

Artigo

Panorama brasileiro da formação continuada de professores de química: a importância das redes de colaboração

Brazilian overview of continuing education of chemistry teachers: the importance of collaboration networks

Panorama brasileño de la formación continua de profesores de química: la importancia de las redes de colaboración

GlauCIA Ribeiro Gonzaga¹, Daniel Costa de Paiva², Marcelo Leandro Eichler³

Universidade Federal Fluminense (UFF), Santo Antônio de Pádua-Rio de Janeiro, Brasil.
Universidade Federal Fluminense (UFF), Santo Antônio de Pádua-Rio de Janeiro, Brasil.
Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre-Rio Grande do Sul, Brasil.

Resumo

A formação continuada do professor contribui de forma significativa para o seu desenvolvimento profissional, e isso se reflete nas suas atividades laborais de forma direta ou indireta. Dessa forma, esse processo formativo sendo rico e construtivo traz inúmeros benefícios para o docente e o público atendido por ele. Uma das formas de enriquecer esse processo de formação é através da troca de experiências e conhecimento estabelecidos nas relações pessoais e profissionais proporcionadas dentro de uma pós-graduação. Para estabelecer um panorama quantitativo sobre o desenvolvimento profissional do professor de Química, esta proposta tem por objetivo contribuir em um problema pouco estruturado, com grande envolvimento da comunidade científica que necessita de modelos que apoiem decisões. Para tal, mapeia grupos de pesquisa atuantes na formação continuada do professor de Química. Através de publicações, busca identificar possíveis redes de colaboração e compartilhamento de informações como fomento de discussões e relações científicas; através de bancos de dados de amplo acesso e conceitos básicos de Data Mining, processamento de linguagem natural (PLN) e teoria dos grafos, foram identificadas relações e proximidades. Redes de colaboração entre pesquisadores e nuvens de palavras dos termos de maior conexão/relevância nas publicações são sugeridas. As informações obtidas ilustram o cenário da pesquisa nacional sobre formação continuada do professor de Química, colabora com discussões no estabelecimento de perspectivas futuras, e também contribuem no fortalecimento das relações de colaboração dos pesquisadores.

¹ Doutora em Educação em Ciências pela Universidade Federal do Rio Grande do Sul (UFRGS). Professora Adjunta na Universidade Federal Fluminense (UFF), Rio de Janeiro, Brasil. Lattes: <http://lattes.cnpq.br/1055446047081201>. Orcid: <http://orcid.org/0000-0002-0552-3770>. E-mail: glauciargonzaga@gmail.com

² Doutor em Sistemas Eletrônicos pela Universidade de São Paulo (USP). Professor Associado na Universidade Federal Fluminense (UFF), Rio de Janeiro, Brasil. Lattes: <http://lattes.cnpq.br/9425925721844499>. Orcid: <http://orcid.org/0000-0002-0093-9902>. E-mail: profDanielpaiva@gmail.com

³ Doutor em Psicologia pela Universidade Federal do Rio Grande do Sul (UFRGS). Professor Associado na Universidade Federal do Rio Grande do Sul (UFRGS), Rio Grande do Sul, Brasil. Lattes: <http://lattes.cnpq.br/2483458145895884>. Orcid: <http://orcid.org/0000-0001-5650-9218>. E-mail: marcelo.eichler@ufrgs.br



Abstract

The continuing education of teachers contributes significantly to their professional development, and this is directly or indirectly reflected in their work activities. Thus, this training process, being rich and constructive, brings numerous benefits to the teacher and the public served by him. One of the ways to enrich this training process is through the exchange of experiences and knowledge established in the personal and professional relationships provided within a postgraduate course. To establish an overview of the professional development of the Chemistry teacher, this proposal aims to contribute to a poorly structured problem, with great involvement of the scientific community that needs models to support decisions. To this end, it maps research groups active in the continuing education of Chemistry teachers. Through publications, it seeks to identify possible networks of collaboration and information sharing as a way of promoting scientific discussions and relationships; through widely accessible databases and basic concepts of Data Mining, natural language processing (NLP) and graph theory, relationships and proximities were identified. Collaboration networks between researchers and word clouds of terms of greater connection/relevance in publications are suggested. The information obtained illustrates the scenario of national research on continuing education of the Chemistry teacher, collaborates with discussions in the establishment of future perspectives, and also contributes to the strengthening of collaborative relationships between researchers.

Resumen

La formación continua de los docentes contribuye significativamente a su desarrollo profesional, y esto se refleja directa o indirectamente en sus actividades laborales. Así, este proceso de formación, siendo rico y constructivo, trae numerosos beneficios al docente y al público por él atendido. Una de las formas de enriquecer este proceso formativo es a través del intercambio de experiencias y conocimientos que se establecen en las relaciones personales y profesionales que se brindan dentro de un posgrado. Para establecer un panorama del desarrollo profesional del docente de Química, esta propuesta pretende contribuir a un problema poco estructurado, con gran involucramiento de la comunidad científica que necesita modelos para sustentar decisiones. Para ello, mapea grupos de investigación activos en la formación continua de profesores de Química. A través de publicaciones, busca identificar posibles redes de colaboración e intercambio de información como una forma de promover discusiones y relaciones científicas; a través de bases de datos ampliamente accesibles y conceptos básicos de Minería de Datos, procesamiento de lenguaje natural (PNL) y teoría de grafos, se identificaron relaciones y proximidades. Se sugieren redes de colaboración entre investigadores y nubes de palabras de términos de mayor conexión/relevancia en las publicaciones. La información obtenida ilustra el escenario de la investigación nacional sobre la formación continua del profesor de Química, colabora con las discusiones en el establecimiento de perspectivas de futuro, además contribuye al fortalecimiento de las relaciones colaborativas entre investigadores.

Palavras-chave: Desenvolvimento profissional, Regime de colaboração, Produção do conhecimento, Análise da informação.

Keywords: Professional development, Collaboration, Knowledge production, Information Analysis.

Palabras claves: Desarrollo profesional, Régimen de colaboración, Producción científica, Análisis de información.



Introdução

A formação de professores voltada para a racionalidade crítica é fundamental para uma educação de qualidade social (SOARES *et al.*, 2012) – “um conceito associado às exigências de participação, democratização e inclusão, bem como à superação das desigualdades e das injustiças” (Gatti; Barreto; André, 2011, p. 38). Essa formação, além da busca da reconstrução da docência, deve envolver a reflexão no decorrer desse processo, e valorizar e favorecer as experiências prévias dos professores, e também as experiências construídas ao longo do processo formativo (Behrens; Fedel, 2020). Esses fatores, que tem relação direta com os meios acadêmico, científico e profissional, contribuem com a formação da identidade docente e estreita ações de colaboração entre profissionais e pesquisadores da educação (Iza *et al.*, 2014).

Essas ações de colaboração entre pesquisadores, através de formação conjunta e criação de núcleos de pesquisa, que contribuem para a construção da identidade docente e, conseqüentemente, da identidade do ensino de química, sob uma perspectiva crítico-reflexiva (Nóvoa, 1997; Imbernón, 2022), já vêm sendo apontadas como fundamentais ao desenvolvimento profissional (dos docentes formados e em formação) por autores como Schnetzler (2002), Echeverría e Belisário (2008) e Maldaner (2020).

Sabendo que o processo de Extração de Conhecimento em Bases de Dados (KDD, do inglês *Knowledge Discovery in Databases*) tem o objetivo de encontrar conhecimento a partir de um conjunto de dados para ser utilizado em processo decisório, o objetivo desse levantamento é compreender a evolução ocorrida no período de janeiro de 2010 a maio de 2021 no cenário da pesquisa em formação de professores, especificamente a formação continuada, e verificar as terminologias mais frequentes comunidade científica, neste mesmo intervalo de tempo, através de um mapeamento textual. Tal levantamento visa responder a seguinte questão de pesquisa: qual nível e a estrutura da rede de colaboração da comunidade científica de química que investiga formação continuada do docente de química? A escolha desse processo ocorreu dado o avanço promissor de pesquisas que utilizam a mineração de dados (*Data Mining*), que é uma técnica bastante utilizada para a realização de KDD, para analisar dados educacionais pontuais e para análise e definição de políticas públicas (Norton, 1999; Lopes *et al.*, 2007; Kampff; Reategui; Lima, 2008; Lorenzetti; Delizoicov, 2009; Oliveira, 2012; Namen; Borges; Sadala, 2013; Semaan *et al.*, 2019a,b). Concordando com Ball (2009, p. 312) quando afirma que “Sem a compreensão da situação, não temos nenhuma possibilidade de agir estrategicamente”, este artigo tem por objetivo contribuir em um problema pouco estruturado, com grande envolvimento da comunidade científica que necessita de modelos que apoiem decisões, evidenciando um panorama de produção científica nacional dos grupos de pesquisa sobre a formação continuada (desenvolvimento profissional) do professor de Química.

2. O Processamento de Linguagem Natural (PLN)

O PLN é um ramo multidisciplinar da Inteligência Artificial que agrupa Ciência da Computação, Linguística, Ciência Cognitiva, e trata de forma computacional diversos aspectos da comunicação em diferentes níveis de processamento: fonéticos/fonológicos, morfológicos, sintáticos, semânticos e

pragmáticos (Conteratto, 2006; Gonzalez; Lima, 2003; Barros; Robin, 2001). Cada nível possui características próprias, e suas aplicações variam com o objetivo da pesquisa desenvolvida. Para Liddy (2003) e Cimiano (2006), os objetivos mais comuns em PLN são: recuperar informações a partir de textos; qualificar o processo de traduções automáticas; proporcionar interpretação de textos; realizar inferências a partir de textos; construir ontologias. E segundo Pinheiro (2009, p. 04, colchetes nossos), o principal foco do PLN é o pré-processamento de dados, que correspondem a “seleção e filtragem de dados, limpeza de dados, normalização e *parsing* [análise sintática], análise semântica e representação numérica dos termos extraídos do documento em um vetor no espaço vetorial (BOW – *Bag-of-Words*)”.

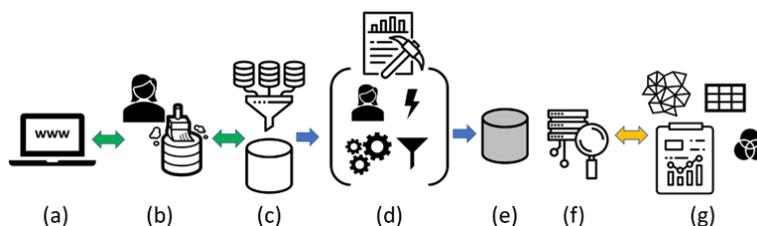
Uma das estratégias de processamento baseadas no conhecimento linguístico, empregada neste levantamento, é a normalização de variações linguísticas, que se divide em três casos: Normalização Morfológica (redução ao radical, ou redução dos verbos ao infinitivo, dos adjetivos e substantivos ao masculino singular), Sintática (frases com semanticamente iguais e sintaticamente diferentes são unificadas à frase de constituição mais frequente) e Léxico-Semântica (os termos são agrupados por similaridade semântica, sendo representados por um item lexical com conceito único). Outra estratégia adotada é a eliminação de *Stopwords* (preposições, conjunções, dentre outros termos que isoladamente não possuem relevância semântica). Uma das formas adotadas neste levantamento para reduzir perdas é a extração de termo compostos por duas (bigramas) ou três palavras (trigramas), além de palavras-chave isoladas.

Trabalhos como o desenvolvido por Acosta, Villegas e Norambuena (2019), que utilizam o PLN para propor um conjunto de heurísticas para aumento de dados na avaliação de resenhas de artigos científicos, indicam resultados positivos no atendimento dos objetivos e na validação da consistência dos dados. São inúmeros os benefícios da aplicação dos conceitos de PLN em diferentes tipos de pesquisas textuais.

3. A meta-análise do processo de Extração de Conhecimento em Bases de Dados

O processo de descoberta sistematizada do conhecimento através da análise de bancos de dados requer a execução de uma sequência de fases, que se inicia com a coleta das informações, o tratamento das mesmas seguindo critérios pré-estabelecidos, e a exposição e análise do resultado final. Um processo mais detalhado de KDD foi estruturado por Semaan e colaboradores (2019a) e adaptado para esta pesquisa (Figura 1), no qual existem 7 etapas (Semaan *et al.*, 2019a; Quoniam *et al.*, 2001; Fayyad; Piatetsky-shapiro; Smyth, 1996).

Figura 1 - Etapas do processo de KDD.



Fonte: adaptado de Semaan *et al.* (2019a).

A primeira etapa se inicia com a seleção da base de dados sólida e a seleção dos dados de pesquisa (a). Para uma melhor qualidade de dados e resultados, a massa dos dados precisa de um pré-processamento para eliminar dados inúteis à pesquisa (*Data Cleaning*) e inconsistências (b). Em sequência tem-se a tabulação dos dados pré-estabelecidos (c), com a padronização e o agrupamento dos dados que de fato contribuirão com os objetivos pré-estabelecidos. Em seguida, se inicia a segunda fase, que consiste na mineração de dados (*Data Mining*) (d), neste caso, na mineração de texto (*Text Mining*). A mineração busca padrões e relacionamentos entre os dados; e para tal, é necessário excluir *stopwords* e agrupar sinônimos (Aranha, 2007; Procaci *et al.*, 2015). Como esta é a etapa mais importante do processo, é necessário escolher bem o método, de acordo com os objetivos do levantamento. Existem basicamente 05 métodos de mineração (Dias, 2002): (1) Regra de associação, para problemas de associação entre dados, estabelecendo correlação estatística entre dados; (2) Árvore de decisão, para problemas de classificação e regressão, hierarquizando e categorizando dados em classes e subconjuntos; (3) Raciocínio baseado em casos, para problemas de classificação e segmentação, comparando e combinando atributos para hierarquizar dados próximos por semelhança; (4) Algoritmos genéticos, para problemas de classificação e segmentação, através da otimização inspirada na melhoria através da evolução; e (5) Redes neurais artificiais, para problemas de classificação e segmentação, para prever e classificar dados segundo um mapa de conexões neurais (não sendo muito eficiente para grande volume de dados).

Por fim, no pós-processamento, os dados são consolidados através da identificação de padrões apresentados na mineração, possibilitando a interpretação dos resultados obtidos (e). Esse conjunto de dados tratados (de publicação e de autores) foram unificados (f), para observar as possíveis correlações existentes. Essas correlações são observadas após a construção dos grafos de colaboração nos grupos e nas produções em coautorias (g) possibilitando identificar correlações não percebidas na análise superficial unitária dos dados iniciais.

Como apontado por Semaan e colaboradores (2019a,b), existem 4 dimensões comumente usadas para analisar acervos bibliográficos. As diferentes combinações entre elas, geram diferentes panoramas de relacionamento de dados: (i) tempo (ano, edição e/ou número da publicação); (ii) título e/ou palavras-chave; (iii) relações entre coautores; e (iv) citações relacionadas dos artigos. Essa análise de relacionamento de dados pode fornecer subsídios capazes de prever tendências e fomentar discussões a articulações (Dias; Pacheco, 2005). Além da adoção do processo de KDD, também são empregados os conceitos básicos PLN (Gonzalez; Lima, 2003),

Teoria dos Grafos (Procaci *et al.*, 2015; Magalhães *et al.*, 2013; Boaventura Netto, 2012; Newman, 2001) e estatísticas gerais.

Em linhas gerais, grafos são objetos matemáticos que se constituem por dois conjuntos de informações (vértices, que no caso deste artigo, são os pesquisadores que atuam com formação continuada de professores de Química; e arestas/arcos, que são as relações simétrica/assimétricas existentes entre esses vértices), representados por esquemas gráficos orientados ou não, para solucionar modelos (representações simplificadas de realidades) (Boaventura Netto, 2012).

4. Procedimento metodológico

Concordando com Chagas, Linhares e Mota (2019, p. 35), quando indicam que é importante “se pautar em novas práticas de pesquisa, para que se possa compreender criticamente a complexidade que envolve a realidade das pesquisas em educação [...]”, a proposta metodológica desta pesquisa reúne métodos emergentes para o campo da educação. Os procedimentos adotados nesta pesquisa seguem as etapas do processo de KDD (Figura 1).

Os bancos de dados selecionados na etapa (a) consistem no Diretório de Grupos de Pesquisa e o Portal de Periódicos da CAPES, a Plataforma *Lattes*, o *Google Scholar*, e as páginas oficiais dos eventos científico-acadêmicos e periódicos relacionados à produção dos autores/pesquisadores levantados. Esse levantamento inicial foi feito de forma manual (etapa A da Figura 1, com a presença humana).

Para o *data cleaning* (etapa (b)), os dados inúteis e repetições foram removidos manualmente através do uso de filtros e buscas por termos chave (“Formação de professores de Química”, “Desenvolvimento Profissional de professores de Química”, “professores de Química” e “Formação permanente de professores de Química”) no processo de levantamento.

A tabulação dos dados (etapa (c)) ocorreu em planilha do Excel®, agrupando dados relevantes para a pesquisa. Para cada grupo focal, foram levantados: (i) Para grupos de pesquisa (GP): ano de criação, quantidade de pesquisadores, quantidade de alunos, nome dos pesquisadores; (ii) Para publicação em eventos: ano, evento, tipo de publicação (trabalho completo, resumo expandido, resumo), título do trabalho e os nomes dos autores; (iii) Para artigos em periódicos: ano, periódico, título, resumo, palavras-chave, nomes dos autores.

Alguns desses dados tabulados foram empregados em análises estatísticas gerais da pesquisa, para gerar informações como frequência de termos chave (quais as palavras que mais se repetem em títulos e palavras-chave), média de publicação (quantidade de publicações anuais por autor, independentemente do tipo de publicação), frequência de coautoria (quantidade de coautores por trabalho publicado).

Antes do início da mineração (etapa (d) da Figura 1), os títulos e as palavras-chave foram fragmentados em palavras/termos isolados (*tokenização*), excluindo-se os *stopwords*. Esses termos selecionados foram consolidados através de normalização por redução ao radical ou equivalência ao termo mais frequente.

Após o processamento desse novo banco de dados (etapa (e) da Figura 1), é possível estabelecer frequência e relevância em relação aos demais termos

através de uma nuvem de palavras (construída no *Infogram*), bem como os autores que os utilizam e quando utilizaram (etapa (g) da Figura 1). Para os (co)autores, a verificação manual também foi necessária para averiguar a correspondência dos nomes dos autores, e atribuir um único identificador para cada indivíduo em autoria. Ainda na etapa (g) da Figura 1, foram identificadas colaborações e coautorias através da construção de grafos, onde cada pesquisador é um vértice, e a existência de colaboração entre pares é uma aresta. Foi considerada a frequência de produções entre os pares e, portanto, se trata de um grafo com peso nas arestas, onde o peso é determinado pela espessura da linha (quanto maior a espessura de linha, maior a quantidade de publicações na colaboração dos dois vértices/pesquisadores).

O grafo (não-orientado) foi estruturado segundo a centralidade de grau (*degree centrality*), que é definida pela quantidade de arestas que um vértice possui e representa o grau de influência desse vértice/pesquisador possui no meio analisado; segundo Borba e Trevisan (2014, p. 02) “um nó importante está conectado com muitos nós”. Esse grafo permitem uma análise visual mais efetiva das redes de colaboração da comunidade científica levantada, além da verificação de possíveis relações de dependência, hierarquias e relevâncias estruturadas pela frequência de publicação, os modelos e tipos de parcerias.

Apesar desses dados, métodos e técnicas aplicados, não há neste trabalho juízo de valor quanto às pessoas envolvidas nos dados obtidos. Os autores aqui se atêm à análise e interpretação do que os dados do período escolhido demonstram. Nesta direção, optou-se por não denominar pesquisadores e grupos, mas sim representá-los por codificações. Não há nenhuma outra intencionalidade além dos objetivos supracitados, se restringindo às informações públicas consultadas nos portais mencionados. O entendimento da área passa pela análise do seu retrato, aprendendo com indicadores, redes e relações.

5. Resultados

Após apresentar o cenário geral, essa seção tem estatísticas com os critérios estabelecidos e contribuições obtidas ao analisar os dados em diversas categorias.

5.1 O cenário geral

Entendendo que a atuação de um GP permite investigações coletivas que promovem o aprofundamento de determinado tema; proporciona a ampliação da compreensão do processo de pesquisa por seus participantes (principalmente os que estão em processo de formação); e permite que participantes desenvolvam uma reflexão crítica sobre a resolução de problemas práticos reais (ROSSIT *et al.*, 2018), foi realizado um levantamento de GPs com o mesmo alvo de interesse (linha de pesquisa, LP) desta pesquisa. Para tal optou-se por um levantamento no Diretório de Grupos de Pesquisa (DGP) no Brasil, da CAPES, sobre os GPs que atuam os quatro descritores apresentados na etapa (b), apresentada na metodologia.

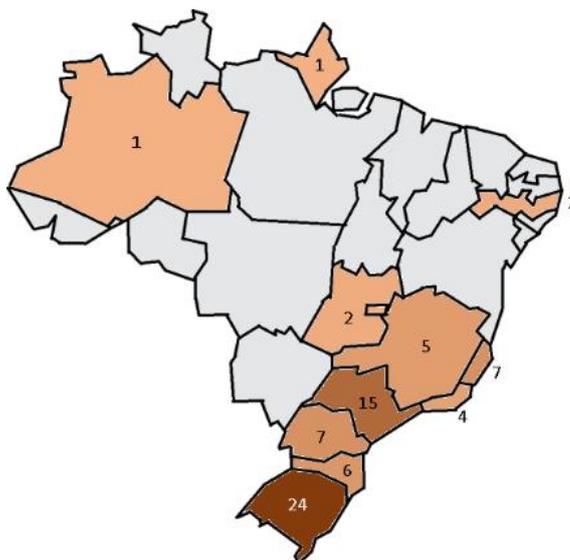
Os descritores foram utilizados em uma busca com todas as palavras, e refinado em uma busca exata, visto que a busca parametrizada por Ensino de Química retornou 346 LPs de 213 GPs. Foram ativados os filtros “Nome do

grupo”, “Nome da linha de pesquisa”, e “Palavra-chave da linha de pesquisa” para que a busca também ocorresse com esses parâmetros. Esse levantamento e triagem ocorreu de forma manual, verificando cada um dos resultados que se enquadrassem nos parâmetros estabelecidos. Após a aplicação dos filtros de pesquisa, foram obtidos 17 GPs, que possuem ao todo 146 pesquisadores. Destes, apenas 53% foram relacionados a GP/LP que atuam em formação continuada do professor de Química. Para este artigo, todas as informações apresentadas dizem respeito às produções desses 77 pesquisadores entre janeiro de 2015 e maio de 2021.

5.2 Os pesquisadores

Nesta e nas próximas seções são discutidos os pesquisadores, os tipos de produção, os títulos e palavras-chave utilizados, as redes de coautorias e os grafos de colaboração. Cada categoria tem seus critérios elencados, os dados obtidos, as análises e pontos de destaque em relação à relevância daquele aspecto para os pesquisadores e grupos de trabalho em formação de professores. As 44 pesquisadoras e 33 pesquisadores possuem diferentes níveis de formação, sendo 4 graduados com cursos de especialização, 2 mestrandos, 8 mestres, 6 doutorados, e 57 doutores. Mesmo se tratando de GPs que envolvem formação continuada do professor de Química, muitos tratam do Ensino de Química dentro de Ciências Naturais, e dentro da grande área de Ciências Exatas, o que pode explicar a participação de 11 matemáticos dentre os 77 pesquisadores identificados. Em relação ao local de formação (Figura 2), é possível fazer algumas inferências quanto a representatividade regional nas pesquisas desenvolvidas.

Figura 2 - Dispersão da formação dos pesquisadores no Brasil.



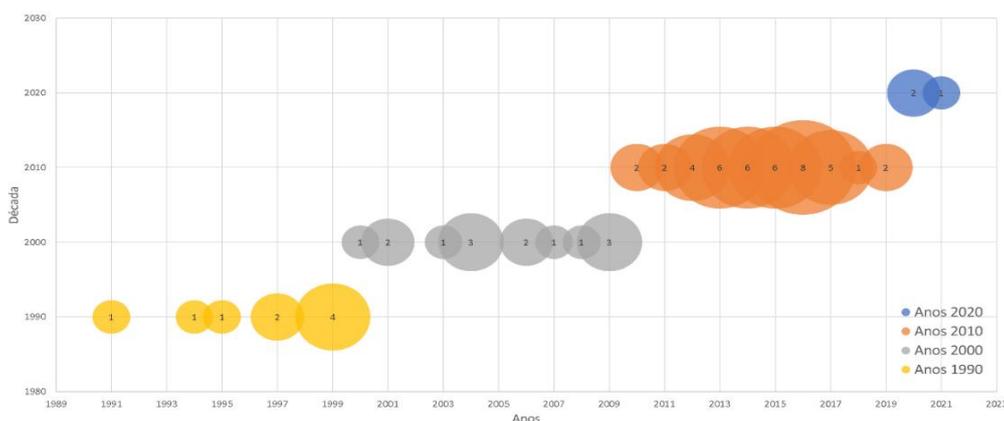
Fonte: os autores.

Analisando regionalmente, nota-se que 48% (37 indivíduos) dos pesquisadores deste levantamento se formaram na região Sul, e juntamente com os 40% (31 indivíduos) formados na região Sudeste, evidenciam a discrepância

científica em território nacional, já conhecida, discutida e questionada no meio acadêmico (Tunes, 2017; Gonzaga, Paiva, Eichler, 2019, 2020). Além destes apresentados, há 2 pesquisadores cujos doutorados foram cursados integralmente em instituições internacionais (Itália e Portugal).

Com relação aos anos de formação, considerando uma década o intervalo de tempo entre o início do ano 0 e o final do ano 9, a variação entre as décadas de 1990, 2000, 2010 e 2020 é crescente e ampla (Figura 3). A maior quantidade de pesquisadores identificados tem sua última titulação na década pós 2010. Para a representação, os profissionais com formação em andamento (9 pesquisadores) não foram contabilizados.

Figura 3 - Dispersão por ano de formação dos pesquisadores dos GPs.



Fonte: os autores.

Sobre as áreas de formação, após a normalização (Figura 4), foi possível quantificar os docentes como pode ser visto na Figura 5.

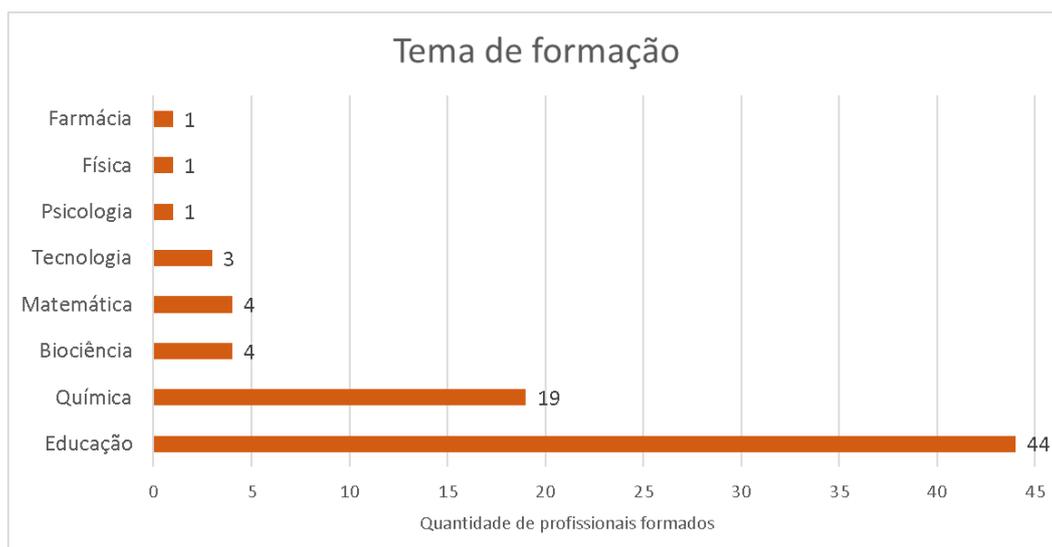
Figura 4 - Mapa de normalização das áreas de formação.



Fonte: os autores.

Em uma correlação dos dados da Figura 5, com a área da última formação dos pesquisadores após a normalização, é possível perceber que as duas maiores áreas de formação estão relacionadas com a área de atuação do GP dos quais participam. Porém, existem áreas diretamente correlacionadas com a atuação do professor de Química, como por exemplo, Informática na Educação. Os pesquisadores com esta formação foram enquadrados na área normalizada de Tecnologia. Cabe a ressalva da presença de pesquisadores de outras áreas, que de alguma forma podem contribuir ativa e efetivamente com uma equipe multidisciplinar, agregando conhecimentos de suas áreas de formação (Farmácia, Física, Psicologia, Matemática e Biociência).

Figura 5 - Pesquisadores por área normalizada de formação.

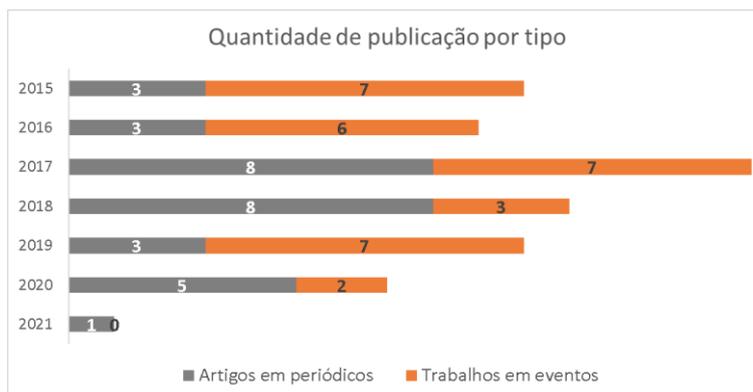


Fonte: os autores.

5.3 Volume das produções

Após a coleta manual e análise dos 77 currículos Lattes, foram identificados, em uma janela de 6,5 anos, 82 publicações dentro do foco de interesse deste levantamento. Na organização desses trabalhos foram considerados dois critérios de exclusão: 1) produções citadas em mais de um currículo, considerar apenas uma delas; 2) produções que não tiverem suas versões completas disponíveis para acesso gratuito online. Aplicando estes critérios, foram descartadas 11 pelo critério 1 de exclusão e 8 pelo critério 2. Ao final foram eleitas 63 publicações, sendo 31 artigos em periódicos acadêmico-científicos e 32 trabalhos publicados em eventos (entre trabalho completo, resumo expandido e resumo), como consta na Figura 6.

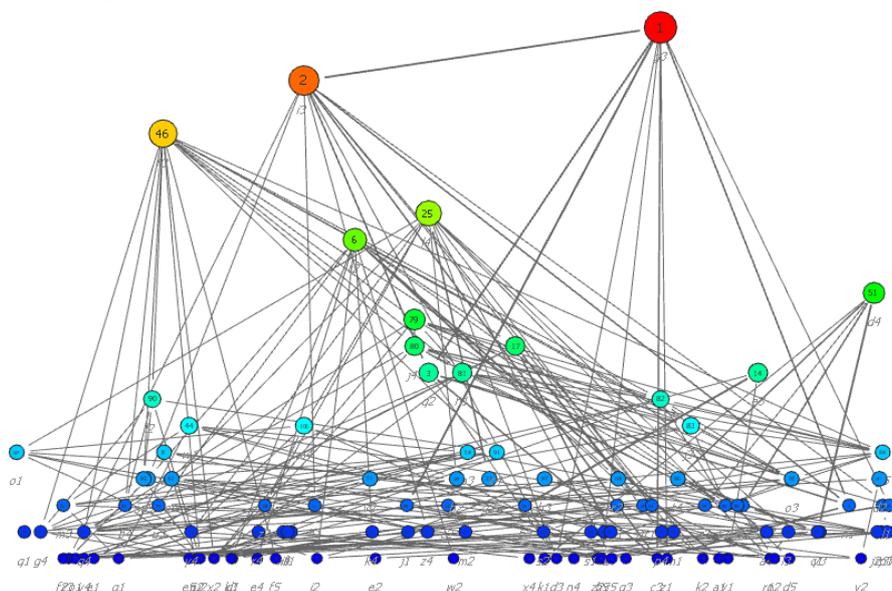
Figura 6 - Produção por ano em periódicos e eventos acadêmicos.



Fonte: os autores.

A Figura 7 representa a relação entre os autores investigados, cujo as produções estão contabilizadas na Figura 6.

Figura 7 - Grafo de coautoria das publicações analisadas.



Fonte: os autores.

O total de publicações na Figura 6, na janela temporal adotada, equivale a média de 5 artigos e 5 trabalhos em evento por ano; é um universo muito pequeno, dada a quantidade de pesquisadores nas LPs dos GPs cadastrados em território nacional. Relacionando a quantidade de publicação (82) com a quantidade de pesquisadores (77), há a constatação importante de uma média inferior a 1 publicação por pesquisador em um período de 6,5 anos. Este número sofre influência, além do baixo índice de publicações com essa temática, da não atualização frequente do currículo, do tempo para avaliação, parecer e publicação dos textos, mas trata-se de uma informação relevante mesmo com todos estes fatores.

O grafo de coautoria da Figura 7 indica pesquisadores (codificados) e suas relações, nas publicações levantadas. Quanto mais acima se encontra o

pesquisador, maior a quantidade de colaboradores relevantes e maior a influência desse pesquisador, dentro de seu meio no tema levantado (Formação Continuada do professor de Química). Nota-se duas particularidades nesse grafo: (i) a quantidade de arestas de maior espessura, indicando a baixa frequência de colaboração com os mesmos pares, porém, com diversificação de pares, e para outros o equilíbrio dentre quantidade de colaborações e frequência de publicações dentro de uma mesma colaboração; (ii) as relações de colaboração entre os pesquisadores para essa temática necessitam maior estreitamento, para que o grafo apresente teias mais relevantes, evidenciando trocas científicas e construções de conhecimento que seriam favorecidas pelas experiências e distintas formações dos integrantes dos GPs levantados.

5.4 Dos títulos e termos

Após organização e normalização dos títulos e termos das produções autodeclaradas, foram elencadas 348 expressões, que se reduzem (retirando as repetições) a 192. É possível representar visualmente uma frequência global de uso das expressões em destaque (Figura 8). Nesta nuvem, com exceção do termo “Formação continuada” (devido ao alto índice de incidência, e para dar maior legibilidade aos demais termos), estão representados todos os termos, após a exclusão das *stopwords*, mesmo que tenham sido empregados apenas uma vez.

Colaboração	4	1,1	-	-	-	-	-	0,3	0,9
Curso	4	1,1	-	-	0,3	0,3	0,3	-	0,3
Ensino de Química	4	1,1	-	0,3	0,3	0,3	0,3	-	-
Portugal	4	1,1	-	-	-	-	0,6	0,3	0,3
Tecnologias	4	1,1	-	-	-	0,3	0,6	-	0,3

Fonte: os autores.

Sobre as palavras-chave, 81% das publicações seguem a recomendação padrão de elencar de 3 a 4 termos que sintetizem o trabalho. Porém há uma variação de 0 (4 publicações) a 7 palavras-chave (1 publicação), que podem ser observadas em suas incidências anuais na Tabela 2.

Tabela 2. Quantidade de trabalhos por palavras-chave por ano.

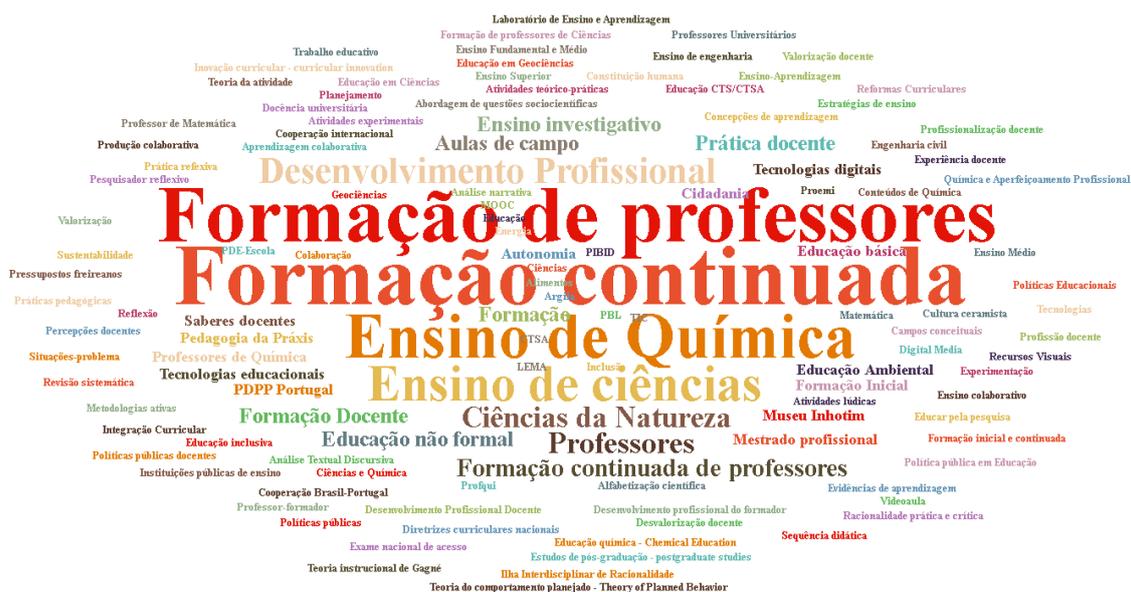
Quantidade	2021	2020	2019	2018	2017	2016	2015	TOTAL
Sem palavras	-	-	-	-	1	3	-	4
3 palavras	-	4	5	7	7	5	5	33
4 palavras	1	3	5	3	3	1	2	18
5 palavras	-	-	-	-	4	-	3	7
7 palavras	-	-	-	1	-	-	-	1
TOTAL	1	7	10	11	15	9	10	63

Fonte: os autores.

Não há um padrão de comportamento na frequência de publicações de acordo com a quantidade de palavras-chave. Porém, desde 2018, as publicações seguem o padrão estabelecido, de 3 a 4 termos (cujas incidências aparecem de forma variável em todos os anos da janela temporal deste levantamento).

As 4 publicações que não apresentam palavras-chave, não fazem parte da próxima análise. Somam-se 213 termos, que, após unificação por equivalência à palavra semelhante mais frequente, se reduzem a 126 termos, dispostos na nuvem de palavras da Figura 9, sem omissões, mesmo o termo de aparição única.

Figura 9 - Nuvens de termos usados como palavras-chave.



Fonte: os autores.

Em uma análise anual de uso do termo, a tabela 3 apresenta as 10 expressões que aparecem ao menos três vezes, com seus percentuais de frequência anual e total. Nota-se pelos quantitativos de repetições que há uma grande variedade de termos empregados pelos pesquisadores (desvio padrão maior na tabela 1 que os dados da tabela 3), tornando o conjunto de publicações mais fluido e difuso. Essa flutuação também pode ser notada pela dispersão das frequências de uso pelos anos.

Tabela 3. Frequência das palavras-chave nas publicações.

Termo	repet	Freq total	2021	2020	2019	2018	2017	2016	2015
			%						
Formação continuada	15	7,0	-	0,5	1,4	0,9	1,4	-	2,8
Formação de professores	14	6,6	0,5	1,4	0,9	0,9	0,9	0,5	1,4
Ensino de Química	11	5,2	-	0,9	0,5	1,9	-	0,5	1,4
Ensino de Ciências	9	4,2	-	0,9	1,9	0,5	0,5	-	0,5
Desenvolvimento profissional	6	2,8	-	-	-	-	1,9	0,9	-
Ciências da Natureza	5	2,3	-	-	0,5	0,5	-	0,5	0,9
Professores	5	2,3	-	-	0,9	-	1,4	-	-
Formação continuada de professores	4	1,9	-	-	0,5	0,5	0,5	-	0,5

Aulas de campo	3	1,4	-	-	-	-	0,9	-	0,5
Educação não formal	3	1,4	-	-	-	-	0,9	-	0,5

Fonte: os autores.

Ao fazer uma análise comparativa das expressões mais empregadas nos títulos e nas palavras-chave, o maior destaque, como esperado pelos termos de filtro, é “formação continuada”, seguido por “professores de Ciências” (18 repetições) em títulos, e “formação de professores” (14 repetições) em palavras-chave. A associação entre os 10 primeiros termos de cada caso demonstra que nos títulos, os termos escolhidos versam mais sobre o professor de Química/Ciências e sua formação/desenvolvimento profissional, enquanto que para as palavras-chave, há um equilíbrio entre a formação e a atuação docente.

Ao contrário das seções de análise anteriores, o levantamento dos termos no título e palavras-chave não indicou um direcionamento necessário seja para aprendizado, seja de grande contribuição na área. A diversidade das abordagens é considerada aqui como adequada à área sem um aspecto necessário para ressaltar.

6. Considerações finais

Como colocado nos objetivos e propósitos desta pesquisa, o intuito desta demonstração não é a discussão a respeito da formação continuada do professor de Química, nem sobre sua importância, histórico ou propósitos. Nosso objetivo foi demonstrar como conceitos básicos de Data Mining, processamento de linguagem natural (PLN) e teoria dos grafos, de forma associada, são capazes de demonstrar as relações (ou não relações) existentes entre pesquisadores e pesquisas de uma determinada temática (neste caso, pesquisadores que atuam com formação continuada do professor de Química). Acreditamos que este conjunto de métodos é capaz de auxiliar tomadas de decisão, e estreitar caminhos para estabelecimento de parcerias e criação de grupos de pesquisa (ou trabalho colaborativo entre eles); e que este é um dos pontos cruciais para melhorar a formação crítica do docente, e seu pleno desenvolvimento profissional, tanto almejado pelos pesquisadores ora mencionados.

Dito isso, o primeiro ponto a ser apresentado nas considerações deste artigo é com relação à escolha e uso dos métodos de levantamento e análise de dados não ter sido anteriormente utilizado em estudos sobre grupos de pesquisa em Educação em Química associado à Formação de Professores. Esse tipo de análise, além de apresentar um modelo que apoia a decisão dentro deste tema pesquisado, é capaz de identificar lacunas existentes e as possibilidades de ação em pesquisas, a fim do fortalecimento de áreas e de Grupos de Pesquisa com interesses comuns. Observar e conhecer o cenário no qual se atua é fundamental para uma atuação mais assertiva e direcionada em pontos científicos estratégicos capazes de contribuir de forma efetiva para a construção e desenvolvimento profissional do professor de química, além de fortalecer a comunidade química (inter)nacionalmente.

A partir de um levantamento da área de formação de professores, esta pesquisa considera grupos de pesquisa, linhas de pesquisa, produções, colaborações e região geográfica e traz uma importante análise. A temática a

qual essa pesquisa se debruçou apresenta baixa quantidade de publicações, diminuindo as possibilidades de diálogo com os pares. Dados dos títulos e palavras-chave demonstram a predominância de enfoques e uma análise de parcerias sugere mais abordagens inter e multidisciplinares com a colaboração de mais pesquisadores do próprio programa, mas também outros alunos e outros parceiros.

As regiões Sul e Sudeste são predominantes em formação de profissionais que atuam na/para a formação continuada de professores de Química, e em volume de publicações. Esse é um fato que a vivência profissional já indicava, e já era esperado; mas, dois aspectos suscitam maior atenção dos pesquisadores. O primeiro ponto tem relação com o baixo volume de produções por pesquisador considerando a janela de 6,5 anos, afinal, mesmo considerando publicação em periódico e em anais de eventos, a média abaixo de 1 é preocupante. E o segundo ponto tem relação com a preferência por determinadas expressões e termos nos títulos e palavras-chave.

Destaca-se que diferentes abordagens nos trabalhos é um ponto relevante identificado nesta pesquisa, assim como a presença de duas titulações obtidas no exterior e a diversidade de áreas de formação dos profissionais que atuam nos cursos de formação de professores.

Como já indicado por Maldaner (2020) e Imbernón (2022), por exemplo, é importante para a formação continuada do professor de química que este tenha acesso a possibilidade de expandir não só seu conhecimento específico, mas também socializar profissionalmente com outros professores para expandir suas construções acadêmicas, científicas, sociais, dentre outras. Esta troca de experiências ajuda a consolidar o conhecimento construído e possibilita o trabalho coletivo e cooperativo entre esses professores e suas respectivas instituições através de “redes de colaboração”. Para que esta rede aconteça, é importante que os pesquisadores que atuam diretamente com esta formação docente também estabeleçam e fortaleçam suas redes colaborativas.

Como sugestões para melhorar os indicadores que foram evidenciados nesta pesquisa, é necessário que todos os profissionais envolvidos nas áreas acadêmico-científicas de ensino de química tenham atenção ao relato e compartilhamento de suas produções, a adequada atualização dos currículos, ao incentivo aos seus discentes para a atuação com ensino, pesquisa e extensão, de modo que haja integração dos diversos níveis de formação e também parcerias com outros pesquisadores da mesma instituição, de outras instituições e até mesmo de outras áreas e países.

Há grande trabalho a ser feito. Aqui foi identificado um primeiro passo, com único propósito de colaborar com a qualidade da área de formação de professores sem qualquer identificação de pesquisador ou grupo de pesquisa; afinal, os aprendizados necessários passam pela análise do cenário atual.

Referências

ACOSTA, Rubén Sánchez; VILLEGAS, Claudio Meneses; NORAMBUENA, Brian Keith. Heurísticas para data augmentation en nlp: Aplicación a revisiones de artículos científicos. **RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação**. n. 34, p.44-53, 2019.

ARANHA, Christian Nunes. **Uma abordagem de pré-processamento automático para mineração de textos em português: sob o enfoque da inteligência computacional.** [Doutorado em Engenharia Elétrica], Pontifícia Universidade Católica do Rio de Janeiro, 2007.

BALL, Stephen J. Um diálogo sobre justiça social, pesquisa e política educacional. **Revista Educação e Sociedade**. v. 30, n. 106, p. 303-318, 2009. DOI: 10.1590/S0101-73302009000100015

BARROS, Flávia de Almeida; ROBIN, Jacques. Processamento de linguagem natural. SBC: **Revista Eletrônica de Iniciação Científica**, Porto Alegre, v. 1, n. 2, p. 1-61, 2001.

BEHRENS, Marilda Aparecida; FEDEL, Tiago Reus Barbosa. Os contributos da reflexão e da experiência vivenciada na formação continuada de professores. **REVEDUC - Revista Eletrônica de Educação**, v. 14, p. 1-13, e3009045, 2020. DOI: <https://doi.org/10.14244/198271993009>

BOAVENTURA NETTO, Paulo Oswaldo. **Grafos: teoria, modelos, algoritmos**. 5.ed. revisada e ampliada. Rio de Janeiro: Blucher, 2012.

BORBA, Elizandro Max; TREVISAN, Vilmar. Medidas de Centralidade em Grafos e aplicações em redes de dados. **Proceeding Series of the Brazilian Society of Applied and Computational Mathematics**, v. 2, n.1, p. 010088-1 - 010088-6, 2014.

CHAGAS, Alexandre Menese; LINHARES, Ronaldo Nunes; MOTA, Marilton Fontes. A curadoria de conteúdo digital enquanto proposta metodológica e multirreferencial. **RISTI**, n. 33, p. 32-47, 2019. DOI: 10.17013/risti.33. 32–47

CIMIANO, Philipp. **Ontology learning and population from text: Algorithms, evaluation and applications**. Heidelberg: Springer-Verlag, 2006.

CONTERATTO, Gabriela Betania Hinrichs. Semântica e Computação: uma interação necessária para o aperfeiçoamento de sistemas PLN. **Letras de Hoje**. Porto Alegre, v. 41, n. 2, p. 353-367, 2006.

DIAS, Maria Madalena. Parâmetros na escolha de técnicas e ferramentas de mineração de dados. **Acta Scientiarum. Technology**, Maringá, v. 24, n. 6, p. 1715-1725, 2002. DOI: 10.4025/actascitechnol.v24i0.2549

DIAS, Maria Madalena; PACHECO, Roberto Carlos dos Santos. Uma metodologia para o desenvolvimento de sistemas de descoberta de conhecimento. **Acta Scientiarum. Technology**. Maringá, v. 27, n. 1, p. 61-72, 2005. DOI: 10.4025/actascitechnol.v27i1.1500

ECHEVERRÍA, Agustina Rosa; BELISÁRIO, Celso Martins. Formação inicial e continuada de professores num núcleo de pesquisa em ensino de ciências. **Revista Brasileira de Pesquisa em Educação em Ciências**, v. 8, n. 3, 2008.

FAYYAD, Usama M.; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. From data mining to knowledge discovery in databases. **AI Magazine, American Association for Artificial Intelligence**, v. 17, n. 3, p. 37-54, 1996.

GATTI, Bernardete Angelina; BARRETO DE SÁ, Elba Siqueira; ANDRÉ, Marli Eliza Dalmazó de Afonso. **Políticas docentes no Brasil: um estado da arte**. Brasília: UNESCO, 2011.

GONZAGA, Glaucia Ribeiro; PAIVA, Daniel Costa; EICHLER, Marcelo Leandro. Habilidades tecnológicas de professores de Química em desenvolvimento profissional: um estudo de caso no âmbito do Mestrado Profissional em Química em Rede Nacional - PROFQUI. **Revista De Estudos E Pesquisas Sobre Ensino Tecnológico**, v. 5, p. 11-29, 2019.

GONZAGA, Glaucia Ribeiro; PAIVA, Daniel Costa; EICHLER, Marcelo Leandro. Desafios e perspectivas atuais na formação do professor de Química: expectativas sobre o Mestrado Profissional em Química em Rede Nacional (PROFQUI). **Química Nova (ONLINE)**, v. 43, p. 493-505, 2020.

GONZALEZ, Marco; LIMA, Vera Lúcia Strube. Recuperação de informação e processamento da linguagem natural. In: **XXIII Congresso da Sociedade Brasileira de Computação**. v. 3, p. 347-395, 2003.

IMBERNÓN, Francisco. **Formação docente e profissional: formar-se para a mudança e a incerteza**. Cortez editora, 2022.

IZA, Dijnane Fernanda Vedovatto; *et al.* Identidade docente: As várias faces da constituição do ser professor. **REVEDUC - Revista Eletrônica de Educação**, v. 8, n. 2, p. 273-292, 2014 DOI: <https://doi.org/10.14244/19827199978>

KAMPFF, Adriana Justin Cerveira; REATEGUI, Eliseo Berni; LIMA, José Valdeni de. Mineração de dados educacionais para a construção de alertas em ambientes virtuais de aprendizagem como apoio à prática docente. **Novas Tecnologias na Educação Rio Grande do Sul, Brasil**, v. 6, n. 2, 2008. DOI: 10.22456/1679-1916.14394

LIDDY, Elizabeth D. Natural Language Processing. In: **Encyclopedia of Library and Information Science**, 2.ed. New York: Marcel Decker, Inc., 2003.

LOPES, Cristiane B.; *et al.* Identificação das características associadas com a aprovação de candidatos de escolas públicas e privadas: Vestibular-2004 UFMG. **Educação em Revista**, v. 46, p. 67-194, 2007, DOI: 10.1590/S0102-46982007000200006

LORENZETTI, Leonir; DELIZOICOV, Demétrio. La producción académica brasileña en Educación Ambiental. **Utopía y Praxis Latinoamericana**, v.14, n. 44, p. 85-100, 2009.

MAGALHÃES, Cleyton Vanut Cordeiro de; *et al.* Caracterizando a Pesquisa em Informática na Educação no Brasil: Um Mapeamento Sistemático das

Publicações do SBIE. In: **XXIV Simpósio Brasileiro de Informática na Educação - SBIE**. v. 24, n. 1, p. 22, 2013.

MALDANER, Otavio Aloisio. **A formação inicial e continuada de professores de química: professores/pesquisadores**. 4 ed. Ijuí: Ed. UNIJUI, 2020

NAMEN, Anderson Amendoeira; BORGES, Sonia Xavier de Almeida; SADALA, Maria da Glória Schwab. Indicadores de qualidade do ensino fundamental: o uso das tecnologias de mineração de dados e de visões multidimensionais para apoio à análise e definição de políticas públicas. **Revista Brasileira Estudos Pedagógicos**, v. 94, n. 238, p. 677-700, 2013.

NEWMAN, Mark E.J. (2001). The structure of scientific collaboration networks. **Proceedings of the national academy of sciences**, v. 98, n. 2, p. 404-409, 2001.

NORTON, Melanie J. Knowledge Discovery in databases. **Library Trends**, v. 48, n.1, p. 9-21, 1999.

OLIVEIRA, Edvaldo Artmann. **Sobre a colaboração na comunidade de sistemas de informação através dos simpósios SBSI**. (Mestrado em Informática), UNIRIO - Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, 2012.

PINHEIRO, Marcello Sandi. **Uma abordagem usando sintagmas nominais como descritores no processo de mineração de opiniões**. (Doutorado em Engenharia Civil) – COPPE, Universidade Federal do Rio de Janeiro, 2009.

PROCACI, Thiago; *et al.* Estudo Exploratório das Produções e Colaborações entre Pesquisadores em Informática na Educação: uma Análise de Publicações do Simpósio Brasileiro de Informática na Educação de 2001 a 2013. In: **XXIV Simpósio Brasileiro de Informática na Educação - SBIE**. V. 26, n. 1, p. 1323, 2015.

ROSSIT, Rosana Aparecida Salvador; *et al.* Construção da identidade profissional na Educação Interprofissional em Saúde: percepção de egressos. **Interface - Comunicação Saúde Educação**, v. 22, n. 4, p. 1-12, 2018. DOI: 10.1590/1807-57622017.0184

QUONIAM, Luc; *et al.* Inteligência obtida pela aplicação de Data Mining em Bases de teses Francesas sobre o Brasil. **Ciência da Informação**, Brasília, v. 30, n. 2, p. 20-28, 2001.

SEMAAN, Gustavo Silva; *et al.* Um panorama geral das contribuições e das redes de colaboração do SBIE. In: **Simpósio Brasileiro de Informática na Educação – SBIE**. v. 30, n. 1, p. 1361-1379, 2019a.

SEMAAN, Gustavo Silva; *et al.* Um panorama dos trabalhos do SIMPEP: 25 anos de contribuições. In: **Simpósio de Engenharia de Produção - SIMPEP**. 2019b.



SOARES, Márton Herbert Flora Barbosa; *et al.* A formação de professores de Química pela pesquisa: algumas ações da área de Ensino de Química do Instituto de Química da Universidade Federal de Goiás. **Espaço Plural**, v. 13, n. 26, p. 70-87, 2012.

TUNES, Regina Helena. O reforço às desigualdades regionais no Brasil no século XXI: concentração espacial do processo de aprendizagem e da produção inovadora. **Confins - Revista Franco-Brasileira de Geografia [Online]**, n. 32, p. 1-16, 2017. DOI: 10.4000/confins.12257