

Resumo: Este estudo reinterpreta a história da bibliometria, das origens documentais aos sistemas algorítmicos contemporâneos, para analisar como a dependência de trajetória moldou as métricas atuais e produziu a inércia avaliativa. Caracteriza-se como um ensaio teórico-histórico fundamentado em revisão integrativa, articulando obras fundadoras e estudos de Ciência e Tecnologia. Os resultados evidenciam que a trajetória do campo é marcada por sedimentação, demonstrando que vieses da era do *mainframe* persistem e que a disseminação tecnológica falhou em promover diversidade epistemológica. Atualmente, a fragmentação via API e altmetria coexiste com lógicas hierárquicas herdadas. Conclui-se que o sistema permanece vinculado à sua trajetória original, fenômeno sustentado pelos processos de inércia epistemológica e atrofia interpretativa, que reforçam a primazia da contagem sobre a interpretação do contexto científico.

Palavras-chave: Bibliometria; Dependência de trajetória; Inércia epistemológica; Infraestruturas do conhecimento.

Abstract: This study reinterprets the history of bibliometrics, from its documental origins to contemporary algorithmic systems, to analyze how path dependence shaped current metrics and produced evaluative inertia. It is a theoretical-historical essay based on an integrative review, articulating foundational works and science and technology studies. Results show that the field's trajectory is marked by sedimentation, demonstrating that mainframe-era biases persist and that technological dissemination failed to foster epistemological diversity. Currently, fragmentation through APIs and altmetrics coexists with inherited hierarchical logics. It is concluded that the system remains tied to its original trajectory, a phenomenon sustained by epistemological inertia and interpretive atrophy, which reinforce the primacy of counting over the interpretation of scientific context.

Keywords: Bibliometrics; Path dependence; Epistemological inertia; Knowledge infrastructures.

Introdução

A onipresença de algoritmos que hoje classificam a ciência representa o capítulo mais recente de uma longa busca pela totalização do conhecimento. Na virada do século XIX para o XX, essa aspiração manifestou-se em duas frentes principais, abrangendo, de um lado, o projeto do *Repertório Bibliográfico Universal* (RBU), estruturado pela *Classificação Decimal Universal* (CDU) de Paul Otlet e Henri La Fontaine, que visava ordenar documentos sistematicamente (RAYWARD, 1997). No outro lado, encontravam-se as tentativas de Francis Galton e James McKeen Cattell de mensurar a eminência intelectual humana, deslocando o foco do documento para o produtor da ciência (GODIN, 2006). Galton, com seu *English Men of Science* (1874), e Cattell, com o diretório *American Men of Science* (1906), sistematizaram o uso de dados biográficos e *rankings* de mérito para quantificar a produtividade científica (GODIN, 2006). Embora distintos, esses projetos partilhavam a crença de que o conhecimento poderia ser administrado por meio

de tecnologias de registro, seja a ficha catalográfica de Otlet ou o verbete biográfico de Cattell, as quais já antecipavam os limites e as possibilidades dos sistemas digitais futuros.

Contudo, a explosão informacional do pós-guerra expôs a incapacidade dessas tecnologias manuais de lidarem com o volume crescente da ciência, exigindo novas soluções para a recuperação da informação. Nesse novo contexto, emergiram duas propostas fundamentais, herdeiras da visão de totalidade anterior: o Memex de Vannevar Bush (1945), um dispositivo orgânico focado na associação cognitiva (RAYWARD, 1997), e o *Science Citation Index* (SCI) de Eugene Garfield (1955), uma solução formal baseada em citações. Enquanto o Memex permaneceu no campo visionário, o SCI consolidou-se como a infraestrutura operacional dominante (KHADEMI, 2018). Sua implementação, contudo, dependia da capacidade de processamento dos *mainframes* da época, computadores de grande porte operados via fitas magnéticas e cartões perfurados, o que forçou a substituição da classificação temática subjetiva por uma análise de conexões puramente numérica e independente da linguagem (KESSLER, 1963).

Este percurso valida o pressuposto central deste estudo ao considerar que a bibliometria funciona como uma infraestrutura do conhecimento (EDWARDS, 2010). Conforme Star e Ruhleder (1996), tais sistemas não são suportes neutros para a prática científica, mas forças que moldam ativamente o que pode ser visto, medido e considerado como conhecimento válido (BOWKER e STAR, 1999). A evolução desses indicadores exhibe uma forte dependência de trajetória (DAVID, 1985), conceito que explica como decisões tecnológicas contingentes do passado restringem as escolhas presentes. Essa dinâmica frequentemente leva a um aprisionamento tecnológico (ARTHUR, 1989) que se cristalizou por meio de uma história humana, marcada pela ascensão e reconfiguração de diferentes atores socioprofissionais. Nesse sentido, o desenvolvimento de serviços *online* é fundamentalmente uma história de pessoas (HAHN, 1996:33), concentrando-se precisamente na intersecção entre artefatos técnicos e agência humana.

Para tornar legível esse percurso de continuidades e rupturas (DALBELLO, 2015), a análise estrutura-se em quatro atos que revelam a coevolução entre agência humana e infraestrutura técnica. Embora delimitados por balizas didáticas (1895, 1960, 1985, 2005), esses marcos não são cortes rígidos, mas lógicas que admitem sobreposições (LYU *et al.*, 2023). O primeiro ato (1895-1960) aborda a hegemonia dos documentalistas e a organização universal via fichas e classificação (RAYWARD, 1997). O segundo (1960-1985) detalha como o *mainframe* e os índices de citação converteram o bibliotecário em um intermediário de sistemas restritos (WOUTERS, 1999). A ruptura seguinte (1985-2005) expõe a descentralização do acesso via CD-ROM e *Web*, exigindo uma postura educativa ante a sobrecarga informacional (TENOPIR, 1991). Finalmente, a era atual (2005-presente) analisa como algoritmos e plataformas redefinem a visibilidade científica, impondo ao bibliotecário o papel de guardião ético contra a governação métrica (MAROLDI, 2025).

A partir desse percurso, emerge a questão central do estudo: por que a disseminação tecnológica, do CD-ROM à *Web*, falhou em produzir diversidade epistemológica, reforçando, ao contrário, a lógica quantitativa herdada do *mainframe*? Para responder a essa questão, o artigo busca (a) reinterpretar a história da bibliometria a partir de sua materialidade técnica e dos atores envolvidos e (b) demonstrar como vieses infraestruturais do passado persistem nas formas contemporâneas de avaliação científica.

Ainda que a bibliometria histórica, conforme propõe Hérubel (1999), busque compreender a evolução das disciplinas científicas e comunidades intelectuais por meio de indicadores e análises de citação, este manuscrito realiza o movimento inverso: em vez de aplicar as métricas à história, aplica a história e a teoria à própria bibliometria, conduzindo uma arqueologia de seus fundamentos conceituais e técnicos. Para os propósitos desta análise, entretanto, adota-se a perspectiva de que a bibliometria e seus desdobramentos posteriores, cientometria, infometria, webometria e altmetria, não são campos estanques, mas expressões evolutivas de uma mesma trajetória sociotécnica de mensuração da ciência (THOMPSON e WALKER, 2015).

Metodologicamente, o estudo assume a forma de um ensaio histórico-interpretativo de orientação sociotécnica, fundamentado em uma revisão crítica da literatura. Os princípios da revisão integrativa (WHITTEMORE e KNAFL, 2005) são mobilizados de modo instrumental, funcionando como diretriz para a articulação de fontes de natureza distinta, sem a adoção de um protocolo sistemático formal. A seleção dos textos baseou-se em levantamentos exploratórios e em referências clássicas da literatura métrica e documental, privilegiando o diálogo com as perspectivas dos estudos de Ciência e Tecnologia (STS) para explorar criticamente a materialidade técnica desses campos. Tal estratégia analítica permite examinar a coevolução entre infraestruturas técnicas, práticas documentais e formas de avaliação científica ao longo do desenvolvimento histórico da bibliometria.

A gênese da infraestrutura métrica: da classificação semântica à era do mainframe

O primeiro ato (1895-1960), a era pré-computacional, foi marcado pelo domínio da materialidade impressa e pela gestão manual da informação. Embora a formalização conceitual do campo bibliométrico tenha ocorrido apenas em 1969 (PRITCHARD, 1969), esse período já possuía uma cultura de quantificação e sistematização. Práticas de mensuração da produção científica circulavam sob designações como *statistical bibliography* (HULME, 1923) e por meio de métodos pioneiros de análise de citações (GROSS e GROSS, 1927) e dispersão documental (BRADFORD, 1934), enquanto a compilação de estatísticas de coleções já existia para fins de gestão bibliotecária (BROADUS, 1987). Contudo, essas iniciativas eram manuais, limitadas pelo suporte físico, e o campo era predominantemente demarcado pelo documentalista e bibliotecário. De fato, alguns dos principais autores de estudos quantitativos, como Hulme em 1923 e Bradford em 1934, eram os bibliotecários-chefes de suas respectivas instituições, e os precursores da análise de citação, como Fussler em 1949, também tinham sua origem na Biblioteconomia (BROADUS, 1987). Nessa era, contudo, uma das principais autoridades intelectuais desses profissionais não residia na contagem, mas na mestria da classificação semântica, uma lógica documental baseada no conteúdo temático, cujo projeto mais ambicioso foi o Mundaneum de Paul Otlet, sustentado pela *Classificação Decimal Universal* (RAYWARD, 1997).

Industrializando essa trajetória, a era do *mainframe* (aproximadamente entre 1960 e 1985) institucionalizou as práticas de quantificação e formou a matriz técnico-discursiva da bibliometria moderna. A ordem anterior, baseada na primazia do conteúdo semântico, foi abalada pela introdução de um novo paradigma técnico e epistemológico. O advento dos computadores de grande porte e a demanda por sistemas capazes de lidar com o volume

crescente da produção científica, uma pressão impulsionada pelo colapso do sonho Alexandrino das bibliotecas de serem totalmente autossuficientes (BROOKES, 1988), criaram as condições para substituir a lógica semântica pela lógica relacional, onde a citação substitui o assunto como unidade de conexão (DE BELLIS, 2009).

Foi nesse contexto que Eugene Garfield (1955) concebeu o *Science Citation Index* (SCI), vislumbrando o uso de automação mecânica, inicialmente cartões perfurados, para viabilizar a indexação de citações em larga escala, uma tarefa impossível para a indexação humana tradicional (DE BELLIS, 2009). Contudo, ele não estava sozinho nessa convergência tecnológica. Como aponta Dalbello (2015), os líderes de iniciativas que moldaram o cenário da informação nas décadas de 1950 e 1960 foram Fred Kilgour (diretor fundador da OCLC) e Eugene Garfield. O movimento foi duplo, pois enquanto Garfield tornava visíveis as estruturas da ciência através de abrangentes índices de citação, Kilgour fazia o mesmo com os catálogos de bibliotecas, ambos utilizando a nova infraestrutura computacional para estruturar o conhecimento (DALBELLO, 2015).

O SCI deslocou o centro de gravidade da análise do conteúdo temático para as conexões formais entre os documentos, permitindo análises independentes da interpretação semântica direta e das ambiguidades da linguagem natural. Essa nova abordagem quantitativa, focada em padrões de frequência e *ranking* (distribuições altamente assimétricas), exigiu também uma diferenciação metodológica das estatísticas tradicionais (DE BELLIS, 2009). Como observado por Brookes (1988), em distribuições de frequência por posto, a ordem dos elementos é determinada pela própria contagem e não por uma escala inerente e independente, tornando inadequadas muitas ferramentas estatísticas convencionais (Gaussianas) para lidar com a bibliometria (BROOKES, 1988).

Essa transformação técnica ensejou uma aliança estratégica: de um lado, Eugene Garfield forneceu a infraestrutura de dados e as ferramentas de quantificação (SCI); do outro, a Sociologia da ciência, via Robert K. Merton, ofereceu o enquadramento teórico necessário para legitimar a prática. Como aponta De Bellis (2009), o edifício filosófico da bibliometria moderna repousa na teoria normativa da ciência mertoniana, na qual o ato de citar é interpretado como um átomo de revisão por pares, validando a contagem de citações como uma medida objetiva de reconhecimento e impacto. Juntos, esses atores redefiniram a avaliação científica, deslocando-a da leitura subjetiva para um problema de mapeamento e análise de redes de citações. Essa mudança resultou em uma clara transferência de autoridade epistêmica, ou seja, a capacidade de definir o que conta como qualidade científica passou das mãos dos especialistas no conteúdo para os especialistas na contagem (WOUTERS, 1999).

Além dessa reconfiguração de poder, a inovação de Garfield (1955) deu origem a um novo modelo de negócio baseado em métricas: o Institute for Scientific Information (ISI). Fundada em 1960, essa empresa privada transformou a estrutura da comunicação científica num produto de informação comercializável, criando um mercado competitivo baseado em revistas principais (WOUTERS, 1999; GUÉDON, 2001). Foi justamente nesse período, com a prática quantitativa já consolidada pelo SCI e suas profundas consequências estruturais se manifestando, que o campo ganhou sua formalização conceitual.

Quase simultaneamente à consolidação comercial do ISI, surgiram termos para nomear essa nova abordagem quantitativa. Em 1969, Alan Pritchard propôs o termo bibliometria para substituir a antiga bibliografia estatística, definindo-o como a aplicação de métodos

matemáticos e estatísticos aos meios de comunicação (PRITCHARD, 1969). No mesmo ano, Robert Fairthorne refinou o conceito como o tratamento quantitativo das propriedades do discurso registrado (FAIRTHORNE, 1969). Paralelamente, na União Soviética, Vassily Nalimov apresentava o termo cientometria, focando na análise da ciência vista como um processo informacional (GRANOVSKY, 2001). O léxico do campo se expandiria uma década depois com a introdução da infometria por Otto Nacke (NACKE, 1979). Essas terminologias batizaram a prática que Garfield havia iniciado e que já moldava a ciência. Esse período não apenas viu a criação de ferramentas fundadoras, mas também a consolidação da bibliometria como campo de estudo, registrando um crescimento de dez vezes em sua literatura especializada entre 1960 e 1980 (PERITZ, 1988).

Essa combinação, a aliança sociotécnica, a transferência de autoridade epistêmica e o modelo de negócio do ISI, solidificou o SCI como a peça arquitetônica central na infraestrutura da ciência (WOUTERS, 1999). É a partir da compreensão do SCI como a fundação dessa nova infraestrutura (disciplinar, comercial e técnica) que podemos desvendar a origem da dependência de trajetória do campo. Uma vez estabelecido, o SCI alcançou rapidamente o *status* de invisibilidade e transparência típico das infraestruturas maduras (STAR e RUHLER, 1996), embutindo-se nas rotinas diárias de avaliação e descoberta científica.

Contudo, essa nova tecnologia de *mainframe* operava sob severas limitações materiais. As restrições impostas pelas interfaces de entrada e armazenamento, como cartões perfurados e fitas magnéticas, geravam o que Edwards (2010) denomina atrito de dados, ditando decisões de *design* duradouras que resultaram em uma estrutura de dados extremamente rígida, baseada em campos fixos e identificadores formais. Como aponta Edwards (2010), os sistemas de informação dessa era eram limitados pela lentidão e custo do processamento mecânico, forçando uma abordagem onde os dados deviam se conformar às limitações da máquina, e não o contrário. Essa rigidez técnica favoreceu uma concepção modular e descontextualizada da literatura científica, na qual cada artigo se tornava uma unidade discreta e contável, desvinculada de sua complexidade semântica original. Conforme argumentam Bowker e Star (1999), toda infraestrutura de informação incorpora uma ontologia implícita, um modo de classificar e ordenar o mundo que, ao se tornar invisível com o uso, define o que pode ou não ser visto como dado válido.

Essas decisões de *design*, embora tecnicamente justificadas pelas limitações de processamento da época, não foram politicamente neutras, exemplificando a tese de que os artefatos têm política (WINNER, 1980). Como analisa Guéron (2001), a decisão pragmática de Garfield de focar em um conjunto selecionado de periódicos teve consequências teóricas profundas ao colapsar os múltiplos núcleos de especialidades em um único grande núcleo científico; Garfield deu substância e realidade à noção de ciência central. O que era originalmente uma ferramenta útil para a gestão de coleções foi reificado como um conceito genérico com reivindicações de universalidade, construindo uma visão de ciência hierárquica e enviesada desde a sua gênese. Tal estrutura lógica atribuiu, por definição, menor centralidade ao conhecimento produzido fora do eixo euro-americano e da língua inglesa. Van Leeuwen *et al.* (2001) demonstram o impacto duradouro desse viés, evidenciando que publicações em línguas diferentes do inglês, mesmo quando cobertas pelo SCI, recebem consideravelmente menos citações, o que dilui artificialmente os indicadores de impacto de nações não anglófonas.

Uma vez consolidada, essa infraestrutura gerou um mecanismo de 'aprisionamento tecnológico' (ARTHUR, 1989) que, impulsionado pela liderança inicial do SCI, criou retornos crescentes de adoção e transformou o conceito de revistas principais em um mercado inelástico de difícil substituição (GUÉDON, 2001; STAR e RUHLER, 1996). Diante da complexidade técnica dessa nova base instalada, onde os primeiros sistemas *online* como DIALOG e ORBIT operavam via interfaces hostis e linguagens de comando complexas, o bibliotecário reemergiu não mais como arquiteto semântico, mas como o intermediário técnico indispensável (HAHN, 1996). Foram estes profissionais que, armados de capacidade de navegação e de zelo, perseverança e charme, tornaram-se os mediadores obrigatórios capazes de superar a resistência inicial e conectar as novas infraestruturas digitais aos cientistas (HAHN, 1996).

A era do cd-rom e da Web 1.0 (aproximadamente 1985-2005): a disseminação e a consolidação da bibliometria

Se a era do *mainframe* estabeleceu a matriz técnico-discursiva da bibliometria, consolidando o SCI como infraestrutura central, porém de acesso restrito e mediado, a era seguinte (aproximadamente 1985-2005) marcou a difusão dessas ferramentas. Impulsionada pela combinação do computador pessoal com o CD-ROM, essa transição tecnológica buscou reduzir o atrito no acesso aos dados e descentralizar a análise. No início dos anos 1980, embora estudos via sistemas *online* conectados a *mainframes* já fossem possíveis (HAWKINS, 1977), tratava-se de um uso tecnicamente complexo e caro, o que gerou demanda por soluções que permitissem a extração e o tratamento local de dados em microcomputadores (MOED, 1988).

A dependência de trajetória consolidada pelos *mainframes* reconfigurou a ideia de documento como unidade processável, e a chegada do CD-ROM respondeu a essa demanda de manuseio local. Visto como um novo papiro (CHISHTI, 1993), o formato alterou radicalmente a forma como bibliotecas geriam coleções e ofereciam acesso, tornando economicamente viável a criação de bases de dados locais e personalizadas para análise bibliométrica (BURNHAM *et al.*, 1992). Entretanto, sob a superfície inovadora e a aparente autonomia do usuário, persistiam os traços da era anterior: o mesmo modelo centralizado de indexação proprietária e a mesma lógica de quantificação, agora apenas encapsulados em discos plásticos em vez de fitas magnéticas. Tal continuidade entre os regimes de organização documental e as métricas de citação expressa o que Star e Ruhleder (1996) denominaram infraestruturas invisíveis, ou seja, arranjos classificatórios que, ao se tornarem transparentes pelo uso cotidiano, moldam silenciosamente e naturalizam o modo como o conhecimento é produzido e legitimado.

A migração do SCI para o formato CD-ROM foi decisiva, pois mais do que democratizar o acesso, ela alterou o regime de prática ao transformar o pesquisador, antes objeto de estudo ou cliente de buscas tarifadas, em um praticante ativo (CHISHTI, 1993). Pela primeira vez, esse ator podia manipular as mesmas ferramentas dos especialistas de forma autônoma e com capacidade de processamento local, livre das restrições de tempo e custo por conexão que caracterizavam os sistemas *online* (CHISHTI, 1993; TENOPIR, 1991). A análise de citações saía do regime de tempo real tarifado dos sistemas online e entrava no domínio do tempo de bancada do pesquisador, constituindo-se como uma prática individual e laboriosa

que se tornava viável através do *download* de dados para *softwares* de gerenciamento pessoal (BURNHAM *et al.*, 1992; CHISHTI, 1993).

Essa nova capacidade de análise autônoma marcou a ascensão do pesquisador como usuário final (LYON, 1984). A redefinição do papel do bibliotecário nesta era, ao transitar de operador técnico para mediador epistêmico, ilustra a previsão central de teóricos da automação daquele período, como Gorman (1980) que antecipou que a tecnologia inevitavelmente erodiria a estrita divisão entre serviços técnicos e públicos e forçaria o especialista a assumir papéis mais integrados. Essa mediação tornou-se crucial porque a democratização via CD-ROM foi puramente técnica, e não epistemológica ao gerar tensões sobre a qualidade da busca do usuário final (LYON, 1984). Embora agora tivessem a ferramenta, os usuários ainda enfrentavam dificuldades metodológicas e cognitivas com a lógica de busca e a cobertura das bases (CHISHTI, 1993). Assim, o bibliotecário foi redefinido como o instrutor da lógica do sistema, evidenciando que o CD-ROM descentralizou a ferramenta, mas não o conhecimento para usá-la, mantendo assim intacta a infraestrutura classificatória herdada da era do *mainframe*.

Enquanto essa reconfiguração profissional redefinia as fronteiras entre mediação e autonomia, no campo técnico emergiam novas formas de organização e exploração dos dados bibliográficos. Pesquisadores começaram a combinar a extração de dados, inicialmente de bases *online* e posteriormente de CD-ROM, com *softwares* de computador pessoal para realizar seus próprios estudos bibliométricos (BURNHAM, SHEARER e WALL, 1992). Ao estruturar dados brutos em bancos relacionais customizados, esses pioneiros buscavam contornar as limitações de software dos grandes sistemas centrais, permitindo uma limpeza e manipulação de dados inviável nos *hosts* comerciais (MALLIG, 2010).

Contudo, o formato físico revelava o paradoxo da época, pois ainda que descentralizasse o acesso, mantinha a centralização lógica do modelo proprietário, com limites rígidos de atualização e armazenamento. Como nota Chishti (1993), o CD-ROM sofria com o atraso inerente à masterização física (atualizações trimestrais x mensais do *online*) e apresentava diferenças críticas na cobertura, como a exclusão de dissertações e variações nos motores de busca em comparação com seus equivalentes *online*. Além disso, a estrutura dos dados muitas vezes era diferente, como apontam Costas e Iribarren-Maestro (2007) que identificaram que versões em CD-ROM e *Web* da mesma base (SCI) apresentavam inconsistências significativas no tratamento de afiliações institucionais e pontuação, complicando a normalização necessária para a bibliometria. A ampliação do acesso estava, portanto, inscrita em uma economia política da informação científica onde o conhecimento circulava sob estritas mediações comerciais. O CD-ROM, entretanto, não rompeu, mas marcou a continuidade de um mercado global de informação, no qual o acesso às ferramentas de mensuração e a própria definição de revistas principais tornaram-se indissociáveis da lucratividade institucional e da criação de mercados inelásticos (GUÉDON, 2001).

Essa falha em traduzir a disseminação técnica em diversidade metodológica é o cerne do paradoxo desta era. O aumento da disponibilidade de dados não apenas falhou em produzir diversidade cognitiva (GUÉDON, 2001), como gerou o oposto: uma convergência e homogeneização do que se entendia por impacto e qualidade. Neste sentido, propomos nomear esse fenômeno de inércia epistemológica, entendida como um processo sociotécnico pelo qual uma infraestrutura de conhecimento, mantém e reforça sua lógica

avaliativa original, mesmo após migrar para novas plataformas como o CD-ROM que, teoricamente, permitiriam maior flexibilidade.

Essa inércia se manifesta, primeiro, no nível da própria infraestrutura, onde o dado democratizado era, em si, uma réplica da infraestrutura centralizada do *mainframe*, carregando todos os seus vieses de cobertura (geográficos, disciplinares, linguísticos) e sua lógica classificatória (BALL, 2017). Mais do que isso, as limitações físicas de armazenamento dos discos ópticos forçaram uma seleção ainda mais restritiva de periódicos em comparação à versão completa do banco de dados, exacerbando a exclusão da ciência periférica (COSTAS e IRIBARREN-MAESTRO, 2007). Não se tratava, portanto, de dados novos ou diferentes, mas dos mesmos dados, filtrados pela mesma lógica de ciência central e reembalados em um novo suporte.

Ademais, a inércia epistemológica consolidou-se através de um mecanismo cognitivo individual. Aparentemente libertadora, a autonomia recém-adquirida pelo pesquisador funcionou, na prática, como o principal vetor para o que propomos chamar de atrofia interpretativa. Este mecanismo define-se pela internalização, talvez inadvertida, da lógica da ferramenta, ao ser capacitado para contar facilmente as citações no computador pessoal, o praticante tornou a complexa tarefa de interpretar o contexto comunicativo metodologicamente secundária. Tal processo pavimentou o caminho para o que Cronin (2005) descreveu como a transformação da citação, que deixa de ser um gesto retórico de reconhecimento para se tornar um número dotado de valor próprio e descolado de seu contexto. A atrofia interpretativa é, portanto, o processo que permite a instalação desse fetiche métrico, e o produto em CD-ROM foi seu catalisador, garantindo que o reempacotamento digital apenas reforçasse a lógica da quantificação

Essa lógica quantitativa, reforçada e disseminada pelas bases de dados em CD-ROM, atingiu sua maturidade com a transição para a *Web* 1.0 no final dos anos 1990. Esse novo ambiente *online*, no entanto, apresentava desafios significativos de recuperação da informação; como apontam Ding *et al.* (2000), o volume crescente de dados trouxe 'sérios problemas que os pesquisadores de recuperação da informação nunca enfrentaram décadas atrás, como a sobrecarga informacional e a incerteza do vocabulário. Em resposta a esse cenário, surgiram propostas de sistemas que buscavam aplicar resultados de pesquisas bibliométricas diretamente na interface de busca, usando mapas de cocitação para ajudar os usuários a visualizar a estrutura intelectual de um domínio e refinar suas consultas (DING *et al.*, 2000). Se, por um lado, a bibliometria era usada experimentalmente para tentar reorganizar a recuperação da informação na *Web* aberta, por outro, seu uso avaliativo tradicional se consolidava em plataformas proprietárias como o Web of Science. Como nota Atkins (1999), foi a confluência do hipertexto com os navegadores *web* que permitiu tornar a indexação de citações conceitualmente acessível e intuitiva. Essa nova plataforma, contudo, não substituiu imediatamente o CD-ROM; ambos coexistiram como versões distintas da mesma base de dados, criando uma tensão metodológica sobre qual versão, a local e estática ou a *online* e dinâmica, deveria ser a fonte autorizada para a avaliação científica (COSTAS e IRIBARREN-MAESTRO, 2007).

Se o CD-ROM oferecia autonomia de processamento local (embora com dados estáticos e limites de atualização), a *Web* 1.0 trouxe a atualização constante, mas retomou o modelo de acesso como serviço, retirando do pesquisador o controle físico sobre os dados brutos e submetendo-o a um regime de licenciamento (GUÉDON, 2001; COSTAS e IRIBARREN-MAESTRO, 2007). Essa nova interface, exemplificada pela migração para o Web of Science,

tornou a busca de citações um ato instantâneo e, pela primeira vez, conceitualmente intuitivo para o usuário não especialista, mas ao custo de tornar novamente invisível a complexidade algorítmica do sistema (ATKINS, 1999).

Paralelamente, a consolidação dos modelos de licenciamento em massa, conhecidos como *Big Deals*, reforçou a centralização nas mãos das grandes editoras e criou um efeito catraca onde as bibliotecas ficavam presas a pacotes cada vez mais caros e abrangentes (GUÉDON, 2001). Para Guédon (2001:47), as plataformas das editoras funcionam como lentes distorcidas que produzem um "astigmatismo informacional", pois quanto mais amplo o acesso via *Big Deal*, mais as citações se concentram nas publicações da própria editora dominante, o que reforça artificialmente sua centralidade e padroniza a visibilidade.

A era da Web 2.0 e das API (aproximadamente 2005-2025): a fragmentação e a resistência

Enquanto a *Web 1.0* havia consolidado o poder de um modelo centralizado, a consolidação da *Web 2.0*, a partir de meados dos anos 2000, iniciou um movimento de força contrária e de fragmentação. Mas essa nova fase não foi apenas técnica, ela também expressou uma resistência às infraestruturas centralizadas de avaliação e visibilidade científica. Esse período pode ser entendido como o advento de uma quantificação 2.0, impulsionada por um “cenário de informação em constante expansão que se intensificou com os avanços da tecnologia digital” (GHOLAMPOUR *et al.*, 2024:1.671). Tal crescimento da publicação desencadeou uma produção de dados bibliométricos sem precedentes, que se tornou progressivamente “desacoplada de questões de utilidade e usabilidade, transformando os próprios dados em um fim autossuficiente” (KRÜGER, 2020:62).

Diferente de sua predecessora, a *Web* de apenas leitura, a *Web 2.0* é fundamentalmente uma *Web* de leitura-escrita e social, caracterizada pela colaboração e pelo conteúdo gerado pelo usuário (BASKARAN, 2019). De fato, as plataformas e ferramentas desta era passaram a oferecer interfaces que permitem até mesmo a “leigos, como gestores de pesquisa e formuladores de políticas, fazer avaliações por conta própria” (KRUGER, 2020:61). Essa mudança consolidou o gestor de pesquisa como o primeiro ator desta era, armado com soluções de inteligência de pesquisa que criaram um ecossistema fértil para novos concorrentes. Esse movimento desafiou a hegemonia de quatro décadas da infraestrutura bibliométrica tradicional (ISI), inaugurando o que Meho e Yang (2007) chamaram de uma nova era na análise de citação e bibliometria, marcada pela necessidade de usar múltiplas fontes para corroborar o impacto.

Essa fragmentação das fontes de dados foi acompanhada por uma diversificação das ferramentas de análise que impulsionaram a ascensão do segundo ator desta era: o tecnólogo/programador. O aumento da disponibilidade de dados, facilitado enormemente pelas API (*Application Programming Interfaces*), permitiu o surgimento de um ecossistema de *softwares* especializados. Ferramentas como VOSviewer e CiteSpace, cuja proeminência cresce visivelmente a partir de 2010 e que hoje dominam, respectivamente, a visualização de redes e a detecção de tendências, — passaram a incorporar a capacidade de realizar consultas diretas a algumas API, democratizando a realização de análises bibliométricas complexas (MALIHA, 2024; PUENTE *et al.*, 2024), passaram a incorporar

a capacidade de realizar consultas diretas a algumas API, ao democratizar a realização de análises bibliométricas complexas (MORAL-MUÑOZ *et al.*, 2020).

Contudo, como ressaltam Torres-Salinas e Arroyo-Machado (2022), é nos ambientes de programação como R e Python que estas interfaces alcançam seu potencial máximo, permitindo automatizar fluxos de trabalho e conectar fontes heterogêneas através de identificadores persistentes (como o DOI). Isso criou uma nova divisão de trabalho e uma ruptura com a era anterior: enquanto ferramentas da era do usuário final, como o BibExcel, eram projetadas para pré-processar dados para planilhas ou outros softwares, a nova geração, exemplificada pelo pacote Bibliometrix para R, exige que o ator domine a linguagem de programação para análises abrangentes, embora tente mitigar essa barreira através de interfaces *Web* como o Biblioshiny para usuários não programadores (MALIHA, 2023:2).

Essa integração técnica concentrou todo o fluxo de trabalho — da recuperação à análise, em um único ambiente, diluindo a dependência de múltiplas aplicações isoladas. A centralidade das API tornou-se tamanha que elas são hoje conceptualizadas como uma linguagem de domínio específico (DSL) para a cienciometria, permitindo a criação de estudos complexos independentemente das interfaces gráficas proprietárias. Essa realidade consolida uma nova divisão do trabalho: os praticantes (gestores) e os investigadores especializados (tecnólogos), capazes de conduzir os cálculos avançados exigidos pela ciência de dados.

Paralelamente a essa diversificação de ferramentas e atores, outro sintoma da fragmentação foi a emergência de concorrentes diretos da plataforma Web of Science (WoS). O primeiro foi a Scopus, da Elsevier (ROTH, 2005), e o segundo, mais disruptivo, o Google Scholar (GS) (MEHO e YANG, 2007). Contudo, essas plataformas oferecem visões ontológicas distintas: enquanto o GS tende à exaustividade ao indexar a *Web* acadêmica, a WoS e a Scopus mantêm-se seletivas e curadas, sendo a WoS historicamente a mais restritiva (HARZING e ALAKANGAS, 2015). Como a sobreposição de registros entre elas é apenas parcial, a escolha da fonte passou a determinar os resultados. Essa competição, ampliada por novas plataformas como o Dimensions, tornou metodologicamente insustentável a dependência de uma única fonte de dados para a avaliação científica.

Em contraste com essas plataformas comerciais e proprietárias, a resistência ao modelo assumiu uma forma ainda mais radical com o surgimento de infraestruturas de dados bibliográficos totalmente abertas, algo inédito até então na bibliometria. Iniciativas como o OpenAlex, construídas sobre dados abertos de fontes como o Crossref e o extinto Microsoft Academic Graph, representam uma tentativa de 'reconfigurar o aprisionamento' ao criar uma alternativa gratuita e aberta às plataformas comerciais (PRIEM, PIWOWAR e ORR, 2022; KLEBEL e ROSS-HELLAUER, 2023). Diferentemente da Scopus (concorrente comercial) ou do Google Scholar (uma ferramenta gratuita, mas com dados proprietários e sem API pública), estas novas infraestruturas visam criar um bem comum digital, onde os dados de citação possam ser acessados, reutilizados e analisados por qualquer pessoa, sem restrições (PRIEM, PIWOWAR e ORR, 2022).

Simultaneamente, a ascensão da *Web* 2.0 deu origem ainda a novos campos métricos, como a webometria, focada na análise quantitativa da própria *Web* (THELWALL, 2008), e, notavelmente, a altmetria. Esta última pode ser interpretada como a resposta sociológica mais direta ao mecanismo da atrofia interpretativa. Definida como um 'contraponto

revolucionário às métricas tradicionais' (GHOLAMPOUR *et al.*, 2024:1.672), a altmetria emergiu ao propor a medição do impacto para além das 'meras citações', capturando novos rastros digitais em redes sociais e blogues (GHOLAMPOUR *et al.*, 2024). Ela representa um esforço para reintroduzir o contexto comunicativo e social que a simplificação das métricas tradicionais havia eliminado (MAROLDI, 2025).

Essa diversificação reflete uma expansão metodológica mais ampla, na qual os *frameworks* tradicionais, baseados exclusivamente na contagem simples de citações, foram gradualmente enriquecidos com técnicas avançadas, como a normalização por campo e a busca por indicadores mais robustos diante das limitações de métricas como o índice h (BREDAHL, 2022; MAROLDI, 2025). Essa busca por análises mais sofisticadas tem sido crescentemente alimentada pela incorporação de técnicas de Inteligência Artificial (IA), que passam a atuar como mediadoras analíticas e metodológicas, sendo esta a tendência emergente significativa do campo (PUENTE *et al.*, 2024). A IA é empregada tanto na análise, em tarefas como classificação de funções de citação e predição de impacto (ZHANG *et al.*, 2025), quanto na coleta, com ferramentas que realizam a triagem de artigos de modo mais eficiente e imparcial (STEFANIS *et al.*, 2025).

Contudo, a pressão por quantificação gerou consequências perversas (lei de Goodhart). Já em 1974, debatia-se como indicadores oficiais poderiam deslocar os objetivos da ciência (CSISZAR, 2020), materializando a transformação do artigo de unidade de conhecimento em unidade contábil de avaliação (GINGRAS, 2016). Economicamente, isso sustentou o modelo de Taxas de Processamento de Artigos (APC), criando uma 'barreira-APC' que estratifica a publicação e exacerba desigualdades globais ao favorecer instituições com mais recursos (KLEBEL e ROSS-HELLAUER, 2023). Nas práticas de pesquisa, a pressão evoluiu do publicar ou perecer para o 'impacto ou perecer', fomentando um ecossistema de manipulações 'pós-produção' (como anéis de citação) que expandem a má conduta para além da fraude de dados (BIAGIOLI e LIPPMAN, 2020).

Em resposta, a alfabetização bibliométrica e o uso responsável tornaram-se centrais, consolidando o bibliotecário como um mediador focado no ensino do uso ético das ferramentas (COX *et al.*, 2019).

O Quadro 1 demonstra essas transformações sociotécnicas, atores e efeitos epistemológicos ao longo das quatro eras. Como síntese das transformações sociotécnicas discutidas ao longo das quatro eras, o Quadro 1 sistematiza as principais inflexões tecnológicas, os atores envolvidos e seus efeitos epistemológicos, funcionando como um instrumento de apoio analítico à interpretação histórica proposta da bibliometria.

Quadro 1 – Eras sociotécnicas da bibliometria: tecnologias, métricas e atores em dependência de trajetória

Característica	Era pré-computacional (~1895-1950)	Era do <i>mainframe</i> (~1960-1985)	Era do CD-ROM / Web 1.0 (~1985-2005)	Era da Web 2.0 / API (~2005-2025)
Tecnologia dominante	Ficha catalográfica e estantes	Computador <i>mainframe</i> e fitas magnéticas	PC, CD-ROM, <i>World Wide Web</i>	Web Social, API, nuvem, algoritmos IA
Suporte material da métrica	Fichas e folhas de papel	Fitas magnéticas e cartões perfurados	Discos óticos (CDs), servidores centralizados	Dados distribuídos na nuvem, infraestruturas interoperáveis e versionamento
Lógica métrica principal	Classificação (CDU)* Bibliografia estatística	Citação e indexação livre de vocabulário controlado	Acesso e recuperação (busca booleana e navegação)	Engajamento e algoritmo (metadados de uso, altmetria)
Consequência sociocultural	Sonho do conhecimento universal	Centralização do poder métrico (ISI)	Democratização do acesso	Datificação da ciência, <i>rankings</i> e vigilância algorítmica
Conceito-chave	Organização universal	Objetividade mecânica e avaliação quantitativa	Explosão da informação e acesso do usuário final (consolidação da lógica)	Plataformização da ciência e visibilidade algorítmica
Dependência da trajetória	Hierarquia semântica da CDU	Estrutura hierárquica do SCI, dependência do <i>mainframe</i>	Arquitetura de busca baseada em palavras-chave	Rede, interoperabilidade e versionamento de dados
Ator(es) humano(s) protagonista	Bibliotecário e documentarista	Cientista da informação, sociólogo da ciência e bibliotecário	Usuário final e bibliotecário	Gestor de pesquisa, bibliotecário e tecnólogo/programador
Mercado predominante	Estatal e não comercial	Comercial monopolista (ISI)	Assinatura e/ou licenciamentos comerciais	Plataformas/bases de dados comerciais e APC

Fonte: Elaborada pelo autor

*Lógica métrica não se refere a uma contagem quantitativa, mas à lógica fundamental de organização e avaliação do conhecimento daquela era. Na ausência de ferramentas computacionais para a contagem, a principal forma de ‘medir’ e situar um documento no universo do saber era através da sua classificação semântica, sendo a CDU a uma ferramenta para essa tarefa (RAYWARD, 1997).

A era da *Web 2.0* e das API culmina, portanto, em uma contradição, onde a fragmentação técnica sem precedentes (Scopus, GS, altmetria) falhou em desalojar a lógica avaliativa anterior. O sistema preserva vieses sistemáticos e uma visão reducionista do impacto baseada na dependência excessiva de indicadores quantitativos (PUENTE *et al.*, 2024). Configura-se não uma substituição, mas um ecossistema de avaliação híbrido e tenso (MAROLDI, 2025), onde novas infraestruturas coexistem com a resiliência das hierarquias tradicionais e perpetuam a atrofia interpretativa.

Considerações finais

Ao final do manuscrito, sustenta-se que os objetivos propostos foram alcançados e a questão de pesquisa respondida. Demonstrou-se que a história da bibliometria é marcada por uma profunda dependência de trajetória sociotécnica na qual os vieses e a ontologia invisível da era do mainframe persistem como fósseis tecnológicos sob a sombra das estruturas do passado (GUÉDON, 2001).

Como contribuição teórica, propõe-se um mecanismo em dois níveis para explicar essa permanência. No nível estrutural, a inércia epistemológica descreve o aprisionamento institucional a um paradigma de contagem, em que novas interfaces e fontes ampliam o acesso, mas preservam a lógica avaliativa original. No nível sociocognitivo, a atrofia interpretativa nomeia o processo pelo qual a facilidade de quantificação rebaixa a interpretação contextual a uma tarefa secundária, reforçando retroativamente a primazia do número (DE BELLIS, 2009; CRONIN, 2005).

A trajetória dos atores evidencia o bibliotecário como a contraparte humana dessa inércia ao adaptar-se como mediador essencial para a literacia métrica (COX *et al.*, 2017). Esse profissional demonstrou resiliência em todas as fases analisadas ao consolidar uma função indispensável que, contudo, não foi suficiente para alterar a lógica dominante do sistema. As práticas institucionais mostraram-se resistentes, persistindo mesmo diante da fragmentação tecnológica e do surgimento de alternativas métricas abertas e API que prometiam democratizar os dados (TORRES-SALINAS e ARROYO-MACHADO, 2022).

Nesse cenário, a Inteligência Artificial constitui a inflexão decisiva, pois ela pode favorecer uma recontextualização do saber, se orientada a leituras qualitativas e à explicitação de pressupostos, ou intensificar, com eficiência inédita, a mesma ontologia avaliativa ao automatizar triagens, rankings e decisões sob métricas herdadas (STEFANIS *et al.*, 2025). A bibliometria do futuro, portanto, depende menos da inovação técnica em si do que da capacidade institucional de enfrentar criticamente as infraestruturas que governam silenciosamente a visibilidade e o valor do conhecimento científico.

Referências bibliográficas

ARTHUR, W. B.

1989 Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal*. 99 (1989) 394.

ATKINS, Helen

1999 The ISI Web of Science: links and electronic journals. *D-Lib Magazine*. 5:9 (1999).

BORGMAN, C. L.

1996 Why are online catalogs still hard to use? *Journal of the American Society for Information Science*. 47:7 (1996).

BOWKER, G. C.; STAR, S. L.

1999 *Sorting things out: classification and its consequences*. Cambridge: MIT Press, 1999.

BRADFORD, Samuel C.

1934 Sources of information on specific subjects. *Engineering*. 137 (1934) 85-86.

BROADUS, R. N.

1987 Early approaches to bibliometrics. *Journal of the American Society for Information Science*. 38:2 (1987).

BROOKES, B. C.

1988 Comments on the scope of bibliometrics. *Informetrics*. 87/88 (1988).

BURGHAM, B.

1994 The New Alexandria?: the role of the CD-ROM in library collections. *The Electronic Library*. 12:3 (1994) 166-170.

BURNHAM, J. F.; SHEARER, B. S.; WALL, J. C.

1990 Combining new technologies for effective collection development: a bibliometric study using CD-ROM and a database management program. *Bulletin of the Medical Library Association*. 78:3 (1990).

BUSH, Vannevar

1945 As we may think. *The Atlantic Monthly*. 176:1 (1945).

CHISHTI, S. H.

1993 CD-ROM vs. online: a comparison of PsycLIT (CD-ROM) and PsycINFO (DIALOG). *The Reference Librarian*. 18:40 (1993).

COX, A. M. [et al.]

2017 Competencies for bibliometrics. *Journal of Librarianship and Information Science*. 51 (2017).

CRONIN, B.

2005 *The Hand of science: academic writing and its rewards*. Lanham, MD: Scarecrow Press, 2005.

DALBELLO, M.

2015 Digital convergence: the past in the present. In *A History of modern librarianship: constructing the heritage of western cultures*. Libraries Unlimited, 2015.

DAVID, P. A.

1985 Clio and the economics of QWERTY. *The American Economic Review*. 75:2 (1985).

DE BELLIS, N.

2009 *Bibliometrics and citation analysis: from the science citation index to cybermetrics*. Plymouth: Scarecrow Press, 2009.

DING, Y. [et al.]

2000 Bibliometric information retrieval system (BIRS): a Web search interface utilizing bibliometric research results. *Journal of the American Society for Information Science*. 51:13 (2000).

EDWARDS, P. N.

2010 *A Vast machine: computer models, climate data, and the politics of global warming*. Cambridge, MA: MIT Press, 2010.

FAIRTHORNE, Robert A.

1969 Empirical hyperbolic distributions (Bradford-Zipf-Mandelbrot) for bibliometric description and prediction. *Journal of Documentation*. 25 (1969).

GARFIELD, Eugene

1955 Citation indexes for science: a new dimension in documentation through association of ideas. *Science*. 122:3.159 (1955).

GHOLAMPOUR, S. [et al.]

2024 Does social media contribute to research impact?: an altmetric study of highly-cited marketing research. *Total Quality Management & Business Excellence*. (2024).

GINGRAS, Y.

2020 The Transformation of the scientific paper: from knowledge to accounting unit. In *Gaming the metrics: misconduct and manipulation in academic research*. Cambridge, MA: The MIT Press, 2020.

GODIN, B.

2006 From eugenics to scientometrics: Galton, Cattell and men of science. *Project on the History and Sociology of S&T Statistics*. 32 (2006).

GORMAN, M.

1980 Technical services in an automated library. In *The Role of the library in an electronic society: Proceedings of the 1979 Clinic on Library Applications of Data Processing*. Illinois: University of Illinois, 1980.

GRANOVSKY, Y. V.

2001 Is it possible to measure science?: V. V. Nalimov's research in scientometrics. *Scientometrics*. 52:2 (2001).

GROSS, P. L. K.; GROSS, E. M.

1927 College libraries and chemical education. *Science*. 66:1.713 (1927).

GUÉDON, J. C.

2017 *Open access: toward the internet of the mind*. Budapest: Open Book Publishers, 2017.

GUÉDON, J. C.

2001 *Oldenburg's long shadow: librarians, research scientists, publishers, and the control of scientific publishing*. Montréal: ARL, 2001.

HAHN, T. B.

1996 Pioneers of the online age. *Information Processing & Management*. 32:1 (1996).

HARZING, A.-W.; ALAKANGAS, S.

2015 Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison. *Scientometrics*. (2015).

HAWKINS, D. T.

1977 Unconventional uses of on-line information retrieval systems: on-line bibliometric studies. *Journal of the American Society for Information Science*. 32:1 (1977).

HÉRUBEL, J.-P. V. M.

1999 Historical bibliometrics: its purpose and significance to the history of disciplines. *Libraries & Culture*. 34:4 (1999).

HULME, Edward Wyndham

1924 *Statistical bibliography in relation to the growth of modern civilization: two lectures delivered in the University of London*. London, 1924.

KESSLER, M. M.

1963 Bibliographic coupling between scientific papers. *American Documentation*. 10 (1963).

KHADEMI, R.

2018 Visualizing Garfield's scientific performance. *Journal of Scientometric Research*. 7:3 (2018).

KLEBEL, T.; ROSS-HELLAUER, T.

2023 The APC-barrier and its effect on stratification in open access publishing. *Quantitative Science Studies*. (2023).

KÖVÁRI, A.; PASIN, M.; MEDUNA, A.

2025 The Dimensions API: a domain specific language for scientometrics research. *Frontiers in Research Metrics and Analytics*. 10 (2025).

KRÜGER, A. K.

2020 Quantification 2.0?: bibliometric infrastructures in academic evaluation. *Politics and Governance*. 8:2 (2020).

KYRILLIDOU, M.; COOK, C.

2008 The Evolution of measurement and evaluation of libraries: a perspective from the association of research libraries. *Library Trends*. 56:4 (2008).

LYU, P.; LIU, X.; YAO, T.

2023 A Bibliometric analysis of literature on bibliometrics in recent half-century. *Journal of Information Science*. (2023).

LYON, S.

1984 End-user searching of online databases: a selective annotated bibliography. *Library Hi Tech*. 6 (1984).

MABE, M.

2003 The Growth and number of journals. *Serials*. 16:2 (2003).

MALIHA, H.

2024 A Review on bibliometric application software. *Scientometrics Letters*. 1 (2024).

MALLIG, N.

2010 A Relational database for bibliometric analysis. *Journal of Informetrics*. 4 (2010).

MAROLDI, A. M.

2025 The Agency of digital platforms in open science: governance, metrics, and ontological implications. *Infonomy*. 3:5 (2025).

MARTÍN-MARTÍN, A. [et al.]

2021 Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage over time. *Scientometrics*. 126:4 (2021).

MEHO, Lokman I.; YANG, Kiduk

2007 A New Era in citation and bibliometric analyses: Web of Science, Scopus, and Google Scholar. *Journal of the American Society for Information Science and Technology*. 58:13 (2007).

MORAL-MUÑOZ, J. A. [et al.]

2020 Software tools for conducting bibliometric analysis in science: an up-to-date review. *Profesional de la Información*. 29:1 (2020).

NACKE, O.

1979 Informetrie: ein neuer name für eine neue disziplin. *Nachrichten für dokumentation*. 30:6 (1979).

PERITZ, B. C.

1988 Bibliometric literature: a quantitative analysis. *Informetrics* 87/88 (1988).

PRIEM, J.; PIWOWAR, H.; ORR, R.

2022 OpenAlex: a fully open bibliographic database. In INTERNATIONAL CONFERENCE ON SCIENTOMETRICS & INFORMETRICS (ISSI), 17th, Leuven, 2022 – *Proceedings*. Leuven: ISSI, 2022.

PRITCHARD, A.

1969 Statistical bibliography or bibliometrics? *Journal of Documentation*. 25:4 (1969).

PUENTE, J. H [et al.]

2024 Tools and methodologies for scientific evaluation: bibliometrics, scientometrics and informatics. In *Seminars in Medical Writing and Education*, 2024.

RAYWARD, W. B.

1997 The Origins of Information Science and the International Institute of Bibliography. *Journal of the American Society for Information Science*. 48:4 (1997).

ROTH, D. L.

2005 The Emergence of competitors to the Science Citation Index and Web of Science. *Current Science*. 89:9 (2005).

STAR, S. L.

1999 The Ethnography of infrastructure. *American Behavioral Scientist*. 43:3 (1999).

STAR, S. L.; RUHLER, K.

1996 Steps toward an ecology of infrastructure: design and access for large information spaces. *Information Systems Research*. 7:1 (1996).

STEFANIS, C. [et al.]

2025 Innovative methodological pillars for bibliometric studies: AI screening, data normalization and dual-tool analysis. *Discover Applied Sciences*. 7 (2025).

THELWALL, M.

2008 Bibliometrics to webometrics. *Journal of Information Science*. 34:4 (2008).

THOMPSON D. F.; WALKER C. K.

2015 A Descriptive and historical review of bibliometrics with applications to medical sciences. *Pharmacotherapy*. 35 (2015).

TORRES-SALINAS, D.; ARROYO-MACHADO, W.

2022 APIs en contextos bibliométricos: introducción básica y corpus exhaustivo. *Anuario ThinkEPI*. 16 (2022).

VAN LEEUWEN, Thed N. [et al.]

2001 Language biases in the coverage of the science citation index and its consequences for international comparisons of national research performance. *Scientometrics*. 51:1 (2001).

WHITTEMORE, R.; KNAFL, K.

2005 The Integrative review: updated methodology. *Journal of Advanced Nursing*. 52:5 (2005).

WINNER, L.

1980 Do artifacts have politics? *Daedalus*. 109:1 (1980).

WOUTERS, P.

1999 *The Citation culture*. Amsterdam: University of Amsterdam, 1999.

Tese de doutoramento - Universidade de Amsterdão, Faculdade de Ciências Sociais e Comportamentais.

ZHAO, Q.; FENG, X.

2022 Utilizing citation network structure to predict paper citation counts: a deep learning approach. *Journal of Informetrics*. 16:1 (2022).

Alexandre Masson Maroldi | alexandremm@ufscar.br

Universidade Federal de São Carlos (UFSCar), Brasil