

# CENTRALIDADE DIGITAL E REDES URBANAS: um estudo da hierarquia digital dos municípios brasileiros<sup>1</sup>

Pedro Mendes Loureiro<sup>2</sup>

Diego Las Casas<sup>3</sup>

Ana Flávia Machado<sup>4</sup>

Rodrigo Ferreira Simões<sup>4</sup>

Artigo submetido para o 44º Encontro Nacional de Economia da Associação Nacional dos Centros de Pós-Graduação em Economia (ANPEC), em Foz do Iguaçu, Brasil, 2016.

**Resumo:** A digitalização é um processo multidimensional de grande relevância econômica e social. Não obstante, é de difícil mensuração e sua relação com outros fenômenos econômicos é pouco compreendida. Este artigo apresenta quatro contribuições inter-relacionadas ao campo: propõe e aplica uma medida de digitalização ao conjunto dos municípios brasileiros, desenvolve uma tipologia dos municípios em termos de centralidade digital, explora a relação entre esta e a centralidade urbana e, por fim, estima os determinantes da centralidade digital. Através de amplo conjunto de fontes, faz-se uma análise de componentes principais e identificam-se duas dimensões da digitalização: uma geral, com contribuição positiva de todas as variáveis, e uma secundária que diferencia entre difusão do acesso digital e desenvolvimento de infraestrutura de ponta. Em seguida, uma análise de *cluster* sobre as mesmas variáveis identifica cinco categorias de municípios: 2,225 digitalmente excluídos, 2,477 com presença marginal da digitalização, 764 usuários passivos de conteúdo digital, 95 desenvolvedores autônomos com centralidade digital – os ecossistemas digitais – e 4 grandes *hubs* digitais. Os *clusters* são validados externamente através de *big data*, e então explora-se a relação entre centralidade urbana e centralidade digital através de um modelo *logit* multinomial. Encontra-se estreita relação entre atributos urbanos e digitalização e demonstra-se que o avanço desta está associado inicialmente a questões educacionais, mas que somente a integração com atividades econômicas de alto valor agregado é capaz de levar a condições de efetiva centralidade digital. Através dos resultados exploram-se as implicações em termos de políticas públicas, ressaltada a necessidade de medidas ativas de inserção digital para evitar o reforço de desigualdades regionais, econômicas e sociais.

**Palavras-chave:** digitalização, centralidade digital, redes urbanas, inclusão digital, economia regional.

**Abstract:** Digitalisation is a multidimensional process of great economic and social relevance. It is, nevertheless, markedly difficult to measure, and its relationship to other economic phenomena is poorly understood. This paper presents four interrelated contributions to this field: it proposes and applies a measure of digitalisation to all Brazilian municipalities; it develops a typology of the latter in terms of digital centrality; it explores the relation between digital and urban centrality; and it estimates the determinants of digital centrality. With data from several sources, a principal component analysis is run and two main dimensions of digitalisation are identified: a general one, with a positive contribution from all variables, and a secondary one that discriminates between the diffusion of digital access and the development of high-end infrastructure. A cluster analysis performed over the same data identifies five categories of municipalities: 2,225 digitally excluded ones, 2,477 with marginal digitalisation, 764 passive users of digital content, 95 autonomous content developers, with digital centrality – which can properly be called digital ecosystems – and 4 large digital hubs. The clusters are externally validated through big data, and then a multinomial logit model explores the relation between urban and digital centrality. The paper finds a strict connection urban attributes and digitalisation. It moreover shows how the latter's initial stages are stimulated by educational improvements, but it is only by becoming integrated with high valued added economic activities that digital centrality is effectively reached. The implications for public policies are then explored, highlighting the need for active initiatives to spur the digital inclusion of the population, under risk of heightening regional, economic and social inequalities.

**Key words:** digitalisation, digital centrality, urban networks, digital inclusion, regional economics.

**Códigos JEL:** R12, B52, L86, O33.

**Área ANPEC:** Área 10, Economia Regional e Urbana.

---

<sup>1</sup> Este trabalho foi inicialmente desenvolvido dentro do laboratório PoliTIC do Departamento de Melhoria de Processo de Software (DMPS) do Centro de Tecnologia da Informação Renato Archer (CTI). Os autores e autora agradecem o fundamental apoio de todos e todas da equipe, e em particular as valiosas sugestões e boas discussões com Mauro Zackiewicz e Renata Menezes.

<sup>2</sup> SOAS, University of London (UK). Contato: lmpedro@gmail.com e pm\_loureiro@soas.ac.uk. O autor agradece o apoio da CAPES (bolsa BEX-0840-14-9), cujo financiamento ajudou na elaboração desse projeto.

<sup>3</sup> DCC/UFMG (Belo Horizonte, Brasil). Contato: diegolascas@gmail.com.

<sup>4</sup> CEDEPLAR/UFMG (Belo Horizonte, Brasil). Contato: afmachad@cedeplar.ufmg.br e limoes@cedeplar.ufmg.br.

## 1 INTRODUÇÃO

A digitalização é um fenômeno multifacetado que envolve a produção, utilização e difusão de bens, serviços e tecnologias digitais nas diversas relações sociais e econômicas: trata-se, em termos gerais, da maior interação entre a esfera digital e as demais dimensões sociais. É um fenômeno possibilitado pela existência de infraestrutura física específica, mas que tem na mesma apenas uma condição de possibilidade. Primeiro, porque, para além da infraestrutura, é necessário que a população tenha as qualificações necessárias para utilizar e, em certa medida, produzir conteúdo digital – o que pode ser denominado letramento digital. E, em segundo lugar, porque a digitalização só se dá na medida em que a camada digital efetivamente se integra às várias relações sociais e, assim, as transforma. Sem conexão à internet e as qualificações individuais necessárias a digitalização não ocorre, mas estes elementos não garantem que haja, por exemplo, governo eletrônico ou trabalho por *home office*.

Tendo essa questão em tela, busca-se, portanto, analisar a distribuição espacial da digitalização no Brasil em suas várias camadas – difusão e qualidade das conexões à internet, acesso a aparelhos digitais, produção de conteúdo digital, letramento digital da população, cobertura de sinal móvel e infraestrutura digital. De uma forma inédita, trabalham-se diversas fontes e dados aplicados em modelos de aglomerados. Os perfis de municípios formados por essa técnica possibilitam não somente sua descrição como também o mapeamento regional, além de análise de fatores associados.

Assim, além desta introdução, o artigo compreende mais cinco seções. A segunda trata da digitalização como marco de uma ruptura tecnológica e o efeito concentrador (ou não) deste processo no espaço. Em seguida, apresentam-se as fontes de dados, variáveis selecionadas e as estratégias empíricas. Na quarta seção, os resultados de aplicação são analisados em termos da distribuição da digitalização, por intermédio de análise de componentes principais, e em termos de uma tipologia de centralidade digital, por meio de análise de aglomerados. Descrita a tipologia desenvolvida e sua associação com a rede urbana, explora-se os fatores associados à transição entre diferentes “estágios” da digitalização na quinta seção e, por fim, a sexta traz as principais considerações.

## 2 DIGITALIZAÇÃO EM CONTEXTO DE DESIGUALDADE

As rupturas tecnológicas impactam o modo de pensar e viver em sociedade, definindo novos rumos para sua constituição no campo da economia, da política e da cultura. Tal reestruturação retroalimenta e determina o desenvolvimento da tecnologia em um contexto histórico de forma circular. No caso das Tecnologias de Informação e Comunicação (doravante TIC's), a natureza de sua vivência não é diferente. A ruptura criada pela digitalização espraia-se em decorrência da dinâmica capitalista desde idos da metade do século XX, sem, contudo, tornar uniforme sua apropriação e utilização. Pelo contrário, são vivenciadas idiossincriticamente por diversos países e segmentos da população intra e interregional.

A digitalização é uma tecnologia que abrange a produção e o consumo, além do modo de viver. A sua eminência como centralidade na dinâmica do capital acarretou várias interpretações que renomearam essa fase do capitalismo como sociedade ou economia “pós-industrial”, do “conhecimento”, “nova economia” e esses nomes dependem da leitura de seus analistas (LIPSEY; CARLAW; BEKAR, 2006). Algo comum nessas análises é a observância de crescente produção e consumo de bens intangíveis, globalização de mercados, crescimento do terciário, em especial de serviços de comunicação, financeiro, de logística, de entretenimento, dentre outros e a maior necessidade de diferenciação de produtos por aumento do conteúdo de criatividade e conhecimento.

De um modo mais geral, as características inerentes à digitalização podem ser traduzidas em mobilidade, confiança, efeitos de rede, tendência à concentração de mercados e volatilidade (OECD, 2014). A mobilidade é entendida em termos da produção intangível, em especial de informação, em termos de usuários e em termos dos negócios que requerem menos trabalho alocado para desempenhar certas funções, assegurando maior flexibilidade espacial e temporal. Desenvolvimento e exploração da produção intangível é, a princípio, algo que demanda elevado volume de investimento em Pesquisa & Desenvolvimento para geração de um protótipo, um roteiro, um software, entretanto, sua distribuição é de custo praticamente nulo. Em vista disso, Baumol (2006) passou a denominá-la *nova economia*, ao buscar

descrever a produção, a distribuição e o consumo de bens simbólicos facilitados pelas Tecnologias de Informação e Comunicação, que dinamizaram a economia tangível, proporcionando a redução dos custos de produção, transação e disseminação do conteúdo. Para além da mobilidade gerada pela produção intangível em si, a digitalização credencia o usuário ou o cliente a acessar remotamente os serviços ou a demanda por bens tangíveis, o que muitas vezes o torna anônimo ou mesmo não identificável em um território. A conectividade por meio de redes pessoais ou a servidores próximos transcende o conceito de um mercado local para global. No âmbito dos negócios, a organização das firmas foi facilitada de tal maneira que a coordenação de atividades de produção em diversos territórios pode ser feita a um custo mais baixo, ainda que tais atividades sejam complexas. Os avanços tecnológicos possibilitam, assim, aumento da atividade econômica sem correspondente aumento de emprego. Em muitos casos, tem-se produção em “escala sem massa”, ou seja, aumenta-se vertiginosamente a produção sem ampliar o pessoal ocupado para gerenciar as operações rotineiras desta produção.

No que concerne à confiança, está-se tratando da geração de informação confiável na perspectiva do “big-data”. As informações podem ser coletadas dos fornecedores, clientes e de toda a operação de produção, gerando um sistema de controle. O acesso à informação é um problema da natureza do bem. A informação em si, considerada um bem público, é armazenada pela experiência, podendo o ato de armazenamento ser voluntário ou não. A capacidade de coletar e analisar a informação é crescente, em função do desenvolvimento de vários dispositivos eletrônicos. O *big-data* acaba por substituir a lógica da escassez pela da abundância, dado a profusão de informação em tempo real (HUTTER, 2003).

Os efeitos de rede são entendidos na perspectiva da interação entre os diversos agentes do sistema econômico, criando-se externalidades tanto positivas, como aproveitamento e aprimoramento de novos produtos pelas redes de contatos, quanto negativas, quando se tem excesso de conectividade nessas redes, reduzindo a velocidade da informação, por exemplo. Os efeitos de rede combinados ao baixo custo marginal podem concentrar ainda mais os mercados, formando-se monopólios e oligopólios. Isto porque o recurso a plataformas amigáveis afeta as preferências dos integrantes das redes, tornando-os “fidelizados”, o que contribui para ampliar o poder de mercado da firma. Ao mesmo tempo, para garantir este mercado, tais firmas tendem a atuar de modo a capturar inovações que são desenvolvidas em nichos deste mercado por empresas menores. A trajetória do processo inovativo é mais rápida, o que amplia a volatilidade da estrutura de mercado, na medida em que se ocupam os interstícios deixados pelas pequenas empresas desenvolvedoras de novos produtos e processos (ZACKIEWICZ, 2015).

O cenário acima descrito traduz uma realidade de países e firmas que, há muito, já se destacam na hegemonia tecnológica, como Estados Unidos, Alemanha, Grã-Bretanha, Japão, Coreia, mais recentemente China e Índia. No caso brasileiro, permanecemos à margem desta configuração, nosso lugar é de o usuários/consumidores e não de produtores das inovações tecnológicas advindas da digitalização.

Ainda que existam diferenciais entre os países, internamente aos mesmos, observam-se desigualdades não desprezíveis. Robinson et al. (2015) salientam que 14% dos americanos adultos permanecem sem acesso à internet, especialmente negros, mulheres e indivíduos situado na base da pirâmide socioeconômica, assim como residentes de áreas rurais. No caso do Brasil, Sorj e Guedes (2005) também evidenciam a dinâmica da exclusão digital sobre os setores mais pobres da população, considerando as diferenças segundo sexo, raça, nível de riqueza e escolaridade e seus resultados mostram que a desigualdade na distribuição da educação e da riqueza são os fatores predominantes. Tal constatação reitera que o ciclo vicioso das desigualdades pode ser ampliado pela digitalização. E tais desigualdades tendem a se distribuir também no espaço, haja visto a prevalência de residência destes grupos populacionais em áreas e/ou regiões mais pobres. Compreender fatores locais torna-se, portanto, mister.

Na literatura de economia regional e urbana, a análise de cluster é comumente aplicada para tratar de amenidades urbanas e efeitos de aglomeração. A concepção teórica remonta a Alfred Marshall, em seus estudos sobre localização industrial. Para ele, os distritos industriais são territórios de economias externas de aglomeração. Tais economias podem ser tratadas por meio da presença de infraestrutura de serviços públicos, tais como acesso a energia, água, transporte, educação e saúde, criando-se assim mercado para trabalhadores. A formação de distritos asseguraria, assim, custos médios de produção e transação decrescentes, gerando economias internas de escala.

Atualmente, como bem ressaltam Machado et al. (2013), a incorporação do fator amenidades urbanas tem se tornado determinante tanto de decisões de localização de residências como também de escalas de centros urbanos. Por exemplo, Glaeser et al. (2001), ao tratarem da atratividade do urbano, sugerem que o padrão de consumo e amenidades urbanas – variedade de bens e serviços, configuração física e estética, qualidade dos serviços públicos ofertados, ou ainda a garantia da velocidade de transporte e informação – definem o perfil de atratividade das cidades. Os autores partem do pressuposto de que o crescimento da renda e os avanços nas tecnologias de comunicação e transporte tornaram as cidades grandes centros de consumo de larga escala, desvinculando-as da existência de economias de escala marshallianas. Prioriza-se, deste modo, a oferta de amenidades para atração de uma população qualificada e de alta renda.

Parte deste foco se deve ao processo de digitalização. E, de maneira reverberativa, a concentração e a intensidade do uso da internet e os impactos de qualificação do trabalho a ela relacionados ampliam o nível de renda e competitividade desses locais. Neste sentido, capital humano, financeiro e cultural afetariam tanto a inclusão digital como essa os retroalimentaria, por intermédio do aumento de produtividade, competitividade e atratividade desses centros urbanos, criando novas classes de ganhadores e perdedores (ROBINSON *et al.*, 2015).

Trazendo essa leitura para o caso do Brasil, pretende-se analisar em que medida o processo de digitalização rompe ou consolida as desigualdades regionais, guardando as melhores posições para os grandes centros localizados no Sudeste do país, assim como se dá com o mercado de trabalho, acesso à infraestrutura básica e aos serviços de saúde e educação. Ademais, busca-se analisar se os determinantes desses clusters “digitais” seriam generalizáveis ao se contrapor à hierarquia urbana.

### 3 ESTRATÉGIA EMPÍRICA

#### 3.1 FONTES E TRATAMENTO DOS DADOS

Mensurar a digitalização passa por coletar uma gama variada de indicadores relacionados à qualidade e difusão das conexões à internet, ao acesso a aparelhos eletrônicos, à utilização e produção de conteúdo digital, ao letramento digital da população, à cobertura de sinal móvel e, por fim, à infraestrutura digital. Deve-se observar que as pesquisas estatísticas “tradicionais” têm cobertura reduzida de fenômenos relativos à digitalização. Com isto em mente, foram levantadas informações de diferentes fontes para compor um quadro balanceado do objeto em tela. Para além disso, coletaram-se também dados referentes às características urbanas dos municípios, que são mais à frente relacionados à digitalização dos mesmos. Sete instituições são responsáveis pelas informações aqui coligidas: a Agência Nacional de Telecomunicações (Anatel); o Instituto Brasileiro de Geografia e Estatística (IBGE); o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP); o Comitê Gestor da Internet no Brasil (CGI.br); o laboratório *Measurement Lab* (M-Lab), associado ao Google; a empresa Teletime e a Thomson Reuters. Para a descrição, fonte, período de referência e sigla das variáveis, ver QUADRO 1.

A conectividade à internet e o acesso às plataformas digitais foram capturados através de dados da Anatel, do Censo Demográfico de 2010 e do *Atlas brasileiro de telecomunicações* (Teletime 2014). O percentual de domicílios com computadores com acesso à internet mede a difusão da acessibilidade pessoal ao conteúdo digital. Avançando nisso, o número de acessos ao Serviço de Comunicação Multimídia (SCM) por habitante, em diferentes velocidades, reflete padrões de usuários distintos (acessos mais rápidos sugerem aplicações intensivas em dados e processamento). Dado também considerarem acessos de empresas, estas variáveis medem a camada digital no âmbito produtivo e doméstico. Já o acesso a tecnologias móveis é capturado em quatro variáveis. O percentual de pessoas com telefone celular reflete a difusão da tecnologia, enquanto o número de antenas por área do município reflete a ubiquidade da cobertura. Por fim, o número de operadoras que disponibilizam sinal 3G e 4G é importante dimensão do potencial de mobilizar tecnologias digitais a partir de dispositivos móveis. Outro aspecto cuja interação com o digital vem crescendo rapidamente é o número de assinaturas de televisão por habitante.



Para os aspectos educacionais foram coletados dados do Censo e do Censo Escolar. Para o letramento digital proporcionado pela educação não especializada, como *proxy*, observou-se a infraestrutura digital das escolas públicas por meio da presença de banda larga. Quanto à mão-de-obra especializada, calculou-se a proporção de pessoas com graduação ou pós-graduação em cursos associados à TI e ao desenvolvimento de software. Para a produção de conteúdo digital, mensurou-se o fenômeno em duas variáveis. Através de dados do Censo, viu-se o emprego total em empresas especializadas em TI e a presença de ocupações relacionadas à atividade (programadores etc.) em empresas com outras atividades-fim.

Para a infraestrutura digital coletou-se, adicionalmente, dados da Anatel e de instituições relacionadas ao Comitê Gestor da Internet no Brasil (CGI.br). Através do número de centrais de fios de Exploração Industrial de Linha Dedicada (EILD) pode-se observar a interconexão de prestadoras de serviços de telecomunicações, relacionada à capilaridade da rede de dados e, conseqüentemente, à extensão da camada digital ao âmbito produtivo. As demais variáveis medem a infraestrutura “interna” do âmbito digital, responsáveis pela própria operação da internet. Um *Autonomous System Network* (ASN) é o ponto nodal de comunicação entre usuários finais e rede mundial; sediá-los é, assim, atividade especializada, associada a espaços com alguma centralidade na hierarquia digital do país e boa infraestrutura. Elemento ainda mais central são os pontos de interconexão de tráfego (PIX), em que as ASN compartilham dados entre si para reduzir transmissões de longa distância – são serviços que demandam larga escala, *hubs* digitais por excelência. Empregou-se o número de ASN participantes de PIX no município, que assume valor 0 quando não há nenhum PIX.

Por fim, mediu-se o governo eletrônico através da interação digital com a administração municipal, contadas as funcionalidades oferecidas na página da prefeitura (acesso a documentos, licitações etc.).

Com o propósito de referendar os resultados, fez-se uso de *big data* referente ao teste *Network diagnosis test* (NDT).<sup>5</sup> Como o nome sugere, trata-se de um teste da qualidade da conexão à internet do usuário final, que registra uma série de características da mesma. Como noção do tamanho da base de dados, nota-se que há, em dados referentes ao mundo inteiro, mais de 26 bilhões de linhas por mês. Foram analisados todos os testes disponíveis na plataforma *BigQuery* do *Google*, para o período 01/2013 a 04/2014, e calculada a mediana dos parâmetros em questão por município brasileiro (1748 municípios foram localizados).<sup>6</sup> Coletou-se a velocidade de download, o percentual de perda de pacotes de dados e de pacotes retransmitidos (importantes dimensões da confiabilidade da conexão) e o número de IP distintos que realizaram testes (o que, dividido pela população, mensura a densidade de conexões).

Também foram compiladas informações sobre as características urbanas dos municípios. Enxerga-se o fenômeno urbano através de medidas de especialização industrial, presença de serviços modernos, diversificação econômica, tamanho da economia local, nível educacional da população, desenvolvimento científico e centralidade urbana. A especialização industrial é definida, com base em Garcia e Simões (2012), como:

$$(1) \quad ind_m = \frac{vi_m}{PIB_m} \left[ 1 - \exp \left( \ln(0.05) \frac{PIB_m}{PIB_{ref}} \right) \right]$$

onde  $vi_m$  é o valor adicionado pela indústria,  $PIB_m$  é o produto interno bruto, o subscrito  $m$  indica o município em questão e  $PIB_{ref}$  é o PIB mais alto do estado ao qual o município pertence. Já a presença de serviços produtivos modernos foi definida como a proporção de trabalhadores em atividades intensivas em conhecimento, típicas de centros urbanos.<sup>7</sup> A definição é baseada em Simões e Freitas (2014), com a diferença de aqui se excluírem as classificações relacionadas à TI, de forma a evitar uma análise circular. A diversificação econômica é medida através do índice modificado de concentração setorial de Hirschman-Herfindahl, calculado a partir do setor de ocupações dos trabalhadores, adaptado de Combes (2000):

<sup>5</sup> Para um estudo que compara os resultados do teste com similares, ver OCDE (2013).

<sup>6</sup> O procedimento de coleta e classificação pode ser detalhado mediante solicitação.

<sup>7</sup> Ver Anexo 1 para listagem das classificações de atividades econômicas.

$$(2) \quad div_m = 1 - HHI_m$$

onde  $HHI$  é o índice de concentração do município, considerando-se as divisões da CNAE 2.0 (seus dois primeiros dígitos) como cada setor. O tamanho da economia local, por sua vez, é captado pela densidade do mercado de trabalho: utiliza-se o total de trabalhadores dividido pela área do município.

A produção científica de cada espaço foi vista através do número de artigos indexados publicados em 2010, disponível na base Web of Science (ISI, 2010).

Para a centralidade dos municípios empregou-se a classificação desenvolvida na pesquisa *Regiões de influência das cidades 2007* (IBGE 2008). Este estudo identifica os locais dotados de centralidade urbana e suas regiões de influência, em cinco níveis. No caso de grandes aglomerações urbanas, contudo, a pesquisa trata como unidade de análise as Áreas de Concentração de População (ACP), e não seus municípios específicos. Para garantir a compatibilidade com as demais informações levantadas, consideramos que todos os municípios de uma dada ACP possuem a mesma centralidade desta.<sup>8</sup> No modelo estimado foi empregada uma variável binária que assume valor 1 quando o município é uma capital regional ou metrópole.

Quadro 1 – Dados coletados, fontes, código das variáveis e período de referência das mesmas

Variável	Código	Fonte	Período de referência
Número de acessos ao SCM de 0 a 12 Mbps por hab.	sacessos_b	Anatel (2014b)	01/2014
Número de acessos ao SCM de > 12 Mbps por hab.	sacessos_a	Anatel (2014b)	01/2014
Número de antenas de celular por área	antarea	Anatel (2014d)	02/2014
Número de operadoras 3G no município	total3g	Anatel (2014c)	02/2014
Número de operadoras 4G no município	total4g	Teletime (2014)	12/2013
Número de assinaturas de TV por hab.	nasspop	Anatel (2014a)	02/2014
Número de centrais de fios de Exploração Industrial de Linha Dedicada (EILD)	ncentr	Anatel (2014e)	10/2012
Número de Sistemas Autônomos (ASN) sediados no município	asn	PTTbr (2014a)	17/04/2014
ASN participantes nos pontos de interconexão de tráfego	pix_part	PTTbr (2014b)	01/06/2014
Proporção de pessoas com celular	celular	IBGE (2010)	2010
QL de trabalhadores da IBSS*, Censo 2010	censo_ibss	IBGE (2010)	2010
QL de trabalhadores PROFSS* fora da IBSS*, Censo 2010	profss_censo	IBGE (2010)	2010
Proporção de domicílios com acesso à internet	comp_net	IBGE (2010)	2010
Proporção de escolas com banda larga	id_banda_larga	Inep (2013)	2013
QL de (pós-)graduados em cursos de PROFSSs**	grad_profss	IBGE (2010)	2010
Funcionalidades da página eletrônica da prefeitura***	vsum_113a125	IBGE (2012a)	2012
Mediana da velocidade de download	p50_ds	M-Lab (2014)	01/2013 a 05/2014
Mediana da proporção de perda de pacotes	p50_ploss	M-Lab (2014)	01/2013 a 05/2014
Mediana da proporção de pacotes retransmitidos	p50_pret	M-Lab (2014)	01/2013 a 05/2014
Número de IP distintos que realizaram testes NDT por hab.	mlabpop	M-Lab (2014)	01/2013 a 05/2014
QL da proporção de pessoas com ensino superior completo	educ_sup	IBGE (2010)	2010
Diversificação econômica@	divers	IBGE (2010)	2010
Serviços modernos†: QL de trabalhadores de serviços modernos	smod	IBGE (2010)	2010
Hierarquia urbana: cinco variáveis binárias que indicam centralidade urbana (1: maior centralidade, 5 menor)	regic_#	IBGE (2008)	Indeterminado
Hierarquia urbana simplificada: variável binária que assume valor 1 quando o município é das categorias 1 ou 2	regicb	IBGE (2008)	Indeterminado
Densidade de empregos: log do número de empregos por km <sup>2</sup>	lnden	IBGE (2010)	2010
Densidade industrial: indicador da participação da indústria no valor adicionado total do município.	pind	IBGE (2012b)	2012
Desenvolvimento científico: artigos publicados no município	artigos	ISI (2010)	2010
Desenvolvimento científico simplificado: variável binária que assume valor 1 quando o número de artigos publicados no município é acima da média estadual.	martigos	ISI (2010)	2010

<sup>8</sup> O procedimento também é adotado pelo IBGE (2013).

Notas: As linhas sem preenchimento são as variáveis incluídas no modelo PCA e análise de cluster, as sombreadas de cinza claro são *big data* usadas para referendar os clusters e as sombreadas de cinza escuro são as variáveis explanatórias do modelo *logit*.

\*: IBSS é indústria brasileira de software e serviços de TI e PROFSS são profissionais de software e serviços de TI (Observatório Softex 2012), ver Anexo 1; \*\*\*: perguntas 113 a 125 da pesquisa Munic de 2012; @: 1 - índice Hirschman Herfindal de concentração econômica (setorial), calculado para divisões (2 dígitos) da CNAE 2.0; †: ver Anexo 1 para listagem das classificações consideradas.

### 3.2 ESTRATÉGIA EMPÍRICA

A análise empírica compreende dois momentos distintos: uma etapa de classificação dos municípios em termos de sua digitalização, que também inclui análise da relação entre os diferentes componentes desta, e outra que busca explicar estes resultados com base nos atributos urbanos das localidades.

O primeiro passo é uma análise de componentes principais (PCA) das dezesseis variáveis relacionadas à digitalização. A técnica permite tanto a redução dimensional do fenômeno (i.e., expressar a digitalização por meio de um pequeno número de fatores) quanto o estudo da relação entre as variáveis empregadas. O procedimento consiste em encontrar vetores ortogonais entre si no espaço  $j$ -dimensional, onde  $j$  é o número de variáveis analisadas, que sucessivamente expliquem a maior variância possível dos dados. Cada observação pode então ser descrita a partir de  $p$  componentes ( $p < j$ ), com pouca perda de informação. Na medida em que os vetores são compostos por uma soma ponderada das variáveis normalizadas, pode-se atribuir a cada componente uma interpretação socioeconômica derivada desta estrutura – e, adicionalmente, a análise desta estrutura lança luz sobre a relação entre as variáveis.

Empregou-se o PCA, portanto, como forma de compreender a relação entre os diferentes aspectos da digitalização – infraestrutura, letramento digital etc. – e expressá-la a partir de um número reduzido de variáveis. Adiantando os resultados, são encontrados dois componentes relevantes, o primeiro associado à digitalização em sentido amplo (contribuição positiva de todas as variáveis) e o segundo que expressa a prevalência ou de infraestrutura digital especializada ou de difusão do acesso (cujas variáveis tem carga de sinal oposto).

Em seguida, busca-se construir uma tipologia dos 5565 municípios brasileiros a partir de algoritmos de *clustering*, também com base nas mesmas variáveis empregadas no PCA. A ideia por detrás deste tipo de procedimento é agrupar as observações em grupos internamente homogêneos e diferentes entre si. Empregou-se o método *k-means clusters*, um procedimento iterativo e não paramétrico de classificação. Para definir o número de clusters experimentou-se com  $k$  entre 2 e 20, e pôde-se observar que  $k = 5$  representa um bom compromisso entre parcimônia e capacidade discriminatória.<sup>9</sup> Quanto à medida de distância, utiliza-se a distância euclidiana, ou:

$$(3) \quad d_{i,l} = (\mathbf{x}_i - \mathbf{x}_l)'(\mathbf{x}_i - \mathbf{x}_l)$$

onde os subscritos  $i$  e  $l$  indicam as duas observações cuja distância se calcula e o vetor  $\mathbf{x}$  é o vetor das  $j$  variáveis incluídas no cluster.

Analisa-se então a classificação resultante com vistas a construir uma tipologia dos municípios brasileiros e observar sua hierarquia. Descrevem-se os cinco grupos em termos das variáveis incluídas na definição dos clusters e dos demais atributos urbanos não incluídos, inclusive a distribuição geográfica dos locais. Como forma de referendar os resultados obtidos, observa-se a distribuição dos dados “orgânicos” (*big data*) entre os clusters. Interpretam-se as categorias como municípios: i) digitalmente excluídos; ii) com inserção digital marginal; iii) usuários digitais passivos, sem grandes transformações do âmbito produtivo; iv) usuários ativos, com ecossistemas digitais; e v) grandes *hubs* digitais.

Finda a fase classificatória, busca-se explicar o avanço da digitalização e a conformação de hierarquia digital como função dos atributos urbanos dos municípios. Foi primeiro estimado um modelo linear, não reportado aqui por questões de espaço,<sup>10</sup> que regride o primeiro componente obtido do modelo PCA contra as variáveis explanatórias descritas adiante. Constatou-se alto grau de ajuste ( $R^2$  de 0.88),

<sup>9</sup> A partir do gráfico *clustergram* (Schonlau 2002) percebe-se que há claros ganhos classificatórios à medida que  $k$  se aproxima de 5, mas que a partir disto a distância entre os centroides varia muito pouco.

<sup>10</sup> Disponível sob demanda.

sinal esperado para as variáveis e grande significância, o que é importante evidência da forte relação entre centralidade urbana e centralidade digital. Contudo, uma limitação do método, por ser linear, é que ele considera que as mesmas variáveis são relevantes nos vários “estágios” da digitalização. Como forma de contornar isto, ajusta-se um modelo *logit* multinomial para determinar o pertencimento aos clusters anteriormente identificados.

Trata-se de um modelo não linear que determina, a partir de um conjunto de variáveis independentes, a probabilidade de as observações enquadrarem-se em um número finito de categorias – no caso, os *clusters*. Não se pressupõe qualquer forma de ordenamento entre as categorias e, assim, o impacto das variáveis independentes pode afetar diferentemente a razão de probabilidades entre quaisquer dois grupos. Exemplificando, o acréscimo de uma dada variável pode estar associada a maior probabilidade de inserção nos clusters dois e três, mas não afetar a transição para o cluster quatro – ou, alternativamente, apenas impactar a probabilidade relativa de inserção nos clusters três ou quatro.

Deve-se observar que, devido ao pequeno número de municípios considerados grandes *hubs* digitais (apenas quatro), não é possível incluí-los. A exclusão é adicionalmente justificada por não se esperar que a transição para este grupo seja um fenômeno recorrente, passível de ser determinado pelas variáveis empregada. Ao contrário da categoria de municípios com ecossistemas digitais (*cluster 4*), na qual é teoricamente plausível que todos se insiram mediante as condições socioeconômicas adequadas, espera-se que *hubs* digitais sejam fenômenos relativamente escassos no espaço. O modelo determina assim o pertencimento aos clusters 1, 2, 3 e 4, e sua forma funcional é descrita pelas equações a seguir:

$$(4) \quad P(c_i = 1) = \frac{1}{1 + \left[ \sum_{z=2}^4 \exp(\mathbf{x}_i' \boldsymbol{\beta}_z) \right]}$$

$$(5) \quad P(c_i = q) = \frac{\exp(\mathbf{x}_i' \boldsymbol{\beta}_q)}{1 + \left[ \sum_{z=2}^4 \exp(\mathbf{x}_i' \boldsymbol{\beta}_z) \right]}, q = 2, 3, 4$$

onde  $P(c_i = q)$  é a probabilidade da  $i$ -ésima observação pertencer ao cluster  $q$ ;  $\mathbf{x}_i$  é o vetor de variáveis explanatórias;  $\boldsymbol{\beta}_z$  ( $2 \leq z \leq 4$ ) são os vetores de parâmetros associados aos clusters 2, 3 e 4. Como se investiga a inserção dos municípios em quatro categorias exaustivas, a probabilidade de pertencimento a uma das categorias pode ser obtida por subtração. No caso em questão, utiliza-se o cluster 1 como referência. As variáveis explanatórias, cuja descrição está disponível no QUADRO 1, são: nível educacional (*educ\_sup*); diversificação econômica (*divers*); densidade do mercado de trabalho (*Inden*); presença de serviços modernos (*smod*);<sup>11</sup> centralidade urbana (*regicb*); industrialização (*pind*); produção científica (*martigos*); e região, com São Paulo separado e tomando o nordeste como referência (*regiao\_u*).

## 4 RESULTADOS

### 4.1 DIMENSÕES DA DIGITALIZAÇÃO E SEU AVANÇO: UMA ANÁLISE DE COMPONENTES PRINCIPAIS

Como primeiro movimento, explora-se a relação entre as diferentes dimensões da digitalização através de um modelo PCA, cujos resultados estão reportados na TAB. 1.

O resultado mais importante do modelo é expresso no primeiro componente, que pode ser visto como uma medida global de digitalização. Observa-se que todas as variáveis têm carga positiva e relativamente alta, com destaque para duas categorias: o acesso “mediano” à camada digital, como proporção de domicílios com acesso à internet e densidade de conexões de baixa velocidade; e a produção

<sup>11</sup> As variáveis *educ\_sup* e *smod* não são incluídas como QL, mas sim como proporções, para facilitar a comparação do impacto do aumento de um ponto percentual.



local de software, visto através da presença de empresas da IBSS. De toda forma, o que o resultado centralmente demonstra é que a digitalização avança, primordialmente, através de um processo multifacetado que inclui suas diversas dimensões – o que é assegurado pela alta proporção da variância explicada pelo primeiro componente. Isto é de grande importância para a elaboração de políticas públicas, na medida em que sugere a necessidade de iniciativas que combinem formação de mão-de-obra especializada, letramento digital amplo da população e infraestrutura dedicada.

Já o segundo componente é dominado pelas variáveis infraestruturais, com destaque para o tamanho do PIX e o número de ASN e de centrais EILD, enquanto variáveis de difusão (e.g., posse de celular) entram com sinal oposto. A partir disso, e em vista da ortogonalidade com os demais componentes, pode-se considerar que o segundo componente expresse a existência de momentos distintos de avanço da digitalização: embora o processo ocorra sobretudo de maneira combinada, há municípios que priorizam o investimento em infraestrutura enquanto outros priorizam a difusão. Associando esta análise aos resultados encontrados com os *clusters*, descritos adiante, observa-se que, para o grupo de municípios de menor digitalização, a correlação entre o primeiro e segundo componentes é negativa (-0.70), crescendo até atingir 0.67 e 1.00 para os dois conjuntos de cidades altamente digitalizadas. Ou seja, pode-se concluir que nos estágios iniciais da digitalização o avanço se dá através de maior difusão, enquanto atingir efetiva centralidade digital está associado especialmente, mas não unicamente, à presença de infraestrutura pesada. Portanto, o desenvolvimento da camada digital até o eventual estabelecimento de um polo ou ecossistema digital não é um processo linear, mas sim com quebras e diferentes gargalos em seu decorrer.

Tabela 1 – Componentes principais da digitalização dos 5565 municípios brasileiros

Variável	Cargas fatoriais			Variância não explicada
	Componente 1	Componente 2	Componente 3	
id_banda_larga	0.192	-0.268	0.377	0.394
nasspop	0.318	-0.101	-0.096	0.304
sacessos_a	0.240	0.142	-0.345	0.422
sacessos_b	0.296	-0.195	0.071	0.326
comp_net	0.319	-0.250	0.142	0.150
celular	0.213	-0.284	0.353	0.338
grad_profss	0.302	-0.093	-0.074	0.377
profss_censo	0.194	0.021	-0.325	0.622
censo_ibss	0.301	-0.024	-0.166	0.377
vsum_113a125	0.237	-0.157	0.049	0.569
antarea	0.251	0.245	-0.187	0.390
total3g	0.279	-0.125	-0.109	0.439
total4g	0.234	0.280	-0.180	0.397
ncentr	0.202	0.393	0.031	0.328
asn	0.198	0.438	0.374	0.067
pix_part	0.134	0.427	0.468	0.131
Proporção da variância explicada pelo componente 1: 0.405				
Proporção da variância explicada pelo componente 2: 0.164				
Proporção da variância explicada pelo componente 3: 0.079				

Fonte: Elaboração nossa a partir de dados de várias fontes.

Não parece haver interpretação socioeconômica clara para o terceiro e demais componentes, em vista do que não são analisados. Dado o baixo poder explicativo dos mesmos, pode-se concluir que a digitalização é razoavelmente bem expressa através de dois componentes: o central, que reporta o desenvolvimento da camada digital como um todo e sugere a necessidade de políticas integradas para seu avanço; e um secundário, que revela como, em diferentes estágios, questões infraestruturais ou de difusão alternadamente diferenciam melhor os municípios. Passa-se ao desenvolvimento de uma tipologia de municípios em termos de sua digitalização.

#### 4.2 CENTRALIDADE DIGITAL: UMA TIPOLOGIA A PARTIR DE K-MEANS CLUSTERS

Empregando-se as mesmas variáveis do modelo PCA, estabeleceu-se uma tipologia de municípios em termos de sua digitalização. Foram encontrados cinco grupos:

1. 2225 municípios com exclusão digital quase completa;
2. 2477 municípios com inserção digital marginal, mas ainda substancialmente excluídos;
3. 764 municípios com certa presença digital, mas usuários passivos que não produzem conteúdo – a camada digital está integrada ao âmbito doméstico, sem transformar a estrutura produtiva;
4. 95 municípios que são efetivamente digitalizados, com utilização e produção de conteúdo digital e integração deste com o restante da economia: são ecossistemas digitais;
5. 4 municípios (São Paulo, Belo Horizonte, Porto Alegre e Rio de Janeiro) que conformam os grandes *hubs* digitais brasileiros, capazes de polarizar grandes espaços e liderar e coordenar a produção de software e serviços de TI no país.

Antes de se observar a distribuição desses *clusters* no espaço e analisar as implicações da tipologia para se pensar em como avançar a inclusão digital no país, justifica-se a descrição dos mesmos nos termos propostos acima. Para tanto, algumas variáveis selecionadas, que foram empregadas na construção dos clusters, têm sua distribuição analisada através de *boxplots* na FIG. 1.<sup>12</sup>

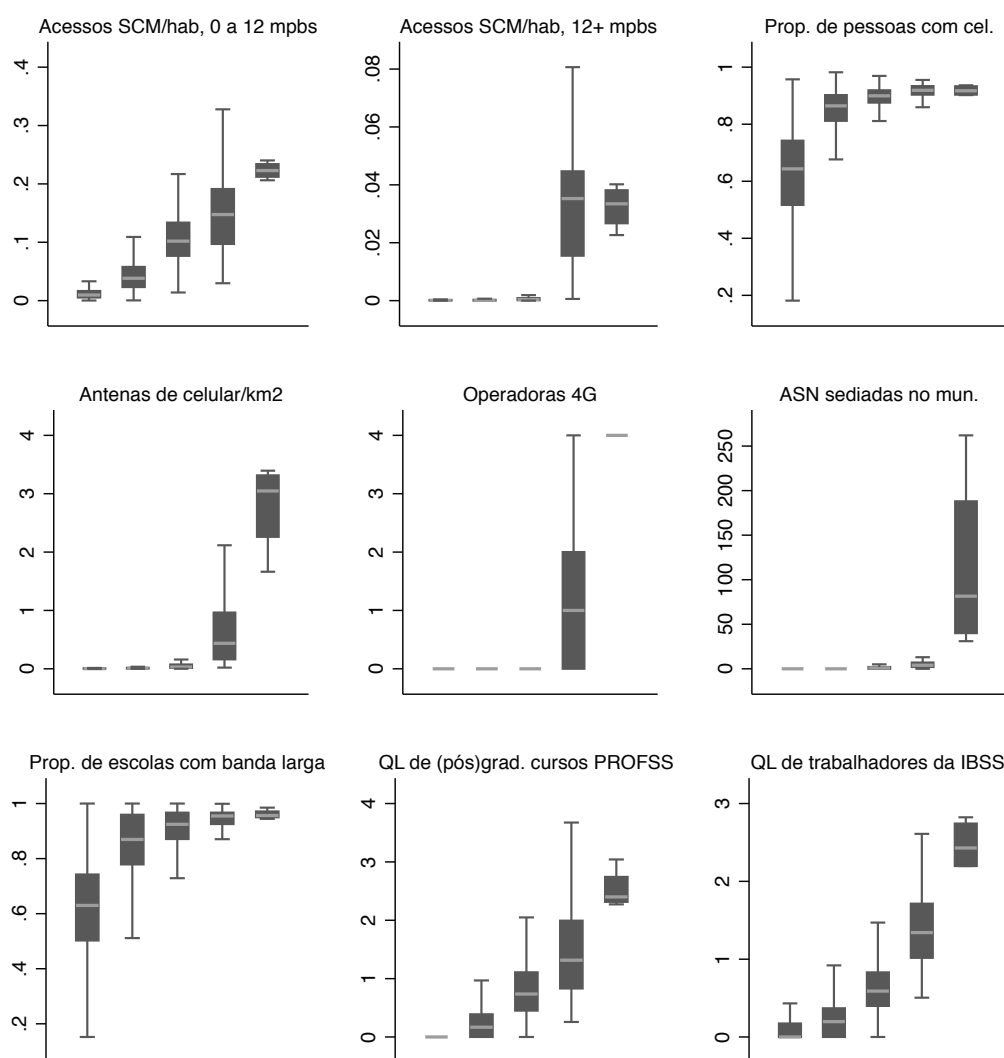


Figura 1 – Distribuição de diferentes aspectos da digitalização nos cinco *clusters*

Fonte: Elaboração nossa a partir de dados de várias fontes.

<sup>12</sup> Neste e nos demais gráficos não foram plotados os *outliers*, definidos como valores acima de  $p_{75} + 1.5 \cdot (p_{75} - p_{25})$  ou abaixo de  $p_{25} - 1.5 \cdot (p_{75} - p_{25})$ , onde  $p_x$  é o  $x$ -ésimo percentil.

Salta aos olhos a não linearidade do agrupamento, em várias dimensões. Observa-se como, no âmbito da difusão da camada digital, o número de acessos a baixas velocidades cresce de maneira razoavelmente linear, enquanto as conexões a alta velocidade são frequentes somente nos ecossistemas e *hubs* digitais (grupos 4 e 5). Por outro lado, a proporção de pessoas com celular diferencia bem os estágios iniciais da digitalização, mas logo deixa de fazê-lo. Já a infraestrutura digital emerge como discriminante apenas em estágios avançados: enquanto não há praticamente nenhum usuário digital passivo (grupo 3) com cobertura 4G, grande número de ecossistemas digitais tinha pelo menos uma operadora que oferecia o serviço. Nota-se também como os elementos mais especializados da infraestrutura são a maior diferença entre os *hubs* digitais e o restante, haja vista a radical quebra no número de ASN sediadas. A densidade de cobertura de sinal telefônico, indicando ubiquidade do acesso ao digital, é também outra importante discriminante nos estágios mais avançados. Em termos do letramento digital e mão-de-obra especializada, um quadro similar emerge. A proporção de escolas com banda larga é bastante baixa nos municípios digitalmente excluídos, enquanto não difere significativamente a partir dos usuários digitais passivos. Por outro lado, trabalhadores qualificados (QL de graduados em cursos de PROFSS) só se fazem presentes em estágios mais avançados e a presença substancial de produção de software (QL de empresas da IBSS) só é encontrada em ecossistemas e *hubs* digitais. É este fato, de que mais de 75% dos municípios do terceiro grupo têm QL de empresas da IBSS abaixo de 1 – e, portanto, abaixo proporção nacional – que estimula a classificação dos mesmos como usuários digitais passivos. Já, no grupo 4, observa-se o inverso, com 75% dos membros com QL acima de 1.

A partir dessas informações, pode-se estabelecer que o primeiro grupo é, efetivamente, marcado pela exclusão digital. Infraestrutura quase inexistente, baixo letramento digital, escassez do acesso pessoal e falta de mão-de-obra qualificada dão a tônica. O grupo de municípios com inserção digital marginal marca algumas diferenças iniciais, a exemplo da proporção razoável de escolas com banda larga e de posse de telefone celular, mas segue essencialmente sem infraestrutura, trabalho especializado ou conexão de alta qualidade. A transição para o conjunto de usuários passivos se dá mediante avanço nos pontos menos sofisticados (celular, escolas com banda larga, acessos de baixa velocidade), em relação aos quais não há diferenças substanciais frente aos dois grupos superiores. Para que se transite em direção aos ecossistemas digitais, por sua vez, é necessário avanço em infraestrutura e a integração da camada digital com o restante da economia: são as conexões velozes, a estrutura que lhes dá suporte e a produção local de conteúdo digital que marca a diferença. Por fim, o status de um *hub* é obtido mediante o avanço intenso na infraestrutura pesada, visto com as ASN sediadas. Não obstante, observa-se que as transições de centralidade digital são processos que compreendem o avanço em todas as dimensões, apenas variando a ênfase em diferentes momentos. Mais uma vez, fica ressaltado o caráter multidimensional da digitalização e a necessidade de políticas tanto integradas quanto atentas ao momento específico de cada município para avançar na mesma.

Como forma de validar os clusters encontrados com dados externos, a FIG. 2 mostra a distribuição de características referentes à qualidade da conexão dos usuários.

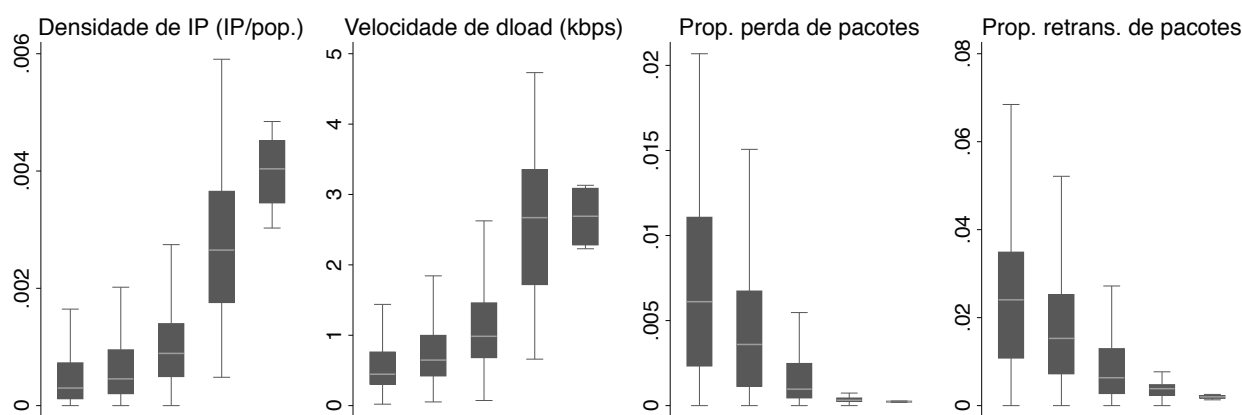


Figura 2 – Características da qualidade da conexão à internet nos diferentes *clusters*

Os resultados evidenciam a robustez dos *clusters* encontrados. Isto é de particular relevância devido à natureza dos dados empregados: trata-se de uma grande base com informação produzida a partir da própria utilização da camada digital. São, assim, dados orgânicos ao nosso objeto, que vão precisamente na direção do que fora estabelecido anteriormente. E, ao contrário das variáveis presentes na FIG. 1, estas não foram utilizadas no cálculo dos *clusters*: logo, o fato de que se distribuem conforme esperado é forte indicativo de que o procedimento estabeleceu uma tipologia sólida, com diferenças relevantes entre os municípios. Tal qual informações anteriores, observa-se que as conexões de velocidade mais alta são presentes essencialmente nos ecossistemas e *hubs* digitais e, mais que isso, que a qualidade da conexão (refletida na proporção de pacotes perdidos e retransmitidos) cresce substancialmente com o avanço na centralidade digital. Comprovada a qualidade da tipologia, passa-se à análise de sua relação com a rede urbana e o espaço no Brasil.

Uma análise visual do mapa (FIG. 3) mostra três fenômenos principais: o quão extensa é a exclusão digital no país; a grande desigualdade da inserção digital; e o quão de perto a hierarquia digital e a hierarquia urbana estão associadas no Brasil. Como já fora indicado anteriormente, os municípios digitalmente excluídos ou com inserção marginal formam a grande maioria, respondendo por, respectivamente, 40.0 e 45.5% do total. Ou seja, quase 90% dos municípios brasileiros não podem ser considerados sequer usuários passivos de conteúdo digital. Este número é tanto maior nas regiões Norte, com 73.5% e 22.7% nos *clusters* 1 e 2, e Nordeste, com 79.8% e 18.0% nestas categorias. A situação é melhor nas demais regiões, com a proporção de municípios pertencentes ao *cluster* 1 variando entre no mínimo 10.4%, no Sul, e no máximo 16.7%, no Sudeste – mas, ainda assim, em todas elas mais da metade dos municípios tem apenas inserção digital marginal. E não há nenhuma região com mais de 2.7% de municípios na categoria de ecossistemas digitais, o que revela com clareza o quão necessárias são políticas ativas de promoção da digitalização.

Voltando os olhos para a relação entre centralidade urbana e digital, nota-se com clareza como a camada digital desenvolveu-se essencialmente a partir de uma rede urbana pré-existente, o que só vem a reforçar tal hierarquia. Primeiramente, vê-se que apenas no Sudeste e Sul há *hubs* digitais. Mesmo Recife, uma cidade com presença de serviços modernos (QL de 1.6), universidades e estrutura produtiva sofisticada, não consegue se estabelecer como tal. Ocorre que sua infraestrutura digital é consideravelmente inferior à dos *hubs*: há apenas 19 ASN sediadas lá, 12 participando do PIX local, e 22 centrais EILD, frente a médias de, respectivamente 152, 114 e 175 para os *hubs*. A capacidade de funcionar como um centro digital de importância nacional é bastante restrita nessas condições, o que sem dúvida impacta o baixo desenvolvimento digital da região Nordeste como um todo.

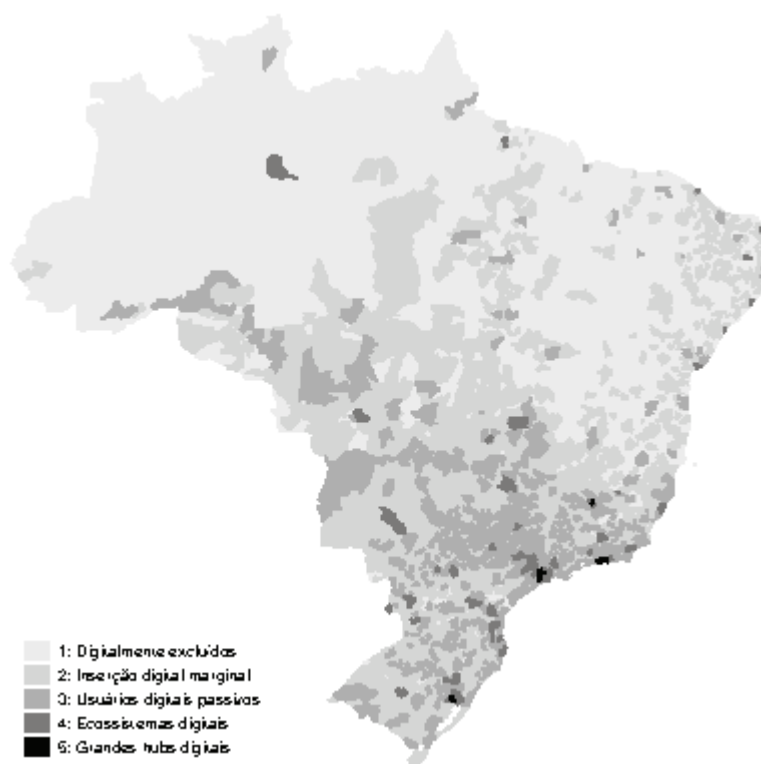


Figura 3 – Mapeamento dos municípios conforme pertencimento aos cinco clusters de digitalização

Fonte: Elaboração nossa a partir de dados de várias fontes.

Ademais, nas regiões Norte e Nordeste, dos 14 municípios considerados ecossistemas digitais (0.6% do total), nove são capitais estaduais e quatro fazem parte de suas regiões metropolitanas, apenas Campina Grande, na Paraíba, fugindo à regra. E, no país como um todo, apenas sete capitais estaduais não são *hubs* ou ecossistemas digitais. Os demais municípios neste *cluster* encontram-se em três categorias principais: entornos de grandes metrópoles, o que é visualmente observável no mapa, a exemplo de Santana da Parnaíba (SP), Nova Lima (MG) e São José dos Pinhais (PR); cidades médias com estrutura produtiva desenvolvida, como Ribeirão Preto (SP), Juiz de Fora (MG) e Maringá (PR); e distritos industriais, centros de serviços ou cidades universitárias,<sup>13</sup> como Ipatinga (MG), São Bernardo do Campo (SP) e Santa Maria (RS). Ou seja, praticamente só os espaços que já tinham alguma centralidade na rede urbana brasileira, seja por razões políticas, econômicas ou educacionais, ou ainda por alguma combinação destas, conseguiram atingir centralidade digital. Outrossim, municípios que transformaram em ecossistemas ou, principalmente, *hubs* digitais, polarizam seu entorno e conseguem estimular sua digitalização.

Esta constatação é reforçada mediante o cruzamento da categorias de hierarquia urbana e os clusters aqui encontrados, o que é feito na TAB. 2. A primeira seção da tabela considera que os municípios pertencentes a ACP tenham o mesmo nível de hierarquia urbana desta, enquanto a segunda seção tabela exclui municípios que sejam partes de ACP, mas não suas sedes.

O grau de ajuste entre as duas hierarquias é bastante alto, excluídos ou não os municípios pertencentes a ACP. O cruzamento entre categorias “de mesmo nível” (e.g., centros locais e municípios digitalmente excluídos) em geral é a categoria mais frequente na linha e, em todos os casos, esta proporção é acima daquela observada para o conjunto dos municípios. Por exemplo, vê-se que (sem) com exclusão de ACP, (23.2%) 41.4% das capitais regionais são ecossistemas digitais, comparado a (1.7%) 0.8% para o país, e (57.3%) 57.3% dos centros subregionais tem inserção marginal passiva, frente a (13.7%) 11.5% para o total. Ademais, parcela muito pequena dos centros locais (6.7%) têm inserção digital mesmo que passiva, e apenas sete municípios que não fossem ao menos capitais regionais conseguiram estabelecer-se como ecossistemas digitais. Uma vez mais, isso reforça a associação entre a

<sup>13</sup> A influência do *backbone* da Rede Nacional de Pesquisa (RNP), que conecta todas as universidades federais com alta velocidade, parece ser um determinante importante neste aspecto.



rede urbana e a centralidade digital dos municípios, a demonstrar que a digitalização tem avançado de maneira estritamente relacionada à hierarquia urbana brasileira. Reconhecido o caráter desigual desta, a análise sugere fortemente a necessidade de que a política regional e de desenvolvimento atente-se para a camada digital e como estimulá-la, sob risco de reforçar traços concentradores de renda, riqueza, acesso à informação e oportunidades de trabalho.

Tabela 2 – Cruzamento entre centralidade urbana e clusters de digitalização

Hierarquia urbana		Cluster de digitalização					Total coluna
		1	2	3	4	5	
Sem exclusão de municípios pertencentes a ACP	Centro local	2078 46.5	2096 46.9	298 6.7	0 0.0	0 0.0	<b>4472</b> <b>100.0</b>
	Centro de zona	123 22.1	255 45.9	177 31.8	1 0.2	0 0.0	<b>556</b> <b>100.0</b>
	Centro subregional	17 10.4	47 28.7	94 57.3	6 3.7	0 0.0	<b>164</b> <b>100.0</b>
	Capital regional	7 3.7	39 20.5	100 52.6	44 23.2	0 0.0	<b>190</b> <b>100.0</b>
	Metrópole	0 0.0	40 21.9	95 51.9	44 24.0	4 2.2	<b>183</b> <b>100.0</b>
	<b>Total linha</b>	<b>2225</b> <b>40.0</b>	<b>2477</b> <b>44.5</b>	<b>764</b> <b>13.7</b>	<b>95</b> <b>1.7</b>	<b>4</b> <b>0.1</b>	<b>5565</b> <b>100.0</b>
Com exclusão de municípios pertencentes a ACP	Centro local	2078 46.5	2096 46.9	298 6.7	0 0.0	0 0.0	<b>4472</b> <b>100.0</b>
	Centro de zona	123 22.1	255 45.9	177 31.8	1 0.2	0 0.0	<b>556</b> <b>100.0</b>
	Centro subregional	17 10.4	47 28.7	94 57.3	6 3.7	0 0.0	<b>164</b> <b>100.0</b>
	Capital regional	0 0.0	6 8.6	35 50.0	29 41.4	0 0.0	<b>70</b> <b>100.0</b>
	Metrópole	0 0.0	0 0.0	0 0.0	8 66.7	4 33.3	<b>12</b> <b>100.0</b>
	<b>Total linha</b>	<b>2218</b> <b>42.1</b>	<b>2404</b> <b>45.6</b>	<b>604</b> <b>11.5</b>	<b>44</b> <b>0.8</b>	<b>4</b> <b>0.1</b>	<b>5274</b> <b>100.0</b>

Fonte: Elaboração própria a partir de várias fontes.

Nota: os percentuais referem-se ao total das linhas.

Passa-se ao último momento deste trabalho, em que se refina a análise da relação entre atributos urbanos e centralidade digital. Comprovada a robustez da tipologia desenvolvida e sua associação com a rede urbana, busca-se explorar com mais detalhes os determinantes da transição entre diferentes “estágios” da digitalização.

## 5 FATORES ASSOCIADOS À CENTRALIDADE DIGITAL

Conforme indicado na seção de métodos, ajustou-se um modelo linear para a digitalização (medida como o primeiro componente principal analisado acima) empregando as mesmas variáveis do *logit* multinomial discutido abaixo, que não é reportado por questões de espaço. Não obstante, nota-se que os resultados encontrados foram bastante satisfatórios, com alto grau de ajuste ( $R^2$  de 0.88) e significância das variáveis. Dada a não linearidade do fenômeno em tela, conforme amplamente evidenciado ao longo deste trabalho, o emprego do *logit* multinomial é mais informativo, por lançar luz sobre os determinantes das transições entre diferentes níveis.

Os modelos *logit* multinomiais tem uma dificuldade de interpretação associada ao fato de que variáveis que são significantes para a transição entre duas dadas categorias podem não o ser para outras duas quaisquer. A não linearidade do modelo também implica que o impacto de determinada variável depende do valor de todas as outras, podendo sofrer grandes variações. Mais do que isso, o fato de um coeficiente ser positivo e significativo *não* indica que a probabilidade da categoria a que se refere aumente, mas apenas que a probabilidade relativa entre tal categoria e a de referência tende a crescer (o que pode se dever, por exemplo, a uma redução simultânea das duas, com a de referência decrescendo

mais rapidamente). Portanto, como forma de auxiliar a interpretação são reportados os resultados do modelo em três formas: i) uma tabela de coeficientes, tomado como referência o *cluster* 1, e com estatísticas de ajuste; ii) uma matriz indicando os coeficientes e a significância de variáveis selecionadas para transições entre todas as categorias; e iii) gráficos simulando o impacto de algumas variáveis.

Os parâmetros de controle dos gráficos são escolhidos para gerar uma situação analítica e politicamente relevante. Tratam-se de condições propícias para um município tornar-se ecossistema digital, em uma região marcada pela exclusão, ao mesmo tempo em que não são altas o suficiente para *garantir* a digitalização. Utiliza-se uma capital regional do Nordeste, com alto número de artigos publicados e demais variáveis em seus valores médios por município ou iguais à média agregada do país. Por fim, cabe lembrar que, em vista do reduzido número de observações referentes a *hubs* digitais (quatro), não é possível incluí-los no modelo.

Tabela 3 – Resultados do *logit* multinomial: transições de centralidade digital como função de atributos urbanos

Variável	Coeficientes, significância * * * * * e (estatística <i>t</i> )				
	Cluster				
	2	3	4		
educ_sup†	90.15*** (15.35)	143.8*** (18.03)	154.4*** (13.01)		
divers	8.132*** (13.82)	26.49*** (11.56)	46.13* (2.04)		
Indensidade	0.774*** (13.89)	1.897*** (17.86)	3.348*** (14.28)		
smod†	29.38*** (4.36)	66.66*** (6.59)	148.4*** (7.21)		
regicb	1.247** (2.58)	2.209*** (4.02)	4.318*** (5.34)		
pind	12.60** (2.76)	25.65*** (4.82)	36.78*** (6.03)		
martigos	-0.241 (-0.65)	1.049* (2.16)	1.319 (1.95)		
Região: Norte	1.225*** (5.95)	2.664*** (4.22)	3.772** (2.64)		
Região: Sudeste (sem SP)	1.905*** (13.18)	4.681*** (11.05)	5.557*** (6.08)		
Região: Centro-Oeste	3.762*** (19.12)	6.703*** (15.08)	9.523*** (9.72)		
Região: Sul	3.748*** (15.53)	6.974*** (13.52)	8.700*** (7.21)		
Região: SP	3.877*** (10.34)	8.069*** (14.27)	8.942*** (8.51)		
_cons	-12.40*** (-24.05)	-40.29*** (-19.69)	-76.76*** (-3.67)		
Pseudo-R <sup>2</sup>	0.643				
Proporção corretamente classificada‡	0.836				
N	5561				
<i>Cluster</i>	<i>Positive predictive value</i>	<i>Negative predictive value</i>	<i>Sensitivity</i>	<i>Specificity</i>	<i>Tjur coefficient of discrimination@</i>
1	0.859	0.918	0.879	0.904	0.687
2	0.823	0.861	0.828	0.857	0.550
3	0.814	0.961	0.753	0.973	0.651
4	0.814	0.995	0.737	0.997	0.704

Nota: região de referência: Nordeste; categoria de referência: cluster 1; ver Tabela 1 para códigos das variáveis; †: medidas como proporções, e não QL; @: estatística entre 0 e 1 que representa o ajuste do modelo à categoria em questão (Tjur 2009); ‡: sucesso definido quando a maior probabilidade prevista corresponde ao *cluster* observado; \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Tabela 4 – Coeficientes e significância de variáveis selecionadas do modelo *logit* multinomial para transições entre diferentes categorias

Variável	Cluster “de destino”	Cluster de base		
		1	2	3
educ_sup†	2	90.15***		
	3	143.84***	53.69***	
	4	154.44***	64.29***	10.59
divers	2	8.13***		
	3	26.49***	18.36***	
	4	46.13*	38	19.64
Indensidade	2	0.77***		
	3	1.9***	1.12***	
	4	3.35***	2.57***	1.45***
smod†	2	29.38***		
	3	66.66***	37.28***	
	4	148.43***	119.05***	81.77***
regicb	2	1.25**		
	3	2.21***	0.96***	
	4	4.32***	3.07***	2.11***
pind	2	12.6**		
	3	25.65***	13.05***	
	4	36.78***	24.18***	11.13***
martigos	2	-0.24		
	3	1.05*	1.29***	
	4	1.32	1.56**	0.27

Nota: ver Tabela 1 para códigos das variáveis; †: medidas como proporções, e não QL; \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

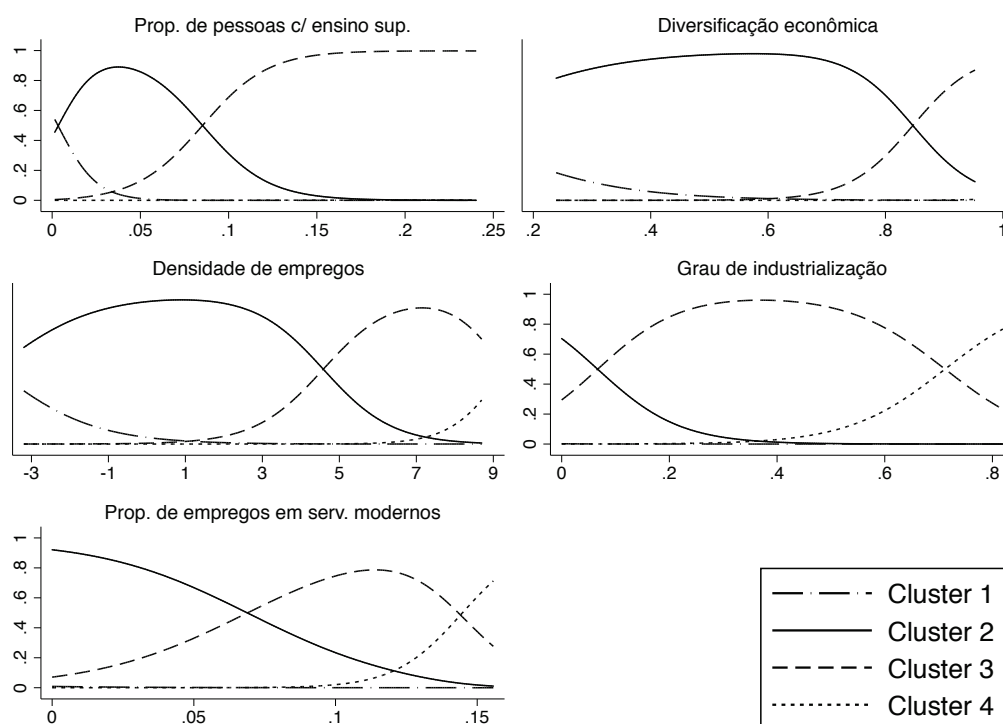


Figura 4 – Impacto estimado da especialização em serviços modernos, diversificação econômica, difusão do ensino superior, densidade de empregos e grau de industrialização sobre a centralidade digital

Notas: as condições de controle são proporção de pessoas com ensino superior e em serviços modernos: agregado do país, QL = 1 ( $educ\_sup = 0.07$ ,  $smod = 0.05$ ); diversificação econômica: média ( $divers = 0.80$ ); densidade de empregos: logaritmo da média ( $Inden = 3.9$ ); capital regional ( $regicb = 1$ ); industrialização: média ( $pind = 0.01$ ); número de artigos publicados: acima da média estadual ( $martigos = 1$ ); região: Nordeste.

Antes de interpretar o modelo, cabe observar o alto grau de ajuste, conforme diversas medidas. Mais que isso, o ajuste é alto para todas as categorias de municípios, a demonstrar que as variáveis

independentes captam adequadamente as várias transições em tela – o que não é trivial em um modelo de quatro categorias com números bastante diferentes de observações em cada uma. A significância de grande parte das variáveis reforça a observação. Isto ressalta, em termos globais, como a digitalização é um processo estritamente associado à urbanização, dado ser possível determinar aquela a partir desta. Outrossim, permite a interpretação dos resultados com a segurança de se estar captando as determinantes essenciais do fenômeno.

Pelas razões expostas acima, o impacto das variáveis independentes será interpretado a partir dos resultados expostos na TAB. 4 e na FIG. 4.

Cabem duas observações gerais antes de nos dedicarmos às análises específicas. A primeira é que não há nenhuma reversão de trajetória, no sentido de que o impacto de um aumento nas variáveis, se observável, vai sempre no sentido de aumentar a probabilidade de inserção nos *clusters* mais altos (do 1 para o 2, 3 ou 4, do 2 para o 3 ou 4...). Isto sugere que todas as dimensões de centralidade urbana – escala, diversificação econômica, presença de serviços modernos, qualificação da mão-de-obra, estrutura produtiva avançada... – estimulam, ainda que de diferentes maneiras, a digitalização. Em outras palavras, o aumento da centralidade urbana jamais reduz a digitalização. A segunda observação geral refere-se à diferenciação do impacto das variáveis consideradas: algumas atuam muito mais nas transições “iniciais”, enquanto outras nas “superiores”. Isto indica que a digitalização é um fenômeno complexo, a requerer uma conjunção de fatores para que os agentes em determinado espaço possam desfrutar de seus diversos ganhos.

Dentro deste quadro, observa-se que as transições iniciais são determinadas essencialmente por questões educacionais. Dadas as condições de controle dos gráficos a única variável capaz de determinar alta probabilidade de exclusão digital é a baixa proporção de pessoas com ensino superior, e esta variável tem os maiores coeficientes de transição para os *clusters* 2 e 3 a partir do *cluster* 1 ou 2 (TAB. 4, como demais coeficientes mencionados). Com menos de 5% da população com ensino superior, condição partilhada por mais de três quartos dos municípios, há mais de 85% de chance de o município inserir-se nos *clusters* 1 ou 2. Isto indica que o determinante fundamental, ou o gargalo principal, para se tornar um usuário passivo de conteúdo digital está na qualificação geral da população: tratam-se de condições básicas de acesso à informação, letramento digital etc.

A diversificação econômica tem papel intermediário. Quando atinge níveis muito baixos, está associada a maior probabilidade de exclusão digital, mas seu impacto principal encontra-se na transição para a categoria de usuários digitais passivos. Isto muito provavelmente está associado ao caráter transversal da camada digital, que pode (e em certa medida necessita) integrar-se às diversas atividades econômicas. Sem que haja algum grau de diversificação, não há escopo o suficiente sequer para utilização passiva do digital, ao mesmo tempo que estruturas diversificadas levam a algum grau de digitalização. O desenvolvimento científico tem papel semelhante, visível através do número de artigos publicados. Trata-se de uma variável de relativamente pouco impacto, que estimula as transições 1, 2 >> 2, 3 – ou seja, transições intermediárias.<sup>14</sup>

Por outro lado, a educação em si não leva a transições mais altas, e nem tampouco o faz a diversificação econômica ou o desenvolvimento científico. A condição de um ecossistema digital, em que há efetiva produção de conteúdo digital, capaz de polarizar o espaço e puxar o desenvolvimento regional, somente é possível a partir de um processo econômico mais amplo. É necessário que o espaço urbano atinja grande densidade, ou, especialmente, que o digital se integre aos serviços e à indústria – é o desenvolvimento econômico do espaço urbano que lhe permite assumir centralidade digital. Em outros termos, a centralidade digital está associada à integração de tal camada aos serviços e atividades produtivas, e não pessoais. Tal fato é claramente perceptível através dos gráficos, que demonstram como a proporção de trabalhadores em serviços modernos e a industrialização são as mais influentes sobre a probabilidade de inserção no *cluster* 4. De forma ainda mais patente, nota-se como os coeficientes da variável educacional são bastante altos para transições 1, 2 >> 2, 3, mas decaem radicalmente e deixa de ser significativo para a transição 3 >> 4. A partir disso pode-se inferir que os altos coeficientes desta variável para o *cluster* 4 a partir dos *clusters* 1 ou 2 referem-se à redução da probabilidade de pertencer a estes, e

<sup>14</sup> O fato de não ser relacionada à transição para o *cluster* 4 pode ser indicativo da pouca integração da universidade brasileira com a economia e a sociedade.

não ao aumento da probabilidade referente àquele. De fato, somente as variáveis de centralidade urbana, escala, serviços modernos e industrialização são determinantes da transição para ecossistemas digitais, com grande destaque para a presença de serviços modernos.

Não se deve concluir, contudo, que a educação é desnecessária para atingir centralidade digital. Embora não determine a transição para o *cluster* 4, mesmo com alta proporção de trabalhadores em serviços modernos (15%, o triplo da proporção nacional) e demais variáveis iguais às dos gráficos, caso haja apenas 1% de pessoas com ensino superior a maior probabilidade é de inserção no *cluster* 2. Ou seja, é necessário conjugar educação com avanço da estrutura produtiva: a digitalização é um fenômeno multifacetado.

O que o conjunto desses resultados mostra é que se deve pensar a centralidade digital como um atributo bastante sofisticado *do tecido urbano*. Trata-se de um fenômeno complexo, que em diferentes dimensões e estágios interage com diversos aspectos da urbanização. Atingir alto grau de digitalização passa, assim, por uma estratégia ampla de desenvolvimento, que necessariamente deve envolver educação, produção local de software e serviços de TI, letramento digital e infraestrutura especializada, mas, mais que isso, deve sobretudo vir a mirar a integração da camada digital com a estrutura produtiva mais ampla. O fato de que somente as atividades econômicas de alto valor agregado, a centralidade urbana e grandes escalas do mercado de trabalho determinam a transição para ecossistemas digitais comprova a associação.

## 6 CONSIDERAÇÕES FINAIS

Neste trabalho, o recurso a múltiplas bases de dados permitiu uma visão ampla do processo de digitalização no Brasil, tendo sido os resultados validados externamente por *big data* relacionado à utilização e difusão da internet, comprovando sua robustez. Em termos dos resultados, o primeiro ponto é a própria mensuração da digitalização ou centralidade digital. Não há nenhum outro estudo para o país com cobertura compreensiva nacional trabalhados na escala municipal; tampouco qualquer estudo que contemple os múltiplos aspectos da centralidade digital, a saber, utilização e produção de conteúdo digital, infraestrutura digital e letramento digital. Os resultados também demonstram o caráter multifacetado da digitalização. Dois procedimentos parecem confirmar esta questão. Primeiramente, todas as variáveis têm *loadings* positivos na primeira dimensão do PCA, mostrando uma correlação direta entre infra-estrutura e iluminação digital. Em segundo lugar, o logit multinomial revela como os diversos atributos da urbanização são importantes para elevar e diversificar a cobertura da digitalização. Contudo tais atributos agem de distintas maneiras. Notadamente, educação é a principal determinante das transições iniciais que superam a exclusão digital por completo. Para atingir a condição de um verdadeiro ecossistema digital, contudo, em que há desenvolvimento autônomo de conteúdo, é necessário que a camada digital se integre a atividades produtivas, sejam industriais ou de serviços modernos.

Segundo, encontramos altíssima desigualdade da digitalização pelo espaço. Aproximadamente 85% dos municípios são marcados pela exclusão digital quase completa, enquanto cerca de 14% são usuários apenas passivos de conteúdo digital. Não mais que 99 municípios são espaços em que há autonomia para a produção de conteúdo, com efetivo desenvolvimento da camada digital. Complementarmente, a centralidade digital vem se desenvolvendo de maneira estritamente associada à rede urbana pré-existente, isto é, reforçando a desigualdade regional. A centralidade urbana, conforme definida pelo IBGE, é fortemente relacionada à mensuração de centralidade digital por meio de *clusters*. Ademais, variáveis de atributos urbanos, como – entre outros – presença de serviços modernos e diversificação econômica, parecem determinar a própria digitalização.

Quanto às possíveis ilações de políticas públicas, de maneira central, fica evidente que somente estratégias amplas e integradas podem permitir a maior difusão da digitalização pelo espaço brasileiro. É necessário combinar, no mínimo, inclusão digital por meio de educação, espraiamento da infraestrutura dedicada e políticas inovadoras para a utilização de tecnologia digital no âmbito produtivo. A ênfase depende, naturalmente, do nível de hierarquia digital de cada espaço, conforme mostrado nos diagramas de transição.



Vale dizer, o que parece ser o grande desafio é estabelecer ações que permitam o adensamento da rede digital tanto por intermédio da diminuição do número de espaços caracterizados por exclusão digital *tout court*, criando um nível médio de acesso e letramento digital robusto.

Mais que isto, é preciso encarar a mais difícil das transições, ou seja, a dos níveis intermediários – acesso irrestrito, mas sem produção de conteúdo, para os de produção de conteúdo digital e completa iluminação digital, que poderia criar sinergias produtivas e de infraestrutura. Tais sinergias são fundamentais para alcançar escalas econômicas e de produtividade necessárias ao desenvolvimento de um sistema de inovação – mesmo que incompleto – capaz de justificar, com potencialidades de mercado, a expansão dos espaços capazes de se caracterizar como ecossistemas digitais completos.

Neste ponto é importante destacar uma questão que, invariavelmente, vem sendo *taken for granted* pela literatura. Vale dizer, a de que a brutal virtualização das relações sociais do mundo contemporâneo faz com que as redes sociais e suas hierarquias subjacentes apresentem uma redução mais que proporcional da necessidade de bases físicas e de verticalidades em sua atuação. O argumento da horizontalização das ações políticas, e suas repercussões na estruturação do território, é presente hoje na quase totalidade da discussão contemporânea do planejamento de base territorial. Este trabalho, contudo, parece confirmar o que Žižek (2016) em **O Sujeito Incômodo**: o centro ausente da ontologia política já afirmava em 1999, i.e., para que a horizontalidade seja viável é necessário um “Estado Invisível” por detrás, capaz de garantir e fornecer não apenas leis e sistemas de saúde e educação, como também, e principalmente, infraestrutura física: energia, estradas, portos, saneamento, e no nosso caso, sistemas de comunicação que viabilizem a existência desta virtualidade. E é no território que esta infraestrutura não virtual se manifesta: nas redes e suas hierarquias urbanas que parecem absolutamente – quando não os exacerba – aderentes aos níveis hierárquicos de base não virtual. No caso brasileiro, como vimos, todos os níveis da hierarquia digital são absolutamente aderentes ao determinado pela REGIC, sendo o nível superior e o nível médio ainda mais concentrados do ponto de vista digital. Os desafios desta conclusão, do ponto de vista das políticas públicas, nos parecem cruciais.

## 7 REFERÊNCIAS

- ANATEL. **Acessos ao Serviço de Comunicação Multimídia**. Brasília: Anatel, 2014a.
- \_\_\_\_\_. **Relação de localização geográfica dos centros de fios - EILD**. Brasília: Anatel, 2014e.
- \_\_\_\_\_. **Cobertura de ERB no Brasil**. Brasília: Anatel, 2014d.
- \_\_\_\_\_. **Área de prestação por faixa de frequência**. Brasília: Anatel, 2014c.
- \_\_\_\_\_. **Acessos a TV por assinatura**. Brasília: Anatel, 2014b.
- BAUMOL, W. J. The arts in the “new economy”. In: GINSBURG, V. A.; THROSBY, D. (Org.). **Handbook of the Economics of Art and Culture**. [S.l.]: Elsevier, 2006, V. 1, p. 339–358.
- COMBES, P. P. Economic structure and local growth: France, 1984-1993. **Journal of Urban Economics**, 2000. v. 47, p. 329–355.
- GARCIA, A. G.; SIMÕES, R. **Desafios e estratégias para a inclusão digital – Programa Banda Larga**. Brasília: CGEE, 2012.
- GLAESER, E. L.; KOLKO, J.; SAIZ, A. Consumer city. **Journal of Economic Geography**, 1 jan. 2001. v. 1, n. 1, p. 27–50.
- HUTTER, M. Information goods. In: TOWSE, R. (Org.). **A handbook of cultural economics**. Cheltenham: Edward Elgar Publishing, 2003, p. 263–268.
- IBGE. **Regiões de influência das cidades 2007**. Rio de Janeiro: IBGE, 2008.
- \_\_\_\_\_. **Censo demográfico de 2010**. Brasília: IBGE, 2010.
- \_\_\_\_\_. **MUNIC 2012**. Brasília: IBGE, 2012a.
- \_\_\_\_\_. **Produto interno bruto dos municípios 2011**. Brasília: IBGE, 2012b.
- \_\_\_\_\_. **Divisão urbano regional**. Brasília: IBGE, 2013.
- INEP. **Censo da educação básica 2013**. Brasília: INEP, 2013.
- ISI. **Web of Science publication database**. [S.l.]: ISI/Thomson Scientific, 2010.

LIMA, A. C. **Desenvolvimento regional e fluxos migratórios no Brasil: uma análise para o período 1980-2010**. Belo Horizonte: CEDEPLAR, Universidade Federal de Minas Gerais, 2012. Doutorado em Economia.

LIPSEY, R. G.; CARLAW, K. I.; BEKAR, C. T. **Economic Transformations: General Purpose Technologies and Long-term Economic Growth**. Oxford: Oxford University Press, 2006.

MACHADO, A. F.; SIMÕES, R. F.; DINIZ, S. C. Urban Amenities and the Development of Creative Clusters: The Case of Brazil. **Current Urban Studies**, 2013. v. 01, n. 04, p. 92–101.

MACKINNON, J. G.; WHITE, H. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. **Journal of Econometrics**, set. 1985. v. 29, n. 3, p. 305–325.

M-LAB. **Network Diagnosis Test (NDT)**. [S.l.]: M-Lab, 2014.

OBSERVATÓRIO SOFTEX. **Software e serviços de TI: a indústria brasileira em perspectiva**. Campinas: Softex, 2012.

OECD. **OECD communications outlook 2013**. [S.l.]: OECD Publishing, 2013.

\_\_\_\_\_. **Measuring the Digital Economy: A New Perspective**. Paris: OECD publishing, 2014.

PTTBR. **Participantes do PTTMetro**. Brasília: PTTbr, 2014b.

\_\_\_\_\_. **Distribuição geográfica de ASNs no Brasil**. Brasília: PTTbr, 2014a.

ROBINSON, L. *et al.* Digital inequalities and why they matter. **Information, Communication & Society**, 4 maio. 2015. v. 18, n. 5, p. 569–582.

SCHONLAU, M. The clustergram: A graph for visualizing hierarchical and nonhierarchical cluster analyses. **Stata Journal**, 2002. v. 2, n. 4, p. 391–402.

SORJ, B.; GUEDES, L. E. Exclusão digital: problemas conceituais, evidências empíricas e políticas públicas. **Novos Estudos - CEBRAP**, jul. 2005. n. 72, p. 101–117.

TELETIME. **Atlas brasileiro de telecomunicações**. São Paulo: Teletime, 2014.

TJUR, T. Coefficients of Determination in Logistic Regression Models—A New Proposal: The Coefficient of Discrimination. **The American Statistician**, 1 nov. 2009. v. 63, n. 4, p. 366–372.

ZACKIEWICZ, M. A economia do software e a digitalização da economia. **Revista Brasileira de Inovação**, 10 set. 2015. v. 14, n. 2 jul/dez, p. 313–336.

ŽIŽEK, S. **O sujeito incômodo: o centro ausente da ontologia política**. Rio de Janeiro: Boitempo Editorial, 2016.

## 8 ANEXO 1: LISTAGEM DAS CLASSIFICAÇÕES EMPREGADAS

Quadro 2 – Grupos ocupacionais de PROFSSs

<b>Código</b>	<b>Titulação</b>
1330	Dirigentes de serviços de tecnologia da informação e comunicações
2153	Engenheiros em telecomunicações
2356	Instrutores em tecnologias da informação
2511	Analistas de sistemas
2512	Desenvolvedores de programas e aplicativos (software)
2513	Desenvolvedores de páginas de internet (web) e multimídia
2514	Programadores de aplicações
2519	Desenvolvedores e analistas de programas e aplicativos (software) e multimídia não classificados anteriormente
2521	Desenhistas e administradores de bases de dados
2522	Administradores de sistemas
2523	Profissionais em rede de computadores
2529	Especialistas em base de dados e em redes de computadores não classificados anteriormente
3511	Técnicos em operações de tecnologia da informação e das comunicações
3512	Técnicos em assistência ao usuário de tecnologia da informação e das comunicações
3513	Técnicos de redes e sistemas de computadores
3514	Técnicos da web
4132	Operadores de entrada de dados

Fonte: Adaptação da proposta do Observatório Softex (2012).

Quadro 3 – Atividades Econômicas da IBSS

<b>Código</b>	<b>Descrição</b>
62000	Atividades dos serviços de tecnologia da informação
63000	Atividades de prestação de serviços de informação
95010	Reparação e manutenção de equipamentos de informática e comunicação

Fonte: Adaptação da proposta do Observatório Softex (2012).

Quadro 4 – Atividades econômicas compreendidas nos serviços modernos

<b>Código</b>	<b>Descrição</b>
61	Telecomunicações
64	Atividades de serviços financeiros
65	Seguros, resseguros, previdência complementar e planos de saúde
66	Atividades auxiliares dos serviços financeiros, seguros, previdência complementar e planos de saúde
69	Atividades jurídicas, de contabilidade e de auditoria
70	Atividades de consultoria em gestão empresarial
71	Serviços de arquitetura e engenharia; testes e análises técnicas
72	Pesquisa e desenvolvimento científico
73	Publicidade e pesquisas de mercado
74	Outras atividades profissionais, científicas e técnicas
85014	Educação superior

Fonte: Adaptação de Simões e Freitas (2014).

Quadro 5 – Áreas de cursos superiores de interesse direto para o setor de software e serviços de TI

<b>Código</b>	<b>Área</b>
481	Ciência da Computação
482	Uso do computador
523	Eletrônica e Automação

Fonte: Observatório Softex (2012).