

Universidade Federal do Rio de Janeiro

Escola Politécnica

MBA em Engenharia de Computação e Sistemas  
(MBCA)

# **Computação Cognitiva: Relações Sociais, Mídias Digitais e o Big Data**

**Autor:**

---

Mario da Costa Bacellar

**Orientador:**

---

**Prof. Edilberto Strauss, Ph.D.**

**Examinador(es):**

---

Prof. Edilberto Strauss, Ph.D.

---

Prof. Flávio Luis de Mello, **D. Sc.**

**MBCA – Turma 11  
Agosto de 2015**



## **AGRADECIMENTO**

Longo foi o caminho que percorri para chegar até aqui. Nem mesmo eu tive essa pretensão. Entre vários abandonos e recomeços do ensino superior, concluo esse curso de pós-graduação com o sentimento de dever cumprido, em uma pequena etapa, da interminável busca pelo conhecimento.

Dedico este estudo à minha mãe, que na condição de mãe, e professora por profissão, incansavelmente me orientou a nunca deixar de evoluir, tanto as faculdades do espírito quanto os ensinamentos disseminados na sociedade em geral. Sua contribuição se deu não apenas nas cobranças das tarefas do ensino fundamental e médio, mas principalmente no incentivo da prática de Boa Leitura. Prática fundamental para a conclusão desse curso, face ao grande volume de textos pesquisados para entregar com excelência todos os trabalhos das disciplinas que envolveram esse curso.

Mãe, creio que não existam palavras significativas o suficiente que permitam lhe agradecer com justiça e o devido merecimento. Sua ajuda e seu apoio foram para mim de valor inestimável, esse estudo foi a forma que encontrei de retribuir e também de lhe prestar esta humilde e singela homenagem.

Com todo o carinho e do fundo do meu coração, Mãe, eu lhe agradeço, e pelo resto da minha vida lhe agradecerei!

Muito obrigado, Regina da Costa Bacellar!

“O descontentamento é o primeiro  
passo na evolução de um homem  
ou de uma nação.” – **Oscar Wilde**

## RESUMO

Vivemos uma nova Renascença. A imprensa é digital, o novo mundo (antes a América) agora é o ciberespaço, e a nova Reforma vem com a computação cognitiva nas redes sociais. Os dados que trafegam no ciberespaço é o novo petróleo, quem souber refinar terá mais vantagem competitiva. Análises preditivas de sentimento, em tempo real, em cima de gigantescas massa de dados trazidos pelas redes sociais, estimulam práticas de marketing cada vez mais invasivas, quebrando os atuais paradigmas de consulta em banco de dados. A computação cognitiva incorpora a mesma forma que nosso cérebro interpreta as informações e toma decisões. A tecnologia tem a missão de tornar a comunicação homem-máquina cada vez mais natural. As pessoas terão soluções para acelerar sua cognição, transformando conhecimento em sabedoria. Este estudo teórico procura desvendar a era dos computadores cognitivos e mostrar como a tecnologia está sendo influenciada pelos processos de cognição do cérebro.

**Palavras-Chave:** novas mídias, mídias digitais, relações sociais, relações virtuais, redes sociais, computação cognitiva, watson, big data.

## ***ABSTRACT***

*We live in a new renaissance period. The press is digital, the new World (before named America) is now the cyberspace and the new Reform come with cognitive computing on social networks. The data who travels over the cyberspace is the new oil, who knows how refine it will have more competitive advantage. Predictive of sentiment analysis, real-time, on top of gigantic amount of data brought by social networks, stimulate increasingly invasive marketing practices, breaking the current query paradigms in the database. Cognitive computing incorporates into the technology the same way that our brain interprets the information and makes decisions. Technology has a mission to make the communication between man and machine more natural. People will have solutions to accelerate their cognition, transforming knowledge into wisdom. Motivated by this scenario, I built this theoretical study to unravel the era of cognitive computers and show how technology is being influenced by the brain cognitive processes.*

**Keywords:** *new media, digital media, social relations, virtual relations, social networks, cognitive processes, cognitive computing, IBM Watson, big data.*

## **SIGLAS**

UFRJ – Universidade Federal do Rio de Janeiro

NM – Novas Mídias

MD – Mídia Digital

BD – Big Data

BI – Business Intelligence

# Sumário

<b>Capítulo 1 Introdução.....</b>	<b>X</b>
1.1 – Tema.....	X
1.2 – Delimitação.....	X
1.3 – Justificativa.....	X
1.4 – Objetivos.....	xi
1.5 – Metodologia.....	xi
1.6 – Descrição.....	xii
<b>Capítulo 2 A Formação do Conhecimento.....</b>	<b>xiii</b>
2.1 – Contexto Histórico.....	xiii
2.2 – Cognitivismo.....	xiv
2.3 – Cognição.....	xv
2.4 – Construtivismo.....	xvii
<b>Capítulo 3 Mídias Digitais.....</b>	<b>xx</b>
3.1 – Contexto Histórico.....	xx
3.2 – A Cognitividade nas Mídias Digital.....	xxii
<b>Capítulo 4 Big Data.....</b>	<b>xxiii</b>
4.1 – Definição.....	xxiii
4.2 – Caraterísticas.....	xxiv
4.3 – Big Data e o Business Intelligence (BI).....	xxviii
4.4 – Análises Preditivas.....	xxx
4.5 – O Sentimento nas Redes Sociais.....	xxxii
4.6 – Transformando Conhecimento em Sabedoria.....	xxxiii
<b>Capítulo 5 Conclusão e Trabalhos Futuros.....</b>	<b>xxxvii</b>
5.1 – Conclusão.....	xxxvii
5.2 – Trabalhos Futuros.....	xxxix
<b>Bibliografia.....</b>	<b>xl</b>



# Lista de Figuras

Figura 2.1.1 – Contingência Tríplice de Skinner.....	10
Figura 2.3.1 – O conhecimento entre crenças e verdades.....	12
Figura 2.4.1 – Esquema do mecanismo de adaptação.....	15
Figura 4.2.1 – Os 3Vs do Big Data.....	18
Figura 4.2.2 – Panorama do Big Data.....	21
Figura 4.4.1 – Google Dengue Trends.....	24
Figura 4.4.2 – Google Dengue Trends – Brasil.....	24

# Capítulo 1 Introdução

## 1.1 – Tema

Todos reconhecemos a forte influência que as mídias digitais exercem nas relações humanas e comerciais. Após uma década de Internet ganhamos um novo eco sistema de dados que inaugura a terceira era dos computadores: a era da computação cognitiva.

Após a consolidação do papel das mídias digitais como meio de transporte de um grande fluxo de dados, vemos agora a cognitividade agregar valor nesse novo cenário tecnológico. Neste novo cenário os dados que viajam na Internet são a principal razão do Big Data. Combustível elementar para a Computação Cognitiva.

## 1.2 – Delimitação

Não é escopo desse estudo desenvolver qualquer tipo de experimento sócio comportamental ou mesmo a construção de qualquer sistema de computador, ou mesmo indicadores que descrevam ou avaliem indivíduos que tenham sofrido exposição às mídias digitais, computação cognitiva ou mesmo ações de Big Data. O estudo será limitado a análises das informações pesquisadas conforme descrito na seção 1.5.

## 1.3 – Justificativa

A sociedade digital já está incorporada à sociedade de carne e osso. O mundo real e virtual se inter-relacionam e criam dependências sociais. Nesse novo cenário de hiperconectividade o indivíduo ganha mais informação e com isso mais poder de decisão. Naturalmente, a capacidade de cognição dos indivíduos aumenta. Entretanto não conseguimos dar conta de tanta informação. É imperativo que os indivíduos recorram a tecnologia, do contrário não conseguiremos analisar tanta informação, a todo momento trazidas pelas mídias digitais.

A importância desse estudo se baseia nos esforços tecnológicos, que tem se valido da cognitividade, para desenvolverem algoritmos capazes de simular o modelo de raciocínio humano na tomada de decisão. Tudo isso para que nós possamos nos beneficiar não apenas de análises mais eficientes mas até mesmo salvar vidas.

## 1.4 – Objetivos

De forma geral, o objetivo deste estudo é expor como as mídias digitais estão sendo utilizadas para trazer cognitividade para a tecnologia.

Entretanto, para facilitar o entendimento do leitor e com base nas pesquisas sobre o Construtivismo<sup>1</sup> de Piaget, explicaremos como se dá o processo de formação do conhecimento no cérebro. Em seguida abordaremos o conceito das Mídias Digitais e o seu papel dentro desse novo eco sistema de dados. Apresentaremos o Big Data como um novo conceito para a gestão desses tipos de dados transportados pela Internet.

## 1.5 – Metodologia

Utilizamos a metodologia de pesquisa bibliográfica, onde foram desenvolvidas pesquisas a partir de materiais publicadas em livros, artigos, dissertações e teses. Foram pesquisados tanto artigos homologados em instituições de ensino superior nacional e internacional quanto textos armazenados na Internet em sites especializados, e que pudessem ter suas bibliografias comprovadas. Pesquisamos sobre cognitividade para entender seu significado tanto epistemológico quanto prático. Em seguida, colhemos informações para entender as práticas de computação cognitiva que vem ganhando espaço no mercado.

## 1.6 – Descrição

No capítulo 1 formalizamos os termos no qual este estudo foi desenvolvido. Introduzindo o leitor nos objetivos de nossos estudos.

---

**1 Construtivismo** - “Construtivismo significa isto: a ideia de que nada, a rigor, está pronto, acabado, e de que, especificamente, o conhecimento não é dado, em nenhuma instância, como algo terminado. Ele se constitui pela interação do indivíduo com o meio físico e social, com o simbolismo humano, com o mundo das relações sociais; e se constitui por força de sua ação e não por nenhuma dotação prévia, na bagagem hereditária ou no meio, de tal modo que podemos afirmar que antes da ação não há psiquismo nem consciência e, muito menos, pensamento.”. “Entendemos que construtivismo na Educação poderá ser a forma teórica ampla que reúna as várias tendências atuais do pensamento educacional. Tendências que têm em comum a insatisfação com um sistema educacional que teima (ideologia) em continuar essa forma particular de transmissão que é a Escola, que consiste em fazer repetir, recitar, aprender, ensinar o que já está pronto, em vez de fazer agir, operar, criar, construir a partir da realidade vivida por alunos e professores, isto é, pela sociedade – a próxima e, aos poucos, as distantes. A Educação deve ser um processo de construção de conhecimento ao qual ocorrem, em condição de complementaridade, por um lado, os alunos e professores e, por outro, os problemas sociais atuais e o conhecimento já construído (‘acervo cultural da Humanidade’).” “Construtivismo, segundo pensamos, é esta forma de conceber o conhecimento: sua gênese e seu desenvolvimento – e, por consequência, um novo modo de ver o universo, a vida e o mundo das relações sociais.” **Publicação:** Série Idéias n. 20. São Paulo: FDE, 1994. **Páginas:** 87 a 93. **Link:** <[http://www.crmariocovas.sp.gov.br/dea\\_a.php?t=011](http://www.crmariocovas.sp.gov.br/dea_a.php?t=011)>. **Acessado em:** 06 de agosto de 2015

No capítulo 2 veremos o conceito de cognição e a formação do conhecimento nos indivíduos. Explicaremos como o cérebro processa e cria o aprendizado. Abordaremos no capítulo 3 o contexto histórico das mídias digitais e a questão da cognitividade. Mais adiante, no capítulo 4, definiremos o Big Data. Esclareceremos seu conceito, utilidade e também suas diferenças com o BI. O capítulo finaliza com análises preditivas de sentimento nas redes sociais e a solução de computação cognitiva da IBM que tem transformado o conhecimento em sabedoria. Por fim, concluiremos este estudo no capítulo 5, juntamente aspirações para o desenvolvimento de trabalhos futuros.

# Capítulo 2 A Formação do Conhecimento

## 2.1 – Contexto Histórico

Para entender a era da computação cognitiva precisamos saber com se dá a formação do conhecimento.

Foi em resposta ao Behaviorismo, que deixava de fora os processos de aquisição do conhecimento, que os estudos do cognitivismo tiveram sua origem no Behaviorismo Radical, iniciados por Burrhus Frederic Skinner (SKINNER, 1974) [6]. “*A maneira como o comportamento se organiza tem origem nas relações [sociais]*”, é o que defende Skinner em seu livro “The Operational Analysis of Psychological Terms”, publicado em 1945. Skinner fala sobre um condicionamento operante, onde observa que as consequências de um comportamento podem influenciar na probabilidade de que ele venha a ocorrer novamente: uma consequência agradável promove o aumento da frequência deste comportamento (reforço), ao passo que um comportamento não agradável registrará uma redução desse comportamento (punição ou extinção). Constitui portanto uma abordagem teórica, voltada para o entendimento da mente, que ganhou destaque nos anos 1950 (LILIENFELD et al, 2010) [8].

Essa abordagem estuda os mecanismos que leva o indivíduo a elaborar o conhecimento e por conseguinte interfere diretamente em seus mecanismos de tomada de decisão. Esta abordagem entende a aprendizagem como sendo algo mais que apenas um produto do ambiente, das pessoas ou de fatores externos. Também leva em consideração as formas pelas quais as pessoas se relacionam com os estímulos do ambiente (real ou virtual), organizam as informações que percebem, sentem e resolvem problemas, adquirem conceitos (criação de esquemas) e empregam símbolos verbais<sup>2</sup>(semiótica).

---

<sup>2</sup>**Semiótica** é a ciência que estuda todos os fenômenos culturais como se fossem sistemas de significação. Estuda os fenômenos de significação que as coisas ganham no cérebro. Investigando todas as linguagens possíveis ela examina os modos de constituição de todo e qualquer fenômeno de produção de significação e sentido. Ela não observa o verbal linguístico, mas sim o extralinguístico, como a música, o teatro, a poesia, o cinema, e ainda, as propagandas das revistas, os jornais e tantos textos que o homem produz.

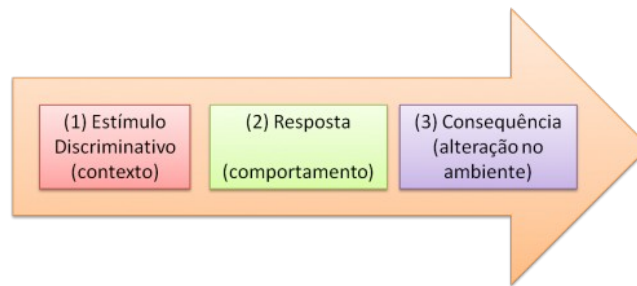


Figura 2.1.1 – Contingência Tríplice de Skinner

Disponível em: [http://www.walden4.com.br/pww4/index.php?title=Todorov\\_2012:O\\_conceito\\_de\\_conting%C3%Aancia\\_tr%C3%ADplice\\_na\\_an%C3%Aalise\\_do\\_comportamento\\_humano\\_-\\_vc](http://www.walden4.com.br/pww4/index.php?title=Todorov_2012:O_conceito_de_conting%C3%Aancia_tr%C3%ADplice_na_an%C3%Aalise_do_comportamento_humano_-_vc)

Acesso em: 10 de agosto de 2015

Essa é a base da teoria que Skinner chamou de *Contingência Tríplice*, onde um determinado comportamento (*resposta*), dentro de um contexto (*estímulo*) gera uma consequência (alteração no ambiente). Conforme explica Márcio Borges Moreira, pesquisador do Instituto Walden4, no vídeo “Contingência Tríplice” disponível no Youtube [7].

Essa foi a principal contribuição de Skinner para a psicologia, onde ele descreve um tipo de relação entre as respostas dos indivíduos (que seriam seu rastro na Web) e o ambiente que o indivíduo se insere ou percebe (as redes sociais, Blogs, as mídias digitais em geral).

Isso posto, temos o Behaviorismo Radical como ponto de partida para compreender o comportamento humano como uma interação entre estímulos do ambiente e respostas do indivíduos.

## 2.2 – Cognitivismo

Vimos que para os behavioristas o pensamento é apenas consequência de um comportamento (*contingência tríplice*). Já para os cognitivistas, a forma como as pessoas pensam impactam diretamente sobre seus comportamentos. Logo, “*o modo de pensar não pode ser um comportamento em si*”(LIDIENSE et al, 2010, pp. 24–28) [8].

O cognitivismo é a área da psicologia – hoje em dia muito voltada para a Neurociência – que se dedica a entender como o ser humano conhece, percebe e entende o mundo. Embora o assunto tenha nascido no campo da psicologia, ele também está diretamente ligado com os estudos do Processamento da Informação. E que teve como teórico principal o estudioso John Von Neumann, responsável pela criação do Modelo de Processamento de Informação para computadores (1953).

A doutora em psicologia Maria Alice d'Avila (BECKER,1997) observa que os estudos do cognitivismo levam em consideração métodos, para descrevem as funções mentais, como os modelos de processamento de informação. Incorporando ainda elementos da teoria dos sistemas, das ciências da computação, da cibernética, da teoria da informação e da robótica.

O desenvolvimento do cognitivismo foi alavancado, principalmente, pelos estudos relacionados com os avanços tecnológicos. E que muito auxiliaram os cognitivistas através de métricas nas análises a respeito de como as pessoas solucionavam problemas difíceis (tarefas mentais). Permitindo assim que eles construíssem modelos para ilustrar suas conclusões. Modelos esses muitas vezes implementados com programas de computador ou fluxos de processamento cognitivo, na observação da realização de tarefas (BECKER,1997)[4].

### **2.3 – Cognição**

Através da Epistemologia, ramo da filosofia que se relaciona com o conhecimento humano e suas relações entre o sujeito e o objeto do conhecimento [5], sabemos que o conhecimento do homem está dividido em três grupos: “*saber que*”, “*saber como*” e o da “*familiaridade*”.

Ao ler um livro de regras um indivíduo descobre que  $1+1=2$  (*saber que*). Contudo, ele só consegue aprender como se anda de bicicleta se ele praticar. Desenvolvendo o conceito do equilíbrio, *saber como* ou conhecimento prático. O que não dá para conhecer apenas lendo livros! E por fim, você pode apenas conhecer uma pessoa e ainda assim não estar *familiarizado* com ela. Dessa forma obtemos dois grandes grupos: Crenças e Verdades.

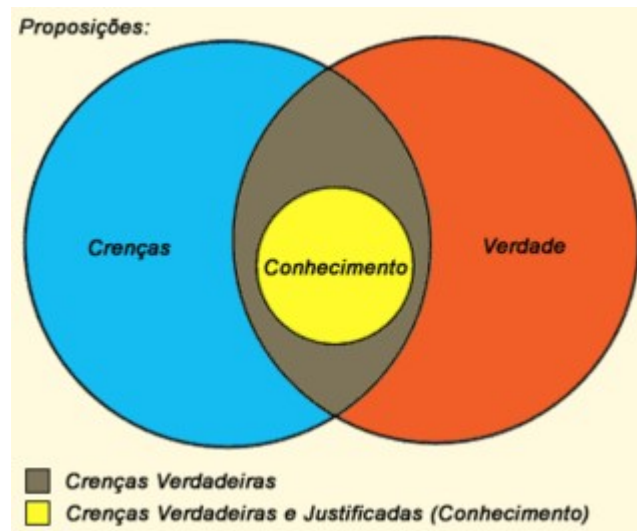


Figura 2.3.1 – O conhecimento entre crenças e verdades

Fonte: <https://pt.wikipedia.org/wiki/Epistemologia>

Acessado em 10 de agosto de 2015

As Crenças representam um discurso verdadeiro, é algo na qual o indivíduo acredita (ou tem confiança). Já o grupo das Verdades é tudo o que já é conhecido. Por exemplo: “... se uma pessoa acredita que a ponte é segura o suficiente para aguentar seu peso e tenta atravessá-la, mas a ponte se quebra devido ao peso, pode-se dizer que a pessoa acreditou que a ponte era segura, mas estava errada. Não seria correto afirmar que ele sabia que a ponte era segura, pois ela, claramente, não era. Em contraste, se a ponte aguentasse seu peso, ela diria que acreditou que a ponte era segura e, agora que cruzou a ponte e provou para si que a ponte é segura, ela sabe que é segura.” [5].

Agora que entendemos como se dá a construção do conhecimento, podemos entender o que é cognição. A cognição é o processo de “criação” do conhecimento que envolve: atenção, percepção, memória, raciocínio, juízo, imaginação, pensamento, linguagem e ação (os cinco sentidos). Fazem parte desse fluxo de criação, etapas como: armazenar; reconhecer; compreender; organizar e utilizar/distribuir a informação (estímulo ou contexto) que se recebe através dos cinco sentidos (informação sensorial). É a partir dos processos de formação do conhecimento que os indivíduos podem desenvolver seus mecanismos de tomada de decisão, fixando objetivos e minimizando suas perdas (consequências negativas), ou seja, adaptando-se ao meio.

Devemos casar ou comprar uma bicicleta? Fazer escolhas é um aspecto fundamental do indivíduo. Mas quais são os processos psicológicos envolvidos na tomada de decisão? Em que são baseados? Durante muito tempo existia a crença de que



os indivíduos tomavam sempre a melhor decisão possível. Esse pensamento se deve ao fato de sabermos que os seres humanos, ao contrário dos animais, são seres racionais (Von Neuman e Morgensten, 1943).

Operar escolhas significa eliminar conflitos. Você tem que escolher entre duas universidades: *A universidade uA tem um excelente corpo docente e uma péssima infraestrutura. Já a uB tem uma excelente infraestrutura, mas seus professores não são lá essas coisas.* Esse tipo de situação pode lhe trazer o Conflito Cognitivo. Estudos observaram que essa questão é comumente resolvida com uma avaliação do risco que cada escolha oferece. Foi observado que quanto maior o risco que uma escolha apresenta, menor são as preferências por uma escolha (Kanheman & Tversky, 1979; Mazur, 1987; Rachlin, Raineri & Cross, 1991) [9].

A chave da questão então é entender como funciona o modelo que classifica o risco de cada escolha. Essa é a missão do mecanismo de cognição dos indivíduos. Assim como, a dos computadores cognitivos. Lembrando ainda que, visões de cada indivíduo podem implicar em escolhas distintas.

E foi Jean Piaget, um dos mais importantes pesquisadores na pedagogia, que propôs uma teoria construtivista, a respeito de como melhorarmos nosso mecanismo de tomada de decisão. Seus estudos sobre pedagogia revolucionaram a educação, derrubando várias teorias tradicionais relacionadas ao aprendizado. Também teve certa relevância na ciência da computação, onde podemos citar Seymour Papert [14], que usou o trabalho de Piaget para desenvolver a base da linguagem de programação Logo [13], muito utilizada como ferramenta de apoio ao ensino regular no desenvolvimento de crianças, jovens e até adultos.

## **2.4 – Construtivismo**

Para o construtivismo o desenvolvimento do conhecimento é criado pelo próprio indivíduo através de sua interação com o meio (PIAGET,1982) [15]. Nesse contexto, a criação do conhecimento se dá quando o indivíduo age (física ou mentalmente) sobre os eventos observados, provocando o desequilíbrio do conhecimento adquirido anteriormente. Ou seja, somente quando o indivíduo confronta o que ele já sabe é que ele aprende.

Piaget não acreditava que o indivíduo já nascesse com todo o conhecimento (apriorismo), tão pouco que todo o conhecimento fosse oriundo do meio (empirismo).

Em sua teoria ele defendia que o conhecimento é gerado através da interação do sujeito com o meio externo a partir de esquemas preexistentes no indivíduo. Assim, a criação do conhecimento depende tanto de estruturas cognitivas inerentes no indivíduo como de sua relação com o fato observado. Essa relação se dá através do processo de *adaptação*, que é subdividido em dois momentos: a *assimilação* e a *acomodação*.

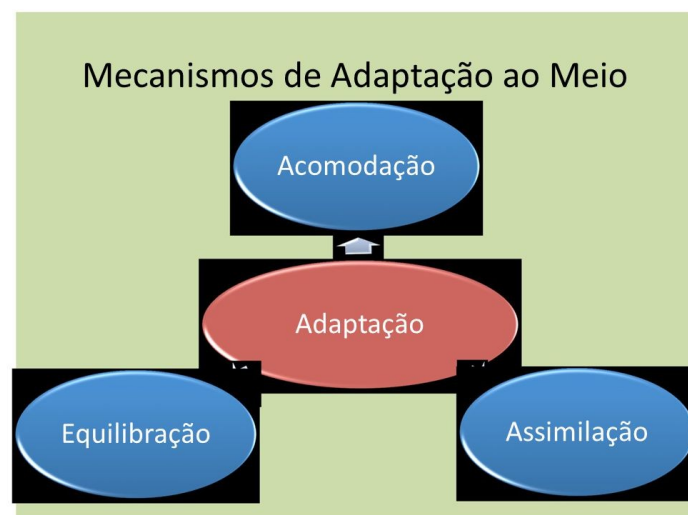


Figura 2.4.1 – Esquema do mecanismo de adaptação  
Disponível em: Slide 17 em <http://slideplayer.com.br/slide/3231723/>  
Acesso em: 10 de agosto de 2015

Na Assimilação o indivíduo toma ações que internalizam o fato observado, classificando novos eventos em esquemas preexistentes. É o processo pelo qual o indivíduo capta e organiza as informações, permitindo a ampliação de seus esquemas a partir das estruturas que já possui. Se o indivíduo não conseguir encaixar esse novo evento (fato observado) em seus esquemas preexistentes ocorre então o desequilíbrio.

E na Acomodação o indivíduo altera suas estruturas cognitivas para melhorar a compreensão do fato observado, ou criando novos esquemas ou alterando os preexistentes. E nas consecutivas relações entre assimilação e acomodação o indivíduo vai se adaptando ao meio externo. Assim, entre a *assimilação* e a *acomodação*, o indivíduo restabelece o equilíbrio.

Para Piaget, e conforme a corrente *construtivista*, o processo de desenvolvimento cognitivo é interminável. Pois os vários níveis de conhecimento da mente humana estão indefinidamente em construção, através das interações consecutivas entre o indivíduo e o meio (PIAGET, 1996) [17].

Fazendo um paralelo com a sociedade digital, o Internauta seria o indivíduo e a Internet o meio externo. E se o conhecimento é criado através das várias interação com o meio, desequilibrando o entendimento do conhecimento adquirido anteriormente, podemos fazer aqui uma relação com o rastreamento das atividades (*várias interação*) dos Internautas e os procedimentos de Big Data, na tentativa de prever o comportamento do usuário (equilibração).

# Capítulo 3 Mídias Digitais

## 3.1 – Contexto Histórico

Durante o século XIX, a sociedade moderna conseguiu grandes avanços, mas que só foram possíveis graças ao aumento do número de pessoas alfabetizadas. Essas pessoas passaram a ter mais participação na sociedade, consumindo jornais e escrevendo cartas. Deu-se então um aumento na transmissão da informação. Com a Internet, novas formas de comunicação surgem a todo momento. Surge um novo espaço virtual e democrático, no qual os indivíduos ganham instrumentos para agilizarem e expandirem o conhecimento, *“...ocorreu a ampliação dos espaços democráticos em que o surgimento das novas mídias evidencia a experiência do século 21, caracterizando, dessa forma, a sociedade da informação”* (NUNES, 2013, p.55). As chamadas Mídias Digitais são o pano de fundo desse novo espaço virtual onde o indivíduo pode se renovar, a qualquer momento, e atuar de outra forma diferente da que ele tinha conhecimento pelos seus antepassados.

Segundo o sociólogo Pierre Lévy (1998) o ciberespaço é a nova forma de democracia, mas que só pode ser exercida através da apropriação das novas tecnologias. O ciberespaço seria o lugar chave onde profundas mudanças sociais, econômicas e políticas se personificariam em um novo modelo de sociedade sem vínculos territoriais, relações institucionais ou laços de poder. Para Lévy o ciberespaço seria um “local” *“para o indivíduo exercer a comunicação aberta, através de uma interconexão mundial dos computadores e das memórias dos computadores”* (LEVY, 1999, p. 92).

Nasce a sociedade informacional, e que por sua vez remodelou o capitalismo e seus conceitos de produção, onde a comunicação e a informação são as principais ferramentas do novo século. A sociedade da informação, *“é uma expressão recorrente nas análises recentes sobre a sociedade. Utilizada em diferentes contextos teóricos, designa geralmente um ambiente de relações sociais mediados por sistemas de informação, cujo aporte tecnológico são as redes de computadores”* (PRUDÊNCIO, 2006, p.17).

Encontraremos muitas definições para a palavra *Mídia*. Em comunicação ela serve fundamentalmente para referenciar *canais*, ou *ferramentas*, que servem para

transmitir informação. Ou seja, ela é utilizada para se referir a um meio, utilizado para comunicar informação para qualquer finalidade.

Segundo o Prof. Lucas Pelaez (PELAEZ, 2012) [2] *"... novas mídias é toda forma de comunicação que encontramos fora dos meios tradicionais: rádio, jornal, televisão, revista."*. Seguindo essa linha, o termo carrega um amplo contexto, um conjunto de novas tecnologias e métodos de comunicação, e que por isso se diferencia dos canais mais tradicionais de comunicação. Sendo este grupo identificado como Internet, Celular e IPTV[3]. Para nós, no entanto, entendemos que tanto a tecnologia quanto a sociedade – em suas inúmeras formas de interação – ainda estão longe de definirem todas as formas.

*"O maior impacto da Internet não foi por [ela] ter surgido como uma nova mídia e sim por ter mudado o comportamento do consumidor e essa mudança se deu principalmente pelo surgimento das comunidades virtuais e intensificação do internauta como ator principal"* (Sandra Turchi) [2]

Apenas a introdução de uma nova tecnologia na sociedade não é fator determinante para promover o aumento da produtividade (FRIEDMAN, 2009)[1]. Sabemos que a lâmpada foi inventada em 1879 e que passaram-se décadas até que a eletricidade pudesse exercer papel decisivo na economia, sendo hoje um dos fatores preponderantes na produção. De certa forma é o que vemos hoje, com o que Thomas Friedman chamou de o *"achatamento do mundo"* [1]. Muitas coisas já vêm ocorrendo há bastante tempo, mas seus efeitos só podem ser sentidos com o surgimento de uma *"massa crítica"* (FRIEDMAN, 2009). Ou seja, um conjunto de coisas (pessoas, tecnologias,...) capazes de desenvolver a colaboração necessária para ocorrer, por exemplo, o Big Data. É o que foi chamado de *"efeito sócio psico midiático"*, cunhado pela mídia e pelo psicólogo Bernard Luskin, autor de *'The Media Psychology Effect'*, onde ele relaciona às implicações socioculturais dos meios de comunicação no comportamento humano.

No mundo virtual, ou mundo mediatizado, terá mais chances de sucesso quem conseguir lidar com os novos hábitos, processos e competências disponíveis a todo o momento nas mídias digitais. A inclusão do indivíduo nessa sociedade digital possibilita ao próprio indivíduo não apenas receber, mas também transmitir e construir conhecimento. O que afeta diretamente a sua cognição. E é aí o foco dos novos sistemas

de computador, que já estão sendo programados para construir a sua própria base de conhecimento, como veremos adiante.

### **3.2 – A Cognitividade nas Mídias Digital**

A velocidade com que as mídias digitais se comportam no cotidiano reduziu drasticamente a noção de tempo e espaço que existia antes da Internet. Porém, a capacidade de reflexão do indivíduo não evoluiu da mesma forma que sua capacidade de percepção (cognição).

Com a variedade de canais e plataformas de comunicação disponíveis o indivíduo passou a ser bombardeado com vários estímulos, e que muitas vezes não são processados mas sim descartados. Para que esse novo indivíduo consiga se inserir efetivamente nas novas formas de comunicação da sociedade atual é necessário que ele desenvolva habilidades de análise e síntese ao mesmo tempo. Durante esse processo de aprendizado o indivíduo é forçado a fazer questionamentos, conectar ideias (esquemas de Piaget), criar relações e elaborar cenários sobre a sua visão da realidade (semiótica).

As mídias digitais transportam milhares de informações nessa nova sociedade digital. Para que fosse possível maximizar o aproveitamento dessas informações, empresas da área de tecnologia como a IBM não tem poupado esforços científicos para cruzar conhecimento das mais diversas áreas, como por exemplo, congregando a neurociência com computação distribuída.

# Capítulo 4 Big Data

## 4.1 – Definição

E parece que não é apenas a psicologia que quer entender o mecanismo de decisão dos indivíduos. Por todos os lados estão se falando sobre o dilúvio de dados, e que já vem sendo coletados por várias empresas e de vários níveis de atividade econômica. Desde financeira e vendas de varejo como também hospitalar. Promovido pelo Marketing, os “dados” agora são o grande negócio. Porém, só vale o esforço se você tiver condições de analisar essa massa de dados e saber como aplicar sua informação de maneira que possa trazer resultados para o negócio. Mesmo que seja apenas contribuições significativas. Como veremos mais a frente no caso do Watson.

Big Data é um termo que, no seu início, veio sendo muito utilizado para referenciar grande volume de dados ou dados muito complexos (desestruturados). E também servia para sinalizar que não faria sentido utilizar as aplicações atuais, que realizam processamento de dados, para analisá-los. Seus principais desafios envolvem: análise, captura, adequação, seleção, compartilhamento, transferência, visualização e informações particulares. A função de uma equipe de Big Data é realizar análises preditivas, na tentativa de extrair algum valor dos dados coletados para o negócio [21].

John Weathington, CEO da Excellent Management Systems (<http://www.xmsystems.com>), uma empresa de consultoria estratégica utilizada por importantes corporações multinacionais, diz que *“para definir Big Data em termos competitivos, você deve pensar sobre o que é preciso para competir no mundo dos negócios. Big data é tradicionalmente caracterizado como um rio que flui grandes quantidades de dados em um ritmo rápido. Para ser competitivo com os clientes Big Data deve criar produtos que são valiosos e únicos. Para ser competitivo com os fornecedores Big Data deve estar disponível livremente, sem obrigações ou restrições.”*

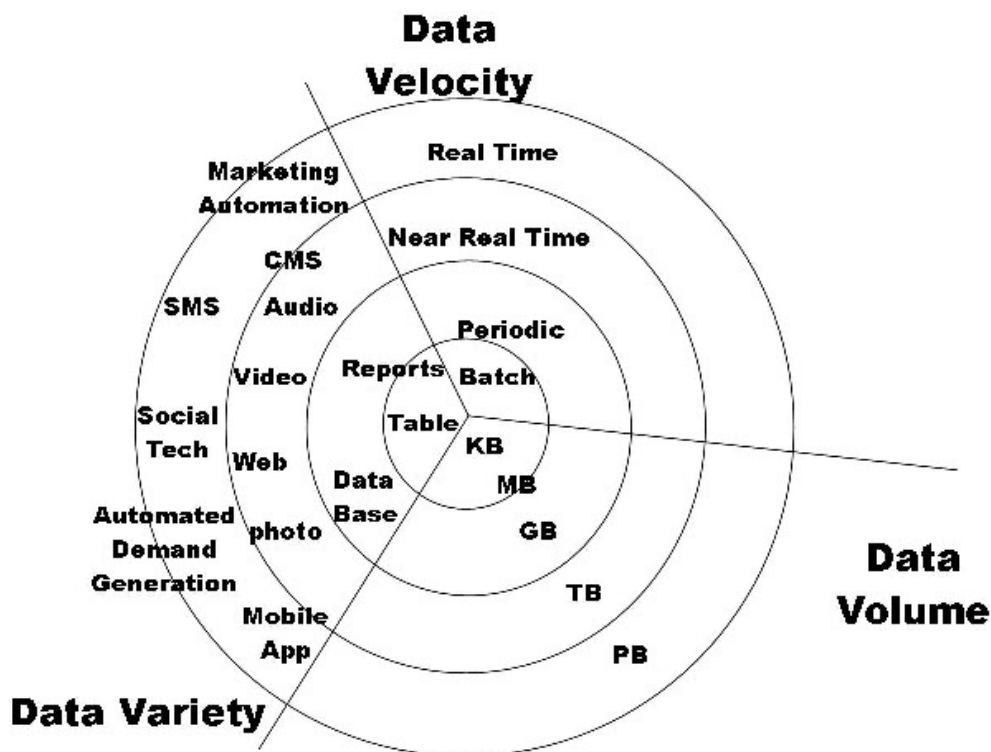
Boyd & Crawford (2012) em seu artigo “Critical Questios for Big Data” [22] observam o Big Data como um fenômeno cultural, tecnológico e acadêmico, onde nos chamam a atenção para três pilares: 1) Tecnologia: O Big data maximiza o poder computacional e precisão algorítmica para coletar, analisar, relacionar, e comparar grandes conjuntos de dados; 2) Análise: Se baseia em grandes conjuntos de dados para identificar padrões, a fim de fazer previsões econômicas, sociais, técnicos e legais; 3) Mitologia: Eles observam uma crença generalizada de que grandes conjuntos de dados

oferecem uma forma mais elevada de inteligência e conhecimento que podem gerar *insights* [mais preciso], e que antes eram impossíveis, com a aura de verdade, objetividade e precisão. Também apontam para a questão da falta de privacidade e da acurácia dos sistemas responsáveis em interpretar os padrões de comportamento das pessoas. Sem que incorram riscos para os cidadãos de serem mal interpretados e vistos como ameaças para a sociedade [22]. Boyd & Crawford, ainda em seu artigo, convocam a sociedade para essa reflexão, principalmente sobre a regulamentação dessa prática na Web.

## **4.2 – Caraterísticas**

Em janeiro de 2012 [23], Doug Laney, membro do Gartner, faz um alerta para as dimensões de valor do Big Data como sendo importantes ativos de informação onde devem ser considerados o volume, a velocidade e a variedade. E que eles necessitam de formas inovadoras para o processamento dessas informações. A fim de que esses dados possam trazer mais informações de maior visibilidade, auxiliando na tomada de decisão e automação de processos. Neste artigo, ele ainda faz referência a um outro artigo seu, publicado em fevereiro de 2001 intitulado “*3-D Data Management: Controlling Data Volume, Velocity and Variety*”[24]. Com os estudos de análise de dados em tempo real em franca expansão, tem se discutido muito a questão dos 3Vs (Figura 4.2.1).





Disponível em: <http://velvetchainsaw.com/2012/07/20/three-vs-of-big-data-as-applied-conferences/>  
 Acesso em: 11 de agosto de 2015

## 1) Variedade

Quanto aos tipos ou forma, englobam tanto dados estruturados (i.e.: formulários) como desestruturados. A exemplo de dados desestruturados podemos citar:

- a) **Dados Transacionais:** São basicamente rastros de atividades das pessoas que navegam nos websites da Web, como: Facepost; e-mail; tweets; check-ins no Foursquare; Localização do usuário no Waze, via GPS; um clique do usuário em um link; itens de desejo ou de preferência, para ver depois em um marketplace qualquer.
- b) **Dados de Conteúdo:** São os diversos tipos de mídia como fotos, vídeos, arquivos de texto, blogs, formulários de solicitação que os websites exibem para enviar um e-book pro usuário;

## 2) Velocidade

Analisar dados em tempo real é o grande diferencial do Big Data. E é o que se tem feito valer, e muito, os conceitos da computação distribuída. O processamento de dados em tempo real ocorre de forma diferente que a dos processos de mineração de dados<sup>3</sup> conhecido atualmente (como veremos mais adiante). As novas ferramentas (figura 4.2.2) que compõe o panorama do Big Data suprem as carências das ferramentas que utilizam as atuais bases de dados relacional. Não apenas o tipo de dado, mas sobre todo a abordagem em tempo real, necessitam de técnicas aplicadas em computação distribuída: como ferramentas de banco de dados distribuídos (i.e.:Hadoop [19]) além de infraestrutura necessária (cloud) para aplicar processamento de paralelismo maciço.

## 3) Volume

A velocidade faz com que o grande volume de dados seja bastante óbvio: petabytes (um milhão de gigabytes). Um arquivo de texto é em kilobytes. Um arquivo de som é geralmente em megabytes e um vídeo em gigabytes. Os dados são gerados por empregados, clientes, potenciais clientes, fornecedores, fornecedores e parceiros. Os dados também podem ser gerados por um grupo de máquinas. Por exemplo, os dispositivos móveis enviam uma variedade de informações para a infraestrutura de rede da empresa através de sites, SMS e chamadas telefônicas. A volumetria dos dados é declaradamente exponencial e gigantesca. Existe um consenso na dificuldade para se armazenar, manter, processar e acessar esses dados;

## 4) Veracidade

Mas os 3Vs já não são mais suficientes. Surge uma quarta variável que chama a atenção dos estudiosos e que indica o quanto o dado é relevante na análise. Existe muito ruído nesses dados. Por essa razão as empresas estão investindo em profissionais como os Cientistas de Dados [31,32,33] ou Chief Data Officer [34]. Pois os dados precisam ter uma significância relativa ao negócio. Muitas das vezes os dados contém informações não tão óbvias, mas ainda podem representar uma vantagem para o negócio. Uma boa fronteira para identificar a capacidade de um dado ser classificado como Big Data é

---

<sup>3</sup>**Mineração de Dados**, mais conhecido com o Data Mining, é um processo de coleta de dados. O objetivo é parecido com a mineração de metais preciosos, onde se procurar separar o metal do detrito. Neste caso, regras e filtros serão implementados para procurar padrões consistentes de informação na massa de dados observada, como regras de associação ou sequências temporais. Formando novos subgrupos de dados. É a base dos processos de BI.

identificar a capacidade que a empresa tem para armazenar, processar e acessar esse dado e confrontar se a empresa necessita desse dados com vantagem competitiva ou servir os clientes. Bowker (2005, pp. 183-184), no mesmo artigo de Boyd & Crawford (2012) [22], diz que devemos discutir sobre as diferentes temporalidades, espacialidades e materialidades na qual podemos representar esses dados em nossos bancos de dados, pois dados crus, sem sentido formal, podem representar ao mesmo tempo uma grande contradição. Por isso os dados coletado devem ser preparados cuidadosamente. É importante alertar que é muito comum encontrarmos nos resultados de análises Big Data uma certa *apofenia*, que “*é um fenômeno cognitivo de percepção de padrões ou conexões em dados aleatórios.*” criador de boatos e mitos [25]. Motivo mais que suficiente para relacionarmos esta abordagem com os estudos da cognitividade, amplamente necessária nas metodologias complexas dos processos de limitação e interpretação dos dados. Outra demonstração de desentendimento de uma fonte de dados de Big Data está no Twitter. É um erro grave assumir que os usuários do Twitter e as “pessoas” sejam a mesma coisa. É preciso considerar os subconjuntos de usuário. Como também não devemos assumir que as contas no Twitter são equivalentes. Algumas contas são usadas por mais de um usuário e alguns usuários usam mais de uma conta. Algumas pessoas não tem uma conta no Twitter, acessando as informações abertas na Web. Outra contas são como robôs, que geram conteúdos indiscriminadamente, sem nenhum relacionamento pessoal. Enquanto que uns usuários postam informações outros apenas leem as publicações. Em 2010 o Twitter revelou que 40% dos usuários ativos se cadastraram apenas para acompanhar as publicações. Os conceitos de usuário, participante e ativo devem ser amplamente examinados! Sem que se entenda corretamente o significado da origem da informação, o dado pode ser qualquer coisa. [26]

O nome Big Data “pegou” mas foi uma escolha ruim face a tudo o que ele representa. Foi inicialmente utilizado nas diversas ciências para se referir ao grande volume de dados que os cientistas utilizavam em suas experiências. O termo Big Data não se restringe a tecnologia e nem uma ou duas ferramentas para análise de dados. Tão pouco representa somente uma forma diferenciada de fazer mineração de dados. Big Data está menos relacionado com o volume de dados e mais com a capacidade de pesquisar, agregar e de realizar referência cruzada com os grandes conjuntos de dados disponíveis.

Big Data é um novo conceito de negócio orientado aos dados. Um eco sistema de dados, que integra tanto novas ferramentas para entender os diversos tipos de dados disponíveis nas mídias digitais quanto novas abordagens e procedimento operacionais, voltada para a geração de informação útil. E que na maioria das vezes precisa ser feita em tempo real. Sem Big Data, por exemplo, não existiria o marketing voltado para o conteúdo digital de massa. Prática já consolidado no mercado como Inbound Marketing<sup>4</sup>.

No website bigdatalandscape.com, podemos ver um panorama das principais empresas que estão trabalhando para viabilizar o Big Data. Clicando em cada imagem no site visualizaremos a definição de cada uma delas.

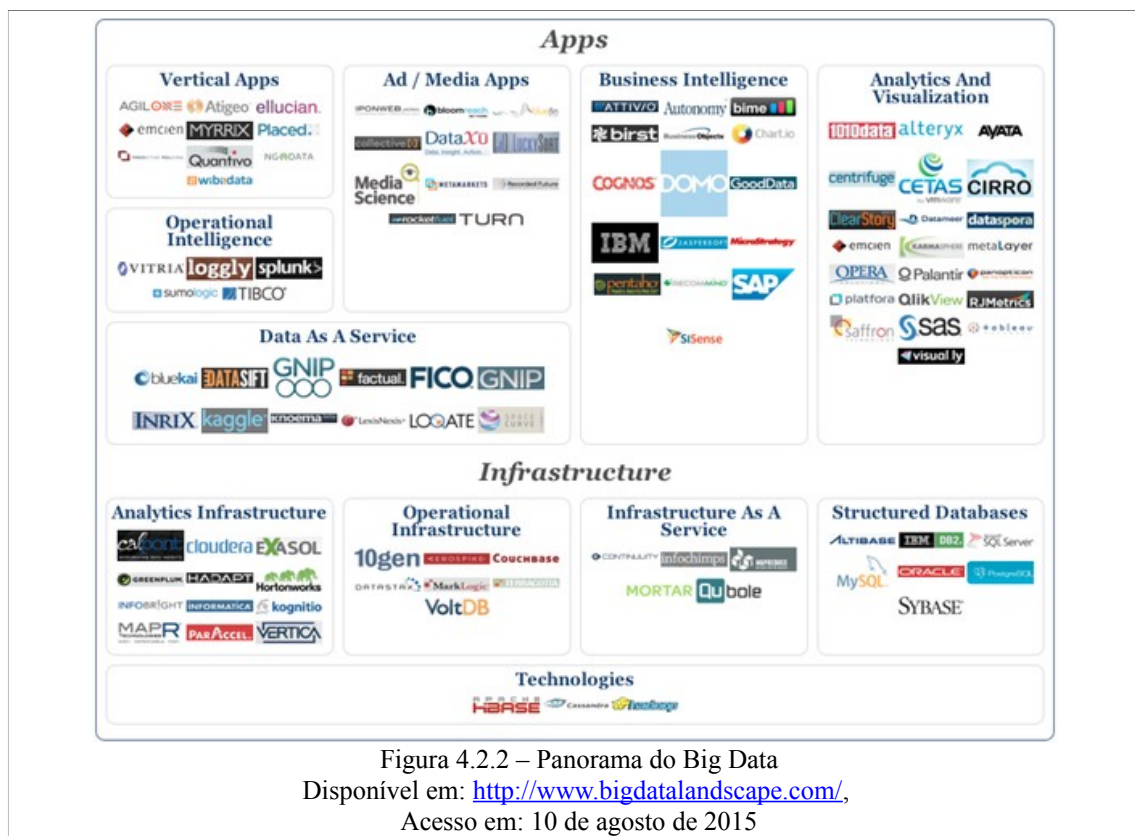


Figura 4.2.2 – Panorama do Big Data  
Disponível em: <http://www.bigdatalandscape.com/>,  
Acesso em: 10 de agosto de 2015

### 4.3 – Big Data e o Business Intelligence (BI)

A ideia por trás do Big Data é muito parecida com a do BI, analisar dados com o objetivo de torná-los estratégicos para as empresas. Porém, não bastasse o grande volume de dados, existem várias diferenças no comportamento dos dados utilizados nas

**4 Inbound Marketing** — É um conjunto de boas práticas, baseada no antigo conceito do Marketing de Conteúdo, criado por Brian Halligan e Dharmesh Shah da hubspot.com. Uma das premissas é falar sobre o que é relevante ao público e não sobre os produtos e a marca. Possui dois fundamentos principais. O primeiro é não interromper as pessoas para exibir um produto. O segundo é fazer com que as pessoas encontrem o que desejam, com o objetivo de criarem algum tipo de relação com os websites que detém determinada informação ou suprem uma certa necessidade.

abordagens de Big Data. Principalmente quanto a forma de captura dessas informações (intrusiva e em tempo real), significado e acurácia desses dados (relevância, volatilidade e qualidade de dados).

Diferente das práticas de mineração de dados utilizadas no BI, ainda não é possível valorar com 100% de certeza a utilidade das diversas fontes de dados utilizados no Big Data. Nas estratégias do Business Intelligence (BI) os dados são capturados (minerados) de tabelas dos sistemas transacionais das empresas, como o sistema de ERP. Esses dados, por sua vez, tem uma característica estática, visto que representam eventos que ocorreram no passado (i. g.: volumes de vendas no mês passado). Esses dados são migrados para uma base de dados que junta todos os dados necessários para que possam ser realizadas as práticas de BI. Essa junção dos dados ocorre para facilitar o acesso aos dados. Tecnicamente, também é possível realizar consultas complexas diretamente no banco de dados transacional, onde está armazenado todo o modelo de dados da empresa. Porém, essas consultas concorreriam com as operações realizadas no dia a dia, ocasionando lentidão em todas as aplicações que alimentam esse bando de dados. Para o leitor entender, seria o mesmo que esperar muito tempo em frente ao caixa eletrônico pelo seu dinheiro sair da máquina, pois o mesmo banco de dados que está registrando seu saque também está sendo usado para responder às consultas de equipes internas da empresa.

Logo, as perguntas que podem ser respondidas pelo BI (nas consultas) são do tipo: “Quanto foi o faturamento do mês passado?” ou “Quantas canetas vermelhas foram vendidas em janeiro do ano passado?”. Esse tipo de abordagem, com base em dados históricos acumulados é determinada conforme as probabilidades matemáticas orientadas por uma Abordagem da Frequência Relativa [20]. O BI ajuda muito a responder o que ocorreu com base no comportamento passado (dados históricos). Tal como psiquiatras corporativos em suas respectivas teorias, as equipes de marketing e planejamento vem projetando comportamentos futuros e realizando suas previsões de quanto devem crescer ou ser investido nessa ou naquela área da empresa com ajuda do BI.

Agora imagine um motor de avião. Digamos que uma peça observada em laboratório quebrou após 13 meses de uso. Suponhamos que, aplicando cálculos de probabilidade com os dados históricos coletados, se chegou a um número mágico de que essa peça deva ser trocada com 4 meses de uso, considerando principalmente a questão

da segurança dos passageiros. Mas e se essa peça, ou o procedimento que a peça é responsável, pudesse ser monitorado em tempo integral? Como nos motores Rolls-Royce AE3007 [29] utilizados nos Jatos da Família ERJ da Embraer.

Nesse caso, mesmo os sensores do motor enviando milhares de dados por segundo, não é nada se comparado ao volume de dados que flutua nas Internet. Mas a coleta de dados, em tempo de real, é o que difere o Big Data do BI. O BI não morreu, ele está dentro do Big Data. Mas para conseguirmos coletar e analisar esse enorme volume de dados que flutua na Web, foi preciso que a TI criasse ferramentas que pudessem processar esses dados de uma outra forma.

Para isso foi preciso a mudanças de paradigmas, como por exemplo a forma de se consultar os dados em tempo real. Todo mecanismo padrão de consulta de dados funciona assim: enviamos uma consulta (i.e.: Qual o valor total das vendas hoje?) para o banco de dados e ele nos responde o valor das vendas referente até o momento que a consulta foi solicitada. O banco de dados percorre todos os registros da Tabela de Pedidos e soma o valor de todos os pedidos com a data de hoje.

Com a quebra desse paradigma, essas consultas ocorrem de outra forma. Taurion (TAURION, 2013) [40] observa que, em vez de se enviar a consulta para uma ferramenta (banco de dados) responder, as novas ferramentas disponíveis (como o Hadoop [19]), que realizam a gestão distribuída de dados, instalam essas consultas no fluxo de dados que estão passando por tal ferramenta, para que essas informações possam ser coletadas em tempo real (*stream computing*) (TAURION, 2013).

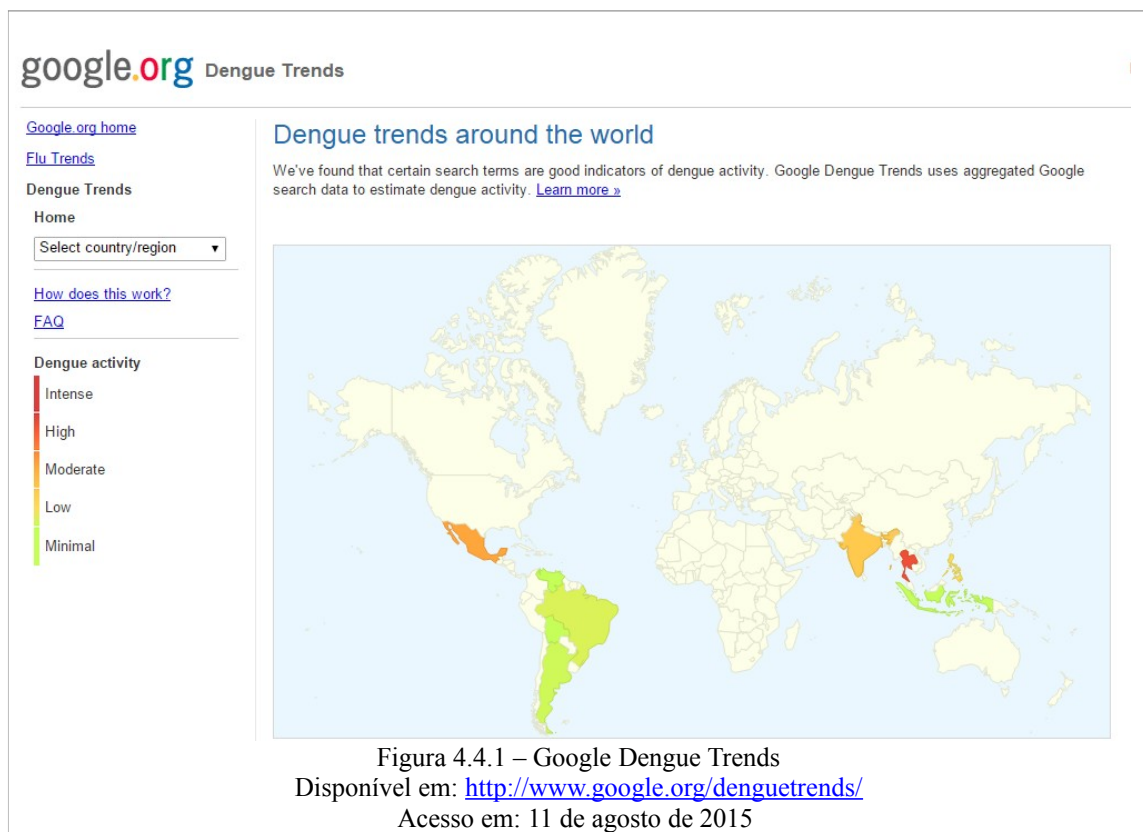
Imagine um rio, onde passam peixes brancos, azuis e vermelhos. Imagine agora uma pessoa contando quantos peixes vermelhos passaram por ela. E outra pessoa contando quantos peixes brancos passaram e qual o tamanho médio deles. Agora transporte essa metáfora para a realidade das mídias digitais, o quanto de informação que passa a cada segundo, as publicações no Facebook!

É por isso que quando se fala em Big Data, além do grande volume, variedade e veracidade dos dados, também se preocupam muito com a velocidade com eles aparecem. Fator fundamental para se determinar o *valor da informação*. A final o que pode ser útil nesse segundo, pode já não ter mais valor uma hora depois! A questão do tempo da informação (volatilidade) foi o divisor de águas do mundo do BI.

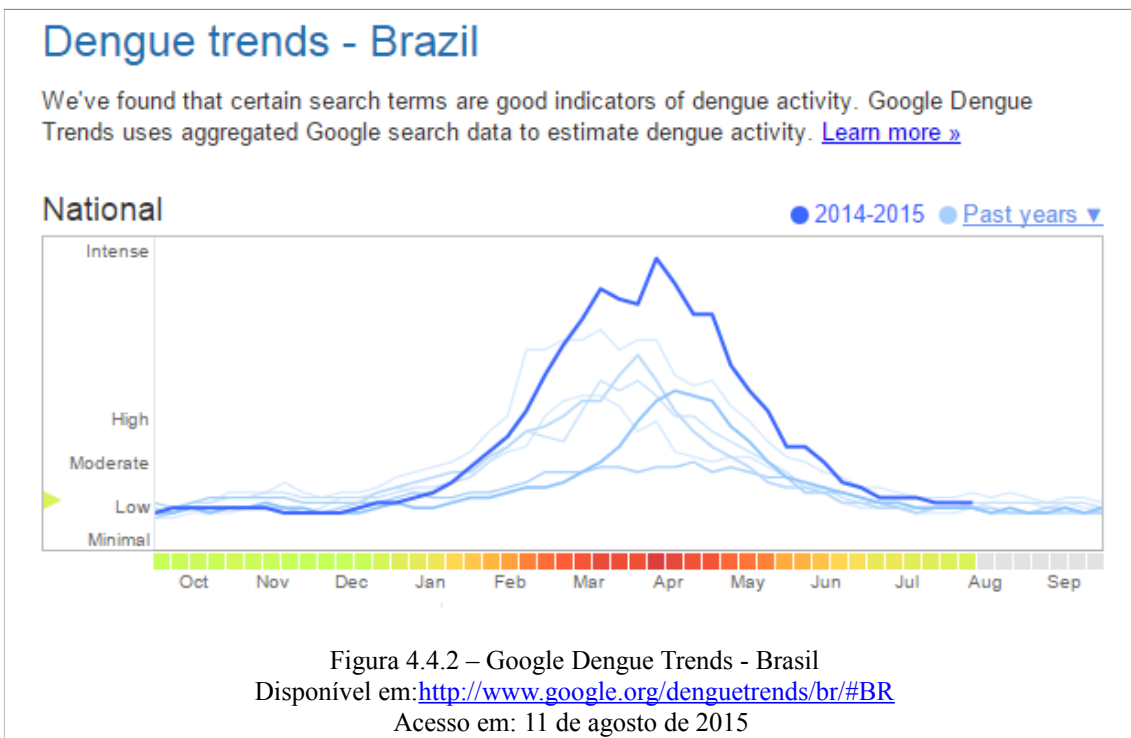
## 4.4 – Análises Preditivas

As análises de Big Data servem para criar modelos preditivos, ou seja, prever padrões de comportamento de eventos que ainda não ocorreram, conforme os conceitos de probabilidade de Abordagem Subjetiva, que são baseados no julgamento pessoal, acúmulo de conhecimento e experiência [20].

Foi o que fez o Google, no projeto Dengue Trends [30], que disponibiliza no site <http://www.google.org/denguetrends> análises em tempo real, tanto de informações que estão sendo pesquisadas no seu próprio site com em outras informações fontes de informação históricas relevantes, como bancos de dados de hospitais locais (BI).



E ao selecionar o Brasil, veremos a distribuição projetada para casos de Dengue no país.



## 4.5 – O Sentimento nas Redes Sociais

Um dos principais diferenciais das análises de dados Big Data é que podemos saber das coisas em tempo real. Podemos fazer uma predição, por exemplo, a respeito de qual candidato de uma eleição se saiu melhor no último debate apenas analisando as milhares de publicações do Twitter e Facebook.

Já sabemos que a maior vantagem competitiva do Big Data está nas análises das mídias digitais em tempo real, como na análise de sentimento feita pela IBM em junho de 2013 no “Projeto Ei!” para a Copa das Confederações, utilizando sua solução de Big Data: Watson [35, 36]. Análise de sentimentos social é uma solução que avalia o que as pessoas estão “falando” nas redes sociais. O objetivo é apontar tendências relacionadas a um determinado assunto. Esse tipo de análise pode revolucionar toda a indústria de marketing. No Projeto Ei, o objetivo era trazer o sentimento dos torcedores para perto do treinador da Seleção Brasileira. A solução da IBM permitiu obter, em tempo real, opiniões sobre os jogadores, táticas e expectativas da torcida [35]. Após “aprender” com um dicionário específico sobre futebol, que permitiu que a solução da IBM identificasse se o tweet era sobre futebol ou não, a solução monitorou tudo que era postado no Twitter sobre o tema “futebol”. Uma vez que o tweet tivesse aderência ao tema ele era analisado em 5 etapas:



1. As palavras que compõem cada tweet foram separadas umas das outras em um processo conhecido como parser;
2. Em seguida, as palavras foram normalizadas, ou seja, erros foram corrigidos e, eventualmente, sinônimos foram empregados;
3. Depois cada palavra foi categorizada de acordo com as regras da gramática portuguesa, identificados adjetivos, substantivos, verbos, etc
4. A seguir, foi identificado o lema de cada verbo;
5. Para finalizar, o sentimento de cada palavra foi retornado com: positivo, negativo ou neutro.

O sentimento referente a cada palavra foi previamente classificado (semiótica). Com todas as palavras previamente classifica, partiu-se para o cálculo do sentimento de todo tweet. Por fim, foi realizada uma etapa de análise estatística para identificar as frequências com que os nomes dos jogadores foram mencionados, os temas mais frequentes, etc.... Contando assim o processo parece bem simples, mas classificar o sentimento das palavras foi uma etapa bastante complexa [35], uma vez que as palavras podem ter um sentido diferente dependendo da forma como são usadas. Para exemplificar, analisando o vocábulo do futebol, eles entenderam que a palavra "vamos" é normalmente utilizada no futebol como "vamos Brasil", e que teria um significado positivo. As demais palavras observadas passaram pelo mesmo desafio. Por isso foi preciso um processamento preliminar manual, onde analistas montaram uma tabela de polaridade (cientistas de dados). A expectativa é de fossem analisados cerca de 5 milhões de tweets por jogo, seria necessário utilizar uma abordagem Big Data. Somente os sistemas que implementam esse conceito possuem capacidade analítica necessária pode processar tamanha quantidade de informações em tempo real. O resultado dessa experiência foi divulgada na emissora de TV Band, com quem a IBM fechou uma parceria para as partidas da Seleção Brasileira. Os tweets expressados pela população posicionou a IBM como o principal *player* de soluções de análise de Big Data do mundo.

#### **4.6 – Transformando Conhecimento em Sabedoria**

Mas a grande fronteira nesse novo eco sistemas de dados é a capacidade de sistemas computacionais de interpretar as palavras em um dado contexto. Entender o real sentido no qual uma palavra foi usada em dado momento. E a partir de um conjunto de contextos (conhecimento) e significados, elaborar uma recomendação (sabedoria) a

respeito do que lhe foi perguntado. É isso que a IBM está fazendo com o Watson, convertendo conhecimento em sabedoria.

Flávio Mendes, um executivo da IBM para soluções corporativas em mídias sociais, em seu artigo de fevereiro de 2014 [37], nos diz que chegamos na era da computação cognitiva. Onde os próximos sistemas computacionais terão a capacidade de aprender. E assim produzirem conclusões, e não apenas um amontoado de números compilados para ainda serem analisados por especialistas. Com aproximadamente 2.000 profissionais trabalhando por trás, o Watson Group é a solução de computação cognitiva da IBM. Lançada em janeiro de 2014. A solução de inovação cognitiva em nuvem já está disponível através de serviço de software. O principal objetivo dessa solução é “ensinar as pessoas” através de respostas qualitativas, a partir de análises de dados históricos e em tempo real. Ou seja, a partir de toda a informação digital disponível o Watson vai encontrar a melhor resposta para uma pergunta.

A solução de computação cognitiva da IBM está inserida 100% no Big Data. Programado para aprender com o usuário, Watson é um sistema computacional que entende a linguagem natural do ser humano. Ele não precisa ser programado via linguagem computacional padrão. Ele foi concebido para aprender com as perguntas que são feitas para ele e com as leituras dos conteúdos que ele faz o tempo todo. Varrendo os dados da Web ele não para de aprender. A medida que esses eventos vão ocorrendo ele vai ficando mais “inteligente”. E então começa a responder as perguntas, a medida que ele erra, ele aprende. Dessa forma ele desenvolve cognição, de forma semelhantes ao nosso cérebro, como vimos na seção 2.2.

Para uma completa explicação sobre o Watson, veja o vídeo lançado pela IBM “2014 IBM Watson Special Event”, disponível no Youtube [39].

A IBM quer compartilhar essa solução e por isso lançou três serviços com inteligência cognitiva, ou fazendo um paralelo com o que aprendemos na seção 2.2, softwares com capacidade de aprendizado:

1. **IBM Watson Discovery Advisor** tem como objetivo revolucionar a forma como indústrias farmacêutica e de publicação realizam suas pesquisas. Seu objetivo é aprofundar a fluência de conteúdo orientado a dados (data-driven) que hoje os pesquisadores se deparam dia a dia, além de descobrir mais conexões que podem acelerar e aprimorar o desempenho deste serviço [38].

2. **IBM Watson Analytics** permite aos usuários explorar ideias de Big Data através de representações visuais, sem a necessidade de simulações de análises avançadas. O serviço remove impedimentos comuns no processo de descoberta de dados, permitindo aos usuários do negócio, de forma rápida e independentemente, descobrirem novos *insights* em seus dados. Guiado por análises sofisticadas e uma interface de linguagem natural, esse serviço prepara automaticamente os dados com os relacionamentos mais importantes e apresenta os resultados em um formato visual interativo e fácil de interpretar.
3. **IBM Watson Explorer** foi concebido para ajudar os usuários a descobrirem *insights* mais rápido, através de análises orientada a dados (data-driven), ajudando as organizações a lançarem suas iniciativas de Big Data com mais rapidez. Esse serviço fornece aos usuários uma visão unificada, exibindo todas as suas informações, bem como toda uma estrutura para o desenvolvimento de aplicações gráficas, disponibilizando visões relevantes de qualquer tópico para os usuários de negócio, cientistas de dados, além de uma variedade de funções de negócio específicas.

Esse conjunto de serviços é uma central de exploração da Internet. A IBM deseja prover esses serviços para as empresas construírem seus próprios aplicativos cognitivos. Com esse novo [meio]ambiente de computação cognitiva, a IBM pretende fomentar uma nova fase da história de computação, auxiliando a criação de milhares de serviços e aplicações cognitivas. A IBM também está fazendo parceria com universidades, o objetivo é formar estudantes voltados para computação cognitiva, big data e ciências analíticas. Iniciativa que também inclui a criação de novos currículos, programas de estágios e novos programas universitários voltados para Watson [38].

Fazendo uma comparação histórica, podemos dizer que a Google permitiu que as pessoas usassem a Internet. O Google tornou o uso da Internet mais amigável ao permitir que encontrássemos as páginas HTML, simplesmente referenciando coisa que queremos encontrar. A Amazon, que ficou conhecida pelas pessoas por vender livros no mundo todo, tem virado de cabeça para baixo os custos de infraestrutura de TI, com seus serviços na nuvem que têm viabilizado cada vez mais pequenas empresas utilizarem grandes soluções relativas a computação distribuída. E agora a IBM com seu Watson, que vai permitir que as pessoas tenham acesso não apenas a dados mas ao

conhecimento, multiplicando exponencialmente nossa capacidade cognitiva. Não é mais a informação em todo lugar, mas sim a sabedoria a onde estivermos.

Um extraordinário exemplo da transformação desse conhecimento foi o experimento da IBM, chefiada pelo Dr. José Baselga, cirurgião-chefe no Memorial Sloan Kettering Cancer Center, para o tratamento do Câncer. No mesmo vídeo, (*“2014 IBM Watson Special Event”*) Baselga fala que os processos normais de atualização dos médicos em relação as novas experiências/resultados publicadas no mundo, já não funcionam mais. São centenas de artigos publicados diariamente, o que torna impossível para os médicos acompanharem tanta evolução. Ocorre que muitos pacientes poderiam ter sido salvos por esse ou aquele artigo de medicina, mas que somente poucos médicos no mundo tem conhecimento. O que faz com que certos tratamentos demorem para ser praticados. De posse dessa enorme base de conhecimento da medicina, uma equipe de médicos oncologistas chefiada por Baselga ensinou o Watson a lógica da medicina. Literalmente, eles colocaram o Watson na universidade para aprender medicina, como foi feito na Copa das Confederações com as palavras no contexto do futebol. A partir daí eles submeteram ao Watson informações de uma paciente. Watson, através de gráficos e textos embasados em todo o conhecimento que ele aprendeu, “disse” para os médicos que eles não tinham informações suficientes para submeter a paciente a um tratamento de câncer de pulmão. Porém, dado as informações que ele tinha, recomendou que os médicos realizassem determinado procedimento. Pois para o Watson, como a paciente era asiática e não fumante, existia a possibilidade de ser um câncer com um certo tipo de mutação.

Como Watson sabia disso? Ele havia “lido” um artigo sobre esse assunto, que coincidentemente havia sido publicado recentemente e que somente poucos médicos no mundo conheciam. Mas para a sorte dessa paciente, esse artigo foi analisado pelo Watson e incorporado na sua rede cognitiva. O que possibilitou aos médicos acesso ao conhecimento desse artigo e de suas afirmações. Para confirmar, os médicos submeteram o tumor a análise de DNA e descobriram a mutação no câncer. O Watson estava 100% certo [39].

# Capítulo 5 Conclusão e Trabalhos Futuros

## 5.1 – Conclusão

As relações sociais evoluíram ao longo dos anos graças a comunicação. Com o indivíduo passando a persistir a informação no papel, nota-se um aumento na velocidade com que as invenções vão surgiram ao longo da História. E as mídias digitais aumentaram essa velocidade. Já é possível observarmos várias soluções e quebras de paradigma nesse início de século.

A computação cognitiva pretende trazer para as soluções de tecnologia o mesmo comportamento que o cérebro tem ao interpretar as informações e tomar decisões. Essa é apenas uma das formas sobre como as mídias digitais estão sendo utilizadas.

Tanto o usuário, que passa a ter disponível essa quantidade de informação, quanto as empresas, têm agora a sua disposição mecanismos para acelerar sua cognição, ou seja, seu processo de formação do conhecimento.

Quanto ao processo de cognição em si, não percebemos alteração. Até porque uma afirmação desse teor deve ser avaliada pelos estudiosos da área. Contudo, está claro que todas as etapas (captação, armazenamento, processamento e tomada de decisão) são impactadas com a velocidade com que agora elas estão sendo demandadas, em detrimento da velocidade e da quantidade de informação que o indivíduo recebe da Internet.

Antes da Internet o processo de cognição não era tão exigido, pois a informação não aparecia instantaneamente na sua frente. Principalmente livros técnicos, os bons em inglês e sempre difíceis de achar. O tempo, e porque não dizer também o custo, para se obter o conhecimento eram muito maior.

Aprender algo novo ficou muito mais fácil, profissões se fundiram em detrimento de não ser mais necessário ter uma pessoa exclusivamente para tal atividade. Essa fusão de atividades só foi possível graças aos impactos cognitivos que os indivíduos vem sofrendo, criando conhecimento cada vez mais rápido a cada dia. A etapa de captação do processo cognitivo foi o principal ponto impactado. Com as novas possibilidades trazidas pelas mídias digitais, o indivíduo é obrigado a melhorar sua

capacidade de avaliar as informações com rapidez, os indivíduos ganham mais conhecimento.

O dado bruto, que brota das mídias digitais, é o novo petróleo. E quem souber refinar terá mais vantagem competitiva. O conhecimento é o fator chave de sucesso para as empresas, e para as pessoas não haveria de ser diferente. O marketing está mais empolgado do que nunca com todas as análises preditivas que podem ser realizadas nos dados que não param de chegar pelas redes sociais.

Como citei no epílogo, a insatisfação possibilita a evolução. Graças ao descontentamento em seguir aos padrões, paradigmas são quebrados e o mundo continua evoluindo. A insatisfação com as análises de dados estratégicas (BI) fizeram surgir as análises de dados em tempo real. Onde as consultas passaram a ser aplicadas diretamente em imensos fluxos de dados e em qualquer tipo de mídia (Big Data) e não apenas nas redes sociais.

Conhecemos também o que talvez seja a mais significativa quebra de paradigmas deste século. A solução de computação cognitiva da IBM cria todo um eco sistema, voltado não apenas para a captura de dados, mas principalmente para a transformação desses dados em conhecimento. Um dos objetivos dessa solução é promover a produção de outras aplicações cognitivas a partir dela, e que por isso pode ser entendida como o ponto de inflexão para a nova tendência do uso da tecnologia.

Vejo a computação cognitiva como uma tendência na humanidade para todo o século XXI. Como uma pessoa que, desde sempre, soube como “dizer” para os computadores o que eles tinham que fazer, entendo que as formas de interação com os computadores serão cada vez mais amigáveis. Não necessitando de tanto conhecimento técnico, ou porque esse conhecimento técnico já estará no cotidiano das futuras gerações ou porque os novos dispositivos terão capacidade cada vez próxima de interagir com o ser humano. Seja na utilização da linguagem natural ou em análises de expressão facial, por exemplo.

A tecnologia torna a comunicação algo cada vez mais e mais natural. Crianças são incentivadas a utilizarem meios mídia digital nas escolas (tabletes). O que faz com elas sejam automaticamente cobradas a terem uma compreensão geral das diversas tecnologias disponíveis, seja oficialmente (escola) ou socialmente (pelos amigos). A Internet é, sem dúvida, o canal mais eficaz de comunicação de todos os tempos. E-mail,

MSN, Facebook, etc, aproximaram incrivelmente as pessoas e criaram novas comunidades de formas nunca antes pensadas

A Internet é um espaço onde o público e o privado, o local e o global, o material e o virtual coabitam. O que promove novas sociabilidades e reorganização das já existentes. Cabe a nós conhecer a melhor forma de usufruir desta, não tão nova, forma de relação.

## **5.2 – Trabalhos Futuros**

Chegamos a um ponto onde não existe retorno. Temos que seguir na transição das formas de comunicação se adaptando às novas plataformas que virão. Na atual era da Informação, como definidas por Paul Virilio em seu livro “L'inertie polaire”, *“as novas mídias trazem consigo a promessa de ampliar as facilidades de uma sociedade tecnocrata, além de determinar o redesenho do espaço doméstico e, em outra escala, os modos comportamentais nas áreas sob influência cultural metropolitana.”*.

Seja em processos de análises de Big Data; seja em experiências de realidade aumentada, como em lojas de roupa com espelhos inteligentes; ou mesmo o simples cenário de uma pessoa de se reunir na sua sala, através de uma TV gigante, com seus amigos de infância que estão do outro lado do mundo; do mais culto à pessoa com menos posses, desde tenha tido qualquer tipo de contato expressivo com as mídias digitais; a tendência, em qualquer estudo sobre a cognitividade e a tecnologia, é que se verifique que a nossa capacidade de cognição, inexoravelmente estão se ampliando.

Por isso, ainda veremos muitas pesquisas envolvendo os temas de Big Data, Computação na Nuvem, Mobilidade e Social Business, tudo isso bastante relacionado com a questão do século: Computação Cognitiva.

# Bibliografia

- [1] FRIEDMAN, T. F. O Mundo é Plano: uma história breve do século XXI. 2a edição. Rio de Janeiro: Objetiva, abril de 2009. 488 p.
- [2] PELAEZ, Lucas. Apresentação da disciplina de Mídia do curso de Publicidade e Propaganda do Instituto de Ensino Luterano de Santa Catarina (IELUSC) com o tema Novas Mídias: Aula Mídia - Novas Mídias: aula 11. In: SlideShare. 22 de novembro de 2012. Disponível em: <<http://pt.slideshare.net/lucaspelaez/aula-mdia-novas-mdias>>. Acesso em: 25 de abril de 2015
- [3] IPTV. In: Wikipédia, a enciclopédia livre. Disponível em:<<http://pt.wikipedia.org/wiki/IPTV>>. Acesso em: 25 de abril de 2015.
- [4] BECKER, Maria Alice d'Avila: Doutora em Psicologia pela Pontifícia Universidade Católica do Rio grande do Sul (PUCRS). Cognitivismo. dezembro de 1997. Disponível em:<<http://penta.ufrgs.br/~maria/cog01.htm>>. Acesso em: 28 de julho 2015.
- [5] Epistemologia. In: Wikipédia, a enciclopédia livre. Disponível em:<<https://pt.wikipedia.org/wiki/Epistemologia>>. Acesso em: 28 de julho 2015.
- [6] SKINNER, B. F. Sobre o Behaviorismo. São Paulo: Cultrix, 1974. pág. 7-11. Introdução.
- [7] Continngencia triplice. Márcio Borges Moreira. Instituto Walden4 - Pesquisa, Educação e Qualidade de Vida. Publicado em 16 de janeiro de 2014. In Youtube. Disponível em:<<https://www.youtube.com/watch?v=xUTmnY5NP9s>>. Acesso em: 30 de julho 2015.
- [8] LILIENFELD, S., LYNN, S. J., NAMY, L., WOOLF, N.. Psychology: A Framework for Everyday Thinking, Pearson, 2010, pp. 24–28, ISBN 978-0-205-65048-4
- [9] Escolhas, tomada de decisão e preferência: uma interpretação analítico-comportamental. apud Referências. In: Instituto Walden4 Pww4. novembro de 2013. Disponível em: <[http://www.walden4.com.br/pww4/index.php?title=Escolhas\\_tomada\\_de\\_decis%C3%A3o\\_e\\_prefer%C3%A2ncia:\\_uma\\_interpreta%C3%A7%C3%A3o\\_anal%C3%ADtico-comportamental#Introdu.C3.A7.C3.A3o](http://www.walden4.com.br/pww4/index.php?title=Escolhas_tomada_de_decis%C3%A3o_e_prefer%C3%A2ncia:_uma_interpreta%C3%A7%C3%A3o_anal%C3%ADtico-comportamental#Introdu.C3.A7.C3.A3o)>. Acessado em: 30 de julho 2015
- [10] What is Inbound Marketing. In: Slideshare. 3 de janeiro de 2014. Disponível em:<<http://pt.slideshare.net/hubspot/what-is-inbound-marketing-29665969>>. Acesso em: 1 de agosto 2015.
- [11] 3 “Big Data” Metrics to Unlock Inbound Marketing ROI. In: V2 Marketing Communications Blog. Chris Kelley. 28 de março de 2014. Disponível em:<<http://blog.marketingv2.com/bid/340834/3-Big-Data-Metrics-to-Unlock-Inbound-Marketing-ROI>>. Acesso em: 1 de agosto 2015.



- [12] In: Optracker: event tracker and analytics. Disponível em:<<http://www.opentracker.net/article/why-do-i-need-website-tracker>>. Acesso em: 1 de agosto 2015.
- [13] Jean Piaget. In: Wikipédia, a enciclopédia livre. Disponível em:<[https://pt.wikipedia.org/wiki/Jean\\_Piaget](https://pt.wikipedia.org/wiki/Jean_Piaget)>. Acesso em: 1 de agosto 2015.
- [14] Seymour Papert. In: Wikipédia, a enciclopédia livre. Disponível em:<[https://pt.wikipedia.org/wiki/Seymour\\_Papert](https://pt.wikipedia.org/wiki/Seymour_Papert)>. Acesso em: 1 de agosto 2015.
- [15] PIAGET, J, O Nascimento da Inteligência na Criança, 4ª edição, Rio de Janeiro, Zahar, 1982.
- [16] MALCON, T., MSc. A construção do conhecimento Segundo Piaget. In: Cerebro e Mente. Disponível em:<<http://www.cerebromente.org.br/n08/mente/construtivismo/construtivismo.htm>>. Acesso em: 1 de agosto 2015.
- [17] PIAGET, Jean. Biologia e Conhecimento. 2ª Ed. Vozes. Petrópolis, 1996.
- [18] Guia\_Michelin. In: Wikipédia, a enciclopédia livre. Disponível em:<[https://pt.wikipedia.org/wiki/Guia\\_Michelin](https://pt.wikipedia.org/wiki/Guia_Michelin)>. Acesso em: 2 de agosto 2015.
- [19] Apache\_Hadoop. In: Wikipédia, a enciclopédia livre. Disponível em:<[https://en.wikipedia.org/wiki/Apache\\_Hadoop](https://en.wikipedia.org/wiki/Apache_Hadoop)>. Acesso em: 2 de agosto 2015.
- [20] BERTOLO, L. Probabilidade: Aula Quatro. In: Site Professor Bertolo. 23 de julho de 2012. Disponível em:<<http://www.bertolo.pro.br/FinEst/Estatistica/EstatisticaNosNegocios/l4.html>>. Acesso em: 3 de agosto 2015.
- [21] Big Data. In: Wikipédia, a enciclopédia livre. Disponível em:<[https://en.wikipedia.org/wiki/Big\\_data](https://en.wikipedia.org/wiki/Big_data)>. Acesso em: 3 de agosto 2015.
- [22] Critical Questions for Big Data. In: Information, Communication and Society. BOYD, B.; CRAWFORD, K.. Published online: 10 May 2012..Disponível em:<<http://www.tandfonline.com/doi/pdf/10.1080/1369118X.2012.678878>>. Acesso em: 3 de agosto 2015.
- [23] LANEY,D. Deja VVVu: Others Claiming Gartner's Construct for Big Data. In: Gartner Blog Network. Publicado online em: 14 de janeiro 2012. Disponível em:<<http://blogs.gartner.com/doug-laney/deja-vvvue-others-claiming-gartners-volume-velocity-variety-construct-for-big-data/>>. Acesso em: 3 de agosto 2015.
- [24] Application Delivery Strategies:3D Data Management Controlling Data Volume Velocity and Variety. In: META Group Inc. Publicado online em: 6 de fevereiro de 2001. Disponível em:<<http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>>. Acesso em: 3 de agosto 2015.
- [25] Apofenia. In: Wikipédia, a enciclopédia livre. Disponível em:<<https://pt.wikipedia.org/wiki/Apofenia>>. Acesso em: 5 de agosto 2015.

[26] Twitter tem 75 milhões de usuários e 40% nunca sequer enviou um tweet. 29 de janeiro de 2010. Disponível em:<<http://www.mdig.com.br/?itemid=10184>>. Acesso em: 3 de agosto 2015.

[27] Malware. In: Wikipédia, a enciclopédia livre. Disponível em:<<https://pt.wikipedia.org/wiki/Malware>>. Acesso em: 5 de agosto 2015.

[28] Otimização para motores de busca. In: Wikipédia, a enciclopédia livre. Disponível em:<[https://pt.wikipedia.org/wiki/Otimiza%C3%A7%C3%A3o\\_para\\_motores\\_de\\_busca](https://pt.wikipedia.org/wiki/Otimiza%C3%A7%C3%A3o_para_motores_de_busca)>. Acesso em: 5 de agosto 2015.

[29] Embraer assina acordo para manutenção de motores Rolls-Royce AE3007 no Brasil. In: Embraer. Acessoria de Imprensa: 13 de agosto de 2014. Disponível em:<<http://www.embraer.com/pt-BR/ImprensaEventos/Press-releases/noticias/Paginas/Embraer-assina-acordo-para-manuten%C3%A7%C3%A3o-de-motores-Rolls-Royce-AE3007-no-Brasil.aspx>>. Acesso em: 5 de agosto 2015.

[30] Google Flu Trends Data: In Google. Disponível em:<<http://www.google.org/denguetrends/>>. Acesso em: 5 de agosto 2015.

[31] BARRETO, M. P. F. L.. Cientista de Dados: a profissão da moda. In: Canaltech. Publicado online em: 11 de fevereiro de 2014. Disponível em:<<http://corporate.canaltech.com.br/noticia/carreira/Cientista-de-Dados-A-profissao-da-moda/>>. Acesso em: 6 de agosto 2015.

[32] PATIL, D. J. Cientista de dados: o profissional mais cobiçado do século 21. In: Harvard Business Review Brasil. Publicado online em: julho 2014. Disponível em:<<http://hbrbr.com.br/cientista-de-dados-o-profissional-mais-cobicado-do-seculo-21/>>. Acesso em: 5 de agosto 2015.

[33] CHEDE, C. Cientista de Dados: o profissional de Big Data. In: IBM Developerworks. Publicado online em: 21 de outubro 2013. Disponível em:<[https://www.ibm.com/developerworks/community/blogs/ctaurion/entry/cientista\\_de\\_dados\\_o\\_profissional\\_de\\_big\\_data?lang=en](https://www.ibm.com/developerworks/community/blogs/ctaurion/entry/cientista_de_dados_o_profissional_de_big_data?lang=en)>. Acesso em: 5 de agosto 2015.

[34] Chief Data Officer. In: Wikipédia, a enciclopédia livre. Disponível em:<[https://en.wikipedia.org/wiki/Chief\\_data\\_officer](https://en.wikipedia.org/wiki/Chief_data_officer)>. Acesso em: 5 de agosto 2015.

[35] MENDESSES, F. IBM Research: Análise de Sentimento durante a Copa das Confederações. In: Social Business e Cloud por Flávio Mendes. Publicado online em: 12 de junho 2013. Disponível em:<<http://fgfmendes.blogspot.com.br/2013/06/ibm-research-analise-de-sentimento.html>>. Acesso em: 7 de agosto 2015.

[36] Análise de Sentimentos em Redes Sociais. In: IBM Big Data & Analytics. Disponível em:<<http://www.ibm.com/big-data/br/pt/big-data-and-analytics/marketing/analise-de-sentimentos-em-redes-sociais/>>. Acesso em: 7 de agosto 2015.

[37] MENDESSES, F. Watson e a Era da Computação Cognitiva. In: Social Business e Cloud por Flávio Mendes. Publicado online em: 11 de fevereiro 2014. Disponível

em:<<http://fgfmendes.blogspot.com.br/2014/02/watson-e-era-da-computacao-cognitiva.html>>. Acesso em: 5 de agosto 2015.

[38] IBM Forms New Watson Group to Meet Growing Demand for Cognitive Innovations. In: IBM News room. Publicado online em: 09 de janeiro 2014. Disponível em:<<http://www-03.ibm.com/press/us/en/pressrelease/42867.wss>>. Acesso em: 8 de agosto 2015.

[39] 2014 IBM Watson Special Event. IBM. Publicado em 09 de janeiro de 2014. In Youtube. Disponível em:<<https://www.youtube.com/watch?v=654xqu7MWzs>>. Acesso em: 5 de agosto 2015.

[40] TAURION, C. Big Data. São Paulo: BRASPORT Livros e Multimídia Ltda., 2013.

[41] NUNES, Denise Silva. A importância das novas mídias na proteção do meio ambiente em face da crise ambiental. In. Revista Eletrônica do Curso de Direito da UFSM, v.8., 2013, p. 51-59. Disponível em <[http://cascavel.ufsm.br/revistas/ojs-2.2.2/index.php/revistadireito/article/view/8217#UdpI\\_FLLZUU](http://cascavel.ufsm.br/revistas/ojs-2.2.2/index.php/revistadireito/article/view/8217#UdpI_FLLZUU)>. Acesso em: 28 de julho de 2015.

[42] LEVY, P. Ciberultura. São Paulo: Ed. 34, 1999.

[43] PRUDÊNCIO, K. C. de Souza. Mídia ativista: a comunicação dos movimentos por justiça global na internet. Florianópolis, SC: UFSC, Tese de Doutorado, p.207, 2006.