging techniques.

Emerging techniques	
IPC Code	Summary based on description found in the IPC database
B01J 3/00	Physical or chemical process or apparatus using sub-atmospheric or super-atmospheric pressure to change chemical or physical properties.
C10L 3/08	Production of synthetic natural gas.
C10L 9/08	Production of synthetic natural gas by heat treatment (e.g., calcining).
C12P 19/02	Fermentation or enzyme-using processes to synthesize a desired chemical compound or composition or to separate optical isomers from a racemic mixture. Preparation of compounds containing saccharide radicals (ketoaldonic acids C12P 7/58) - Monosaccharides.
C12P 7/10	Preparation of oxygen-containing organic compounds, substrate containing cellulosic material.
C13K 1/02	Sugar industry - saccharides obtained from natural resources or by hydrolysis. Glucose-containing syrups, obtained by saccharification of cellulosic materials.

Loading magnitudes having weight values of less than 1.5, occurring in only one period, were considered part of the catchment area, resulting from the shared concepts of the IPC codes. The patent examiners decide the relevant IPC codes for classification during the filing process. However, a technology can have multiple applications. Hence examiners allocate multiple IPC codes to one patent. This indicates that a single patent can cover multiple techniques.

Exploratory techniques are represented by loadings with magnitudes between 0.2 and 0.5 in the last period, between 2016 and 2020. Innovation intensities for these techniques are low, and their developments are still at the infant stage. This area of interest can be an outcome of aligning techniques with commercial and social demands. It supports the circular economy concept, in which the sewage feedstock derived from wastewater treatment plants can be paired with hydrothermal processes, thus, minimizing energy use and contributing to sustainable resource allocation (Fan et al., 2021; Mikulčić et al., 2021). Additionally, from a performance perspective, municipal sludge substrates showed similar heat recovery rate values when used in continuous-flow reactors (Anastasakis et al., 2018). Table 4 summarises the IPC codes classified under this group.

Low weight magnitude loadings, such as those below 0.2 weight units seen in the last period, can be associated with the mid-range loadings in the same period. One explanation is due to their hierarchy association to IPC codes and targeting similar trends in demand.

4.3. Further discussion

The method applied in this study showed that hydrothermal biomass conversion is done through various techniques. We placed them into four groups based on their publication times and the loading magnitudes. The method is unique on its own through the interpretation of the PCA algorithm but also in this subject area. As far as we know, no one investigated the "hydrothermal biomass" processing at a patent level.

During the analysis it was noted that Shell Oil Company accounted for a significant number of patents in the dataset. These were part of the matrices covering the periods between 2011 to 2015 and 2016 to 2020. A side analysis determined that the impact of this cluster of patents on the PCA results does not change the overall trend, instead it can provide additional information from a reduced data perspective. Analysis results are available in the Supplementary material (Figures S10, S11, S12, S13, and S14).

The patenting trends we identified mirror the evolution of scientific publications across time for biomass as a renewable energy source. A total of 10000 articles were retrieved between 2000 and 2019 from the Web of Science database. Publication numbers in the early 2000s were below 100, increasing to almost 800 in 2019, with a rapid ascent between 2010 and 2017. Authors in the US are significant contributors to the field, whereas other regions such as China and Western Europe also made important contributions (Ferrari et al., 2020).

The analysis methods used by other authors focused on bibliometric studies and classic literature reviews. Authors that conducted bibliometric analyses targeted specific techniques such as thermochemical conversion, hydrochar and bio-oil, and hydrothermal liquefaction, which was a scope narrower than our search criteria. With reference to the thermochemical conversion of biomass it is found that keywords such as gasification, pyrolysis and combustion are frequently used.

Table 4. Lists the IPC codes and exploratory techniques.

Exploratory techniques	
IPC code	Summary based on description found in the IPC database
C02F 11/04 C02F 11/10	Treatment of sewage communal water to obtain methane gas using processes such as pyrolysis.
C10L 5/40 C10L 5/42	Obtaining solid fuels from materials of non-mineral origin, such as animal derived origin.

Thermochemical biomass studies are deliberated to a greater extent than biochemical conversion studies (Osman et al., 2021).

Research of techniques to convert biomass into hydrochar and bio-oil found that most cited articles are about the chemical and structural properties of solid residues (610 citations at an article published in 2009), followed by processing of municipal waste (230 citations at an article published in 2011) and solid residues resulted from processing of lignocellulosic materials (216 citations at an article published in 2012) (Mikhail et al., 2020). The top 20 most cited research articles were from 2010 to 2014. Hence, we deduce that they are an essential contributor to the knowledge at their time of publication. This could explain why we had the highest number of patents in our dataset during the 2011 to 2015 period. It is in this period that we grouped the *intense* and the *emerging* techniques.

Bibliometric analyses of research related to hydrothermal liquefaction revealed that studies between 2015 and 2021 focused on optimizing process parameters through surface methodology or application of machine learning algorithms (Sahoo et al., 2021). Interestingly, under our search criteria, no patents that contain claims on artificial intelligence were filed or published. It is usually the case that when technology has multiple applications, multiple IPC codes are assigned to that patent, and our search criteria should have also captured those with applications in hydrothermal biomass processing.

In another bibliometric study, it was reported that catalysts are thoroughly researched for cost improvement purposes. Among "catalysts", feedstocks such as "algae," "food waste," and "sludge" have a high occurrence frequency in recent years (Yang et al., 2021). The use of grown for purpose and waste feedstock represents a shift from lignocellulosic-based biomass toward a circular economy model and supports the techniques grouped under the *exploratory* techniques. Conversion process evaluation focuses on reaction conditions for yield improvements. However, it is unclear which technique is of most interest in the literature. From a patent perspective, the IPC allowed us to categorize with a certain degree of confidence, therefore we gained more precise insights.

Recent literature reviews reported that technologies for biofuel production composed of thermo-bio-chemical processes and biomass combustion via coal-fired technologies are emerging. Bio-chemical conversions such as anaerobic digestion, fermentation esterification, and photo fermentation are also reportedly studied. Studies in biophotolysis, photo-fermentation, dark fermentation, and hybrid systems for bio-hydrogen production are considered emerging (Ambaye et al., 2021). Currently, in international patents, we do not identify protection covering these techniques. Instead, under the *emerging* techniques, some areas, such as microwave co-processing are a topic of interest to current inventors.

From a hardware processing perspective, batch reactors dominate the scientific research studies, while continuous processing is encouraged. Studies investigating slurry viscosity and its control through catalysts addition such as ammonia or carboxymethyl cellulose have attracted attention. Evidence of this research in our study is found under the *intense* techniques. Scientists are still to find solutions posed by environmental issues associated with the biomass volumes required to facilitate this green energy transition. High-density biomass is preferred to lignocellulosic feedstocks, and waste streams are recommended for further studies. Additionally, downstream process capabilities also experience development deficiencies, with pain points such as product filtration, that are economical and time-efficient, and hydrothermal conversion techniques that are suitable for large-scale operations (Filipe et al., 2020). From our patent dataset, we understand that continuous technologies are intensely researched, and downstream technologies are emerging, such as methods of feeding the reactors and purification steps combined with the conversion techniques.

4.4. Limitations and future work

i. It is yet unclear if patenting in this domain carries an anti-patent culture or is considered as important as other industries, where they have a pro-patents view (e.g., pharmaceutical or chemicals). Some strategies involve focusing on first-mover advantages and lead time, while others might prefer secrecy (Scherer, 2014). Vimalnath et al. (2020) found that green innovators engage in semi-opened and closed IP strategies, and no fully open IP strategy has come to light yet. Thus, some innovations might not have been captured in this dataset analysis for those reasons.
ii. The arbitrary choice of loading magnitudes gives another limitation for deriving the four life cycle groups. User interaction with domain knowledge was necessary to position the technologies within the respective groups. A classification method would be beneficial for automatic categorization, perhaps based on the information entropy concept.

5. Conclusion

The green energy transition is an ongoing process and can be accelerated through a technology-driven approach. Hydrothermal biomass conversion is a contender to support these efforts through the use of renewable energy feedstocks. In addition, this type of energy transition could mitigate the environmental impact of other renewable energy sources while creating a circular economy system. However, understanding the methods, the environmental impact, and raw material utilization was necessary at the industrial level through insights into the past and upcoming trends. Using PCT patents, we identified technology developments at different rates. The variation in progression implied the need for a differentiation scheme.

Using the PCA algorithm to determine the loading magnitudes, we categorized the techniques into *dormant, intense, emerging,* and *exploratory*. Based on this dataset and the resulted groups, we established that traditional fermentation processes and improvements of traditional fossil fuels, which fall under the *dormant* techniques, have little commercial interest. In the next group, patents under the *intense* techniques focused on improving process efficiency, highlighting the need to drive the innovation into scale-up and lower the production costs. Possible ways of achieving these targets refer to continuous processes and economical reactor designs. *Emerging* techniques focus on producing synthetic gas, using multiple processing steps to harvest multiple compounds, and using solar energy to power the hydrothermal processes. These techniques are in alignment with the carbon neutrality set by the Net Zero state. Lastly, the *exploratory* techniques are going even further in achieving Net Zero by creating a circular economy system through the use of sewage communal water and conversion of animal-derived materials into solid fuels.

Data under this group also supports the need to reduce the stress on the supply chain by implementing locally sourced feedstock. Four applicants were identified as protecting their techniques under the *emerging* group, two universities and two commercial companies located in three different regions. In conclusion, the green transition supported by hydrothermal biomass processing could see an early implementation of techniques described by the *intense* group with a shift later on toward those in the *emerging* and *exploratory* groups, which are gradually gaining traction.

Declarations

Author contribution statement

Silviu Florin Acaru: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Rosnah Abdullah: Analyzed and interpreted the data; Wrote the paper.

Daphne Teck Ching Lai & Ren Chong Lim: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Funding statement

Mr Silviu Florin Acaru was supported by Universiti Brunei Darussalam [UBD/RSCH/URC/RG(b)/2019/013].

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2022.e10738

Acknowledgements

The authors would like to thank Wan Mohammad Faris Zaini for helpful discussions related to the deployment of principal component analysis. Open-source Python packages such as selenium, pytesseract, NLTK, RegEx, pandas, NumPy, scikit-learn, and plotly were used to carry out natural language processing and principal component analysis described in this article.

References

Aharonson, B.S., Schilling, M.A., 2016. Mapping the technological landscape: measuring technology distance, technological footprints, and technology evolution. Res. Pol. 45, 81–96.
Ambaye, T.G., Vaccari, M., Bonilla-Petriciolet, A., Prasad, S., van Hullebusch, E.D., Rtimi, S., 2021. Emerging technologies for biofuel production: a critical review on recent progress, challenges and perspectives. J. Environ. Manag. 290.
Anastasakis, K., Biller, P., Madsen, R.B., Glasius, M., Johannsen, I., 2018. Continuous hydrothermal liquefaction of biomass in a novel pilot plant with heat recovery and hydraulic oscillation. Energies 11, 1–23.
Anderson, P., Tushman, M.L., 1990. Technological discontinuities and dominant designs: a cyclical model of technological change. Adm. Sci. Q. 35, 604.
Aristodemou, L., Tietze, F., 2018. The state-of-the-art on Intellectual Property Analytics (IPA): a literature review on artificial intelligence, machine learning and deep learning methods for analysing intellectual property (IP) data. World Patent Inf. 55, 37–51.
Baruch-Mordo, S., Kiesecker, J.M., Kennedy, C.M., Oakleaf, J.R., Opperman, J.J., 2019. From Paris to practice: sustainable implementation of renewable energy goals. Environ. Res. Lett. 14, 024013.
Camia, A., Giuntoli, Jonsson, J., Robert, R., 2021. The Use of Woody Biomass for Energy Production in the EU. Publications Office of the European Union, Luxembourg.
Castello, D., Pedersen, T.H., Rosendahl, L.A., 2018. Continuous hydrothermal liquefaction of biomass: a critical review. Energies 11.
Chanchetti, L.F., Oviedo Diaz, S.M., Milanez, D.H., Leiva, D.R., de Faria, L.I.L., Ishikawa, T.T., 2016. Technological forecasting of hydrogen storage materials using patent indicators. Int. J. Hydrogen Energy 41, 18301–18310.
Choi, J., Hwang, Y.S., 2014. Patent keyword network analysis for improving technology development efficiency. Technol. Forecast. Soc. Change 83, 170–182.
De, S., Saha, B., Luque, R., 2015. Hydrodeoxygenation processes: advances on catalytic transformations of biomass-derived platform chemicals into hydrocarbon fuels. Bioresour. Technol. 178, 108–118.
Dziallas, M., Blind, K., 2019. Innovation indicators throughout the innovation process: an extensive literature analysis. Technovation 80– 81, 3–29.
EIA, 2021. Total Energy Monthly Data - U.S. Energy Information Administration (EIA) [WWW Document]. U.S. Energy Inf. Adm. accessed 6.12.22 https://www.eia.gov/totalenergy/data/monthly/.
Fan, Y. Van, Romanenko, S., Gai, L., Kupressova, E., Varbanov, P.S., Klemeš, J.J., 2021. Biomass integration for energy recovery and efficient use of resources: tomsk Region. Energy 235.
Ferrari, G., Pezzuolo, A., Nizami, A.-S., Marinello, F., 2020. Bibliometric analysis of trends in biomass for. Energies 13, 3714.
Filipe, R., Hu, Y., Shui, H., Charles, C., 2020. Biomass and Bioenergy Hydrothermal liquefaction of biomass to fuels and value-added chemicals : products applications and challenges to develop large-scale operations. Biomass Bioenergy 135, 105510.
G7, 2021. Carbis Bay G7 Summit Communiqué. G7, Cornwall UK, p. 6.
Gielen, D., Boshell, F., Saygin, D., Bazilian, M.D., Wagner, N., Gorini, R., 2019. The role of renewable energy in the global energy transformation. Energy Strategy Rev. 24, 38–50.
Goldman, T., Group, S., 2022. The Rise of Power in European Economies.
Haarlemmer, G., Guizani, C., Anouti, S., Déniel, M., Roubaud, A., Valin, S., 2016. Analysis and comparison of bio-oils obtained by hydrothermal liquefaction and fast pyrolysis of beech wood. Fuel 174, 180–188.
Han, Q., Heimerl, F., Codina-Filba, J., Lohmann, S., Wanner, L., Ertl, T., 2017. Visual patent trend analysis for informed decision making in technology management. World Patent Inf. 49, 34–42.
Hu, Y., Gu, Z., Li, W., Xu, C., Charles), 2020. Alkali-catalyzed liquefaction of pinewood sawdust in ethanol/water co-solvents. Biomass Bioenergy 134, 105485.
IEA, 2021. Patents and the Energy Transition. EPO Report, Paris.
Ingrao, C., Bacenetti, J., Adamczyk, J., Ferrante, V., Messineo, A., Huisingh, D., 2019. Investigating energy and environmental issues of agro-biogas derived energy systems: a comprehensive review of Life Cycle Assessments. Renew. Energy 136, 296–307.
Kalthaus, M., 2020. Knowledge recombination along the technology life cycle. J. Evol. Econ. 30, 643–704.
Kiesecker, J., Baruch-Mordo, S., Kennedy, C.M., Oakleaf, J.R., Baccini, A., Griscom, B.W., 2019. Hitting the target but missing the mark: unintended environmental consequences of the Paris climate agreement. Front. Environ. Sci. 7.
Legendre, P., Legendre, L., 1998. Numerical Ecology, Second Eng.
Mathanker, A., Das, S., Pudasainee, D., Khan, M., Kumar, A., Gupta, R., 2021. A review of hydrothermal liquefaction of biomass for biofuels production with a special focus on the effect of process parameters, Co-solvents, and extraction solvents. Energies 14, 4916.
Mikhail, S., Anand, A., Kannan, S., Raghavan, V., 2020. Bibliometric evaluation of research in hydrochar and bio-oil. J. Scientometr. Res. 9, 40–53.
Mikulčić, H., Baleta, J., Klemeš, J.J., Wang, X., 2021. Energy transition and the role of system integration of the energy, water and environmental systems. J. Clean. Prod. 292.
Moehrle, M.G., Caferoglu, H., 2019. Technological speciation as a source for emerging technologies. Using semantic patent analysis for the case of camera technology. Technol. Forecast. Soc. Change 146, 776–784.
Nurton, J., 2020. Patenting trends in renewable energy. WIPO Mag 50–56.
Okolie, J.A., Rana, R., Nanda, S., Dalai, A.K., Kozinski, J.A., 2019. Supercritical water gasification of biomass: a state-of-the-art review of process parameters, reaction mechanisms and catalysis. Sustain. Energy Fuels 3, 578–598.
Osman, A.I., Mehta, N., Elgarahy, A.M., Al-Hinai, A., Al-Muhtaseb, A.H., Rooney, D.W., 2021. Conversion of Biomass to Biofuels and Life Cycle Assessment: a Review, Environmental Chemistry Letters. Springer International Publishing.
Ozdenkci, K., Koskinen, J., 2017. Hydrothermal Conversion of Biomass. PCT/FI2017/050928.
Powell, J.B., Chheda, J.N., 2015. Biomass Treatment for Hydrothermal Hydrocatalytic Conversion. PCT/US2014/071039.
Roberts, V.M., Daage, M., Wang, K., 2014. Methods for Processing Biomass-Derived Feedstocks. Patentscope.
Sahoo, A., Saini, K., Jindal, M., Bhaskar, T., Pant, K.K., 2021. Co-Hydrothermal Liquefaction of algal and lignocellulosic biomass: status and perspectives. Bioresour. Technol. 342, 125948.
Saini, M., Chahal, K.K., Verma, R., Singh, A., 2020. Customer reviews as the measure of software quality. IET Softw 14, 850–860.
Scherer, F.M., 2014. First Mover Advantages and Optimal Patent Protection (No. RWP14-053), HKS Working Paper.
Sharma, H.B., Sarmah, A.K., Dubey, B., 2020. Hydrothermal carbonization of renewable waste biomass for solid biofuel production: a discussion on process mechanism, the influence of process parameters, environmental performance and fuel properties of hydrochar. Renew. Sustain. Energy Rev. 123, 109761.
Shlens, J., 2014. A Tutorial` on Principal Component Analysis. Retrieved from the arXiv database, Version 3.02.
Suominen, A., Toivanen, H., Seppänen, M., 2017. Firms' knowledge profiles: mapping patent data with unsupervised learning. Technol. Forecast. Soc. Change 115, 131–142.
Vimalnath, P., Tietze, F., Jain, A., Prifti, V., 2020. IP Strategies for green innovations - an analysis of European Inventor Awards. Cent. Technol. Manag. Work. Pap. Ser. 1–33.
Wang, L., Jiang, S., Zhang, S., 2020. Mapping technological trajectories and exploring knowledge sources: a case study of 3D printing technologies. Technol. Forecast. Soc. Change 161, 120251.
Wang, L. yin, Zhao, D., 2021. Cross-domain function analysis and trend study in Chinese construction industry based on patent semantic analysis. Technol. Forecast. Soc. Change 162, 120331.
WIPO, 2019. Guide to the International Patent Classification. IPC Publ.
Wu, J.L., Chang, P.C., Tsao, C.C., Fan, C.Y., 2016. A patent quality analysis and classification system using self-organizing maps with support vector machine. Appl. Soft Comput. J. 41, 305–316.
Yang, J., Hong, C., Xing, Y., Zheng, Z., Li, Z., Zhao, X., Qi, C., 2021. Research progress and hot spots of hydrothermal liquefaction for bio-oil production based on bibliometric analysis. Environ. Sci. Pollut. Res. 28, 7621–7635.
Yoon, B., Magee, C.L., 2018. Exploring technology opportunities by visualizing patent information based on generative topographic mapping and link prediction. Technol. Forecast. Soc. Change 132, 105–117.
Zhang, Y., Minaret, J., Yuan, Z., Dutta, A., Xu, C.C., 2018. Mild hydrothermal liquefaction of highwater content agricultural residue for bio-crude oil production: a parametric study. Energies 11.
Zhou, X., Huang, L., Porter, A., Vicente-Gomila, J.M., 2019. Tracing the system transformations and innovation pathways of an emerging technology: solid lipid nanoparticles. Technol. Forecast. Soc. Change 146, 785–794.
Zhu, D., Porter, A.L., Porter, A.L., 2002. Automated extraction and visualization of information for technological intelligence and forecasting. Technol. Forecast. Soc. Change 69, 495–506.