

DESMISTIFICANDO O MUNDO DO BIG DATA

Matheus Batista Furlan

mbfurlan@hotmail.com.br

Abstract: With the advancement and popularization of the Internet, especially social networks, a gigantic number of unstructured data began to be produced daily in extreme speed, this is called Big Data. However, in order to derive value and information from this volume of data, data mining and post-relational database storage techniques, known as the NoSQL concept, are required since traditional banks do not support this volume and variety of data. The purpose of this research was to present the concepts of Big Data, Data Mining, which is the database mining process and the NoSQL movement, which aims to replace relational databases in the process of high-scale data storage. As a result show a case where these concepts were applied and how important they were to have a great return.

Resumo: Com o avanço e popularização da Internet, principalmente das redes sociais, um número gigantesco de dados não estruturados passou a ser produzido diariamente em velocidade extrema, a isso se dá o nome de Big Data. Porém, para se tirar valor e informação desse volume de dados são necessárias técnicas de mineração de dados (Data Mining) e o armazenamento em Banco de Dados Pós Relacionais, conhecidos pelo conceito NoSQL, visto que os bancos tradicionais não suportam esse volume e variedade de dados. O objetivo dessa pesquisa foi apresentar os conceitos de Big Data, Data Mining que é o processo de mineração da base de dados e o movimento NoSQL, que tem com objetivo substituir os bancos de dados relacionais no processo de armazenamento de dados em escala elevada. Como resultado mostrar um caso em que esses conceitos foram aplicados e como eles foram importantes para se ter um ótimo retorno.

Palavras – chave: Big; Data; Banco; NoSQL; Mining.

1. Introdução

O termo Big Data está cada vez mais popular, porém, ainda não está bem claro o seu significado, a sua aplicabilidade e a sua finalidade. Para uma melhor compreensão, é essencial entender a definição dos 3 V's: Volume + Variedade + Velocidade. O *Volume* representa a grande quantidade de dados gerados por sistemas corporativos, por mídias sociais, sensores e outros dispositivos; a *Variedade* representa os dados estruturados e não estruturados, obtidos do twitter e facebook, dados de uma empresa com grande volume de geração e movimentação de dados; a *Velocidade*, que representa a resposta quase que em tempo real para agir no próprio evento gerador das informações (TAURION 2013).

Big Data vem chamando atenção pela acelerada escala em que volumes cada vez maiores de dados são criados pela sociedade. No entanto, existem muitas dúvidas de como tangibilizar o conceito, ou seja, como sair do conceitual e criar soluções de negócio que mineralizem esta massa de dados, já que a cada dia são gerados dezenas de petabytes de dados, em uma escala real e não mais imaginária e futurista.

O Big Data nos permite a utilização de diversas tecnologias de gerenciamento dessa massiva geração de dados como, aplicação de Data Warehouse, Data Mart, Data Mining, Web Data Mining, BI (Business Intelligence) e Cloud Computing.

1.1 Problematização

Atualmente a quantidade massiva de dados com a estrutura do Big Data vem de sistemas estruturados (que por sua vez são minoria) e não estruturados (maioria), como gerados por e-mails, mídias sociais (Facebook, Twitter, YouTube, dentre outros), documentos eletrônicos, mensagens instantâneas, etc.

No entanto, as tecnologias atuais de gerenciamento de dados, como o modelo relacional proposto por Edgar F. Coode em 1969, não são tão adequadas para suportar os dados com a estrutura do Big Data. O modelo relacional é apropriado para acessar dados estruturados, gerados por sistemas internos das corporações, ou seja, não foi projetado para dados não estruturados, já que não era esta a realidade da época, e nem para volumes na casa dos petabytes de dados como se tem nos dias de hoje.

Para tratar dados na escala de Volume, Variedade e Velocidade do Big Data, precisa-se de outros modelos mais apropriados como os sistemas gerenciadores de bancos de dados NoSQL, projetados para tratar imensos volumes de dados estruturados e não estruturados.

Existem diversos modelos como sistemas colunares como o Big Table, usado internamente pelo Google ou o modelo Key/value como DynamoDB da Amazon e o “document database” baseado no conceito proposto pelo Lotus Notes da IBM e aplicado em softwares como MongoDB.

Esta diversidade de alternativas demanda conhecimento e estudo sobre as tecnologias existentes com o intuito de aprimora-las para que seja possível tratar os dados não estruturados de uma maneira eficaz, considerando o Volume, Variedade e Velocidade com que esses dados são acessados.

1.2 Objetivos

1.2.1 Objetivo geral

Essa pesquisa teve início no Programa de Iniciação Científica de 2014, tendo como resultado, em forma de artigo, a desmistificação dos conceitos relacionados ao Big Data também como a apresentação de um estudo de caso pronto, no sentido de ilustrar como Big Data e Data Mining podem ser essenciais e até mesmos vitais para as empresas.

Porém, um ano foi pouco e como ainda há muito para pesquisar, estudar, etc., pretende-se com a renovação do prazo dessa pesquisa por mais um ano, avançar ainda mais no que diz respeito às tecnologias de Data Mining.

Dessa forma, esse segundo objetivo teve por finalidade explorar as tecnologias que envolvem o processo de Data Mining, uma forma de mineração de dados que abrange técnicas, métodos e etapas para aplicação, no sentido de auxiliar os especialistas na identificação de problemas com armazenamento de dados no dia a dia.

Além disso, como os Bancos de Dados Pós Relacionais vem crescendo e atualizando-se a cada dia, outro objetivo desse segundo ano de pesquisa pretendeu-se, apresentação e exemplificação esse novo conceito de banco de dados, que está se tornando cada vez mais popular no mercado de trabalho, conhecido como NoSQL.

1.2.2 Objetivos específicos

Os objetivos específicos que figuraram na segunda etapa da pesquisa foram:

- Compreensão e divulgação dos conceitos envolvidos na mineralização de dados;

- Desenvolver uma explicação sobre o conceito de Big Data;
- Desenvolvimento de uma análise sobre Sistema Gerenciador de Banco de Dados (MongoDB) – SGDB para iniciação prática ao Big Data;
- Demonstrar um estudo de caso onde foram utilizados os conceitos de Big Data e Data Mining.

1.3 Relevância/Justificativa

O intuito desta pesquisa foi sintetizar os principais conceitos relacionados ao Big Data e tecnologias que o abrangem, tais como Banco de Dados NoSQL.

Espera-se que essa pesquisa apresente como resultado uma base introdutória para conhecimento do tema pesquisado, podendo ser utilizado como um guia, para familiarização com o assunto.

Tendo em vista o interesse e as dificuldades para a compreensão do Big Data, a existência de uma pesquisa que aborde e analise o tema como este pode ser relevante e justo pelo fato de auxiliar na introdução de profissionais ou leigos ao tema ampliando sua grade de conhecimento voltada para o armazenamento de dados.

Este segmento existe há mais de cinco anos no mundo sendo que no Brasil começa aparecer agora com maior destaque (DOURADO, J.). Nesse período, no mundo, aconteceram grandes avanços nas empresas que se utilizam dessas tecnologias. É o caso de empresas como Google, Yahoo e Apple, dentre outras, ou seja, não está se falando apenas do armazenamento de grandes volumes de dados, mas também de processamento intensivo dessas informações.

Tendo em vista que o mesmo vem sendo estudado e divulgado no país este projeto poderá ser utilizado como guia introdutório o profissional da área de TI que terá maior conhecimento sobre Big Data, mesmo que seja um tanto superficial, olhando pela imensa quantidade de tecnologias para aplicação e seus conceitos.

2. Big Data

O conceito de Big Data está a cada dia em mais evidência devido a volume de dados que estão sendo gerados pela interação entre homem e máquina ou exclusivamente pelas máquinas. Em um

artigo TAURION resume o Big Data sendo volume + velocidade + variedade e acrescenta, ainda, dois V's: veracidade e valor.

O volume é a quantidade gigantesca de dados que são gerados; A velocidade é necessária para agir quase em tempo real sobre este imenso volume de dados; A variedade se refere à forma em que os dados aparecem de forma estruturada, que são minoria, e não estruturados, que são a maioria; Veracidade, visto que é necessário ter certeza que os dados fazem sentido e são autênticos; Valor porque é absolutamente necessário que a organização que implementa projetos de Big Data obtenha retorno destes investimentos.

Apesar do Big Data chamar a atenção pela quantidade de dados, sua maior importância está no o que será feito com essas informações. Esses dados, se analisados, podem trazer respostas que permitam a redução de custo e tempo, desenvolvimento de novos produtos e otimização de ofertas e tomadas de decisões mais inteligentes, que antes não poderiam ser vistas.

Outro ponto importante é que o Big Data sozinho não revela seu verdadeiro potencial, é necessário encontrar tecnologias que tirem proveito dessa enorme quantidade de dados. Essas tecnologias devem possuir algumas características como armazenamento barato de grandes quantidades de dados, processadores mais rápidos, Open Source acessível e plataformas distribuídas de Big Data, processamento paralelo, clusterização, virtualização, grandes ambientes de grid, alta conectividade e altas taxas de transferência, computação em nuvem e outros arranjos de alocação flexível de recursos.

O artigo de TAURION (2013) discorre sobre a abordagem do tema Big Data e sua importância na atualidade quanto à imensa geração de dados estruturados e não estruturados a cada dia. E uma breve abordagem sobre a tecnologia para minerar dados, Data Mining, mostrando como sua aplicação em um servidor que armazena dados em nível de Big Data se faz possível à sumarização das informações diminuindo o tempo de processamento das mesmas.

3. Data Mining

Diariamente volumes enormes de dados são gerados e armazenados, segundo uma pesquisa do EMC de 2014 o volume de dados ultrapassa os 4,4 ZB (zettabytes) e a tendência é que esse número aumente a cada dia.

Essa grande quantidade de dados passa a ter valor a partir do momento em que se torna possível a extração de conhecimento a partir deles, o que é realizado pelo processo de Descoberta de Conhecimento em Banco de Dados ou KDD (Knowledge Discovery in Databases). Neste processo, a Mineração de Dados (MD) é a principal etapa e consiste na combinação de métodos tradicionais de análise de dados com algoritmos sofisticados para processamento de grandes volumes [Tan et al. 2006].

Segundo Han et al. (2011), Fayyad et al. (1996) e Tan et al.(2006), a mineração de dados consiste em uma etapa no processo de KDD que busca descobrir ou obter padrões a partir de um grande volume de dados, os quais representem informações úteis. Sua realização consiste na aplicação de algoritmos que se utilizam de técnicas estatísticas e de inteligência artificial, para realizar a classificação de elementos de um conjunto de dados e prever valores de variáveis aleatórias.

A premissa do Data Mining nasce a partir do usuário definir um problema, selecionar os dados e as ferramentas para analisá-los. Essas ferramentas iram pesquisar os dados à procura de erros e possíveis relacionamentos, encontrando falhas que antes não podiam ser vistas.

O processo do Data Mining nasce a partir de 3 técnicas. A primeira delas é a estatística, sem ela não seria possível o Data Mining, já que ela é base de outras tecnologias, a partir das quais, o Data Mining foi projetado. A segunda é a Inteligência Artificial que busca imitar a forma como o homem pensa na resolução de problemas. A terceira e última é o Machine Learning, que pode ser resumida como a união entre a estatística e a inteligência artificial. O foco do Machine Learning está no desenvolvimento de algoritmos que aprendam a partir de uma base de dados, reconheçam padrões e assim possam tomar decisões inteligentes, baseadas nas informações coletadas.

O Data Mining, também, possui ramificações, alguns exemplos são:

- Redes neurais que são baseadas na computação de ligações. Os nós simples são interligados para formar uma rede de nós;
- Indução de regras é a detecção de tendências dentro de um grupo de dados ou de regras definidas sobre os dados;
- Árvores de decisão baseiam-se em uma análise que testa todos os valores de uma base de dados para identificar aqueles que são fortemente associados com os itens de saída selecionados para o exame;

- Visualização mapeia os dados que estão sendo minerados de acordo com dimensões especificadas. O programa de Data Mining cuida apenas da manipulação da estatística básica, então é o usuário quem tem que interpretar os dados enquanto olha para o monitor.

A apresentação de BARTH (2012) tem início com uma rápida abordagem sobre a mineralização de dados utilizando a tecnologia Web Data Mining. O autor centra o seu esforço nessa última categoria, o Web Data Mining tem como objetivo mineralizar e aperfeiçoar o gerenciamento de operações corporativas diárias. Ainda de acordo com BARTH, o Web Data Mining é capaz de sumarizar os dados em nível de Big Data e extrair do banco de dados apenas a informação sumarizada para a consulta feita no momento.

Os autores GOLDMAN *et al* (2011) em pesquisa realizada no artigo “Apache Hadoop: conceitos teóricos e práticos, evolução e novas possibilidades” tratam a mineração de dados das redes sociais utilizando o framework Apache Hadoop para o mesmo mostrando a eficácia obtida por esta plataforma onde pode ser constatada ao verificar a quantidade de importantes empresas, de diferentes ramos, que estão usando Hadoop para fins educacionais ou de produção, como exemplo a Yahoo!.

4. Banco de Dados NoSQL

O aumento gigantesco de dados que são gerados diariamente gerou uma necessidade que os bancos de dados tradicionais não conseguiam suporter, por isso o movimento NoSQL ganhou espaço, com a motivação de resolver o problema de escalabilidade dos bancos de dados relacionais.

Esse movimento ficou bem enraizado no Open Source, porém ganhou muita força com a divulgação das implementações proprietárias do Google BigTable e o Amazon Dynamo, ambas empresas que lidam com grande volume de dados.

O termo NoSQL é usado para descrever bancos de dados não relacionais ou pós relacionais. Suas características são registros, schema-free, tolerância à falha, escalabilidade, clusterização, mapreduce, sharding. As vantagens de um banco de dados NoSQL são facilidades de desenvolvimento, desempenho escalável, alta disponibilidade e resiliência.

Tiwari (2011) comenta que, nem todas as bases de dados NoSQL são similares e criadas para resolver o mesmo problema, mas é essencial entender, qual se encaixa à determinada situação.

Os bancos NoSQL são subdivididos pela forma como ele trabalha com os dados. Alguns exemplos são:

- Key/Value Store seu conceito é de uma chave e um valor para ela, porém, ele quem aguenta a carga dos dados.
- Wide Columns Store foi inspirado no BigTable, do Google, ele suportam grande quantidade de linhas e colunas, além de permitir subcolunas.
- Document Store é baseado em documentos XML ou JSON podem ser localizados por *id* único ou por qualquer registro que tenha o documento.
- Graph Store esse tipo de banco guardam objetos e a busca dos itens é feita pela busca desses objetos.
- Column Oriented Store esses bancos são relacionais, porém, possuem características do NoSQL, sendo que a principal diferença é o armazenamento de dados em colunas, o que ajuda na escalabilidade.

Bases de Dados NoSQL possuem tamanhos, e formas diferentes e muitas soluções para eventuais problemas que podem ser facilmente adequados a suas características. Portanto, NoSQL nos dias de hoje não passa apenas de uma nova adoção, que adequa-se ao fato de estar ligado diretamente às aplicações que necessitam de alto desempenho na web (BARASUOL, 2012).

O exemplo de um banco não relacional é o MongoDB, que é orientado a documentos diferente do modelo relacional.

Chodorow e Dirolf (2010), explicam que MongoDB é um banco de dados orientado a documentos, e não relacional, onde a ideia básica é trocar o conceito de “linhas”, por algo mais flexível como “documentos”, essa orientação a documentos permite uma representação hierárquica e complexa de relações em apenas um registro.

Este Banco de Dados tem como característica ser de código-fonte aberto, licenciado pela GNU AGPL (Affero General Public License) versão 3.0, possuir alta performance, não possuir esquemas, ser escrito em C++, multiplataforma e ser formado por um conjunto de aplicativos JSON.

As vantagens de utilizar o MongoDB são uma melhor performance, consultas simples de serem realizadas, o escalonamento horizontal com Sharding é muito bem implementado, é muito simples converter instruções SQL em chamadas de função de consulta e, por fim, ele também possui a funcionalidade chamada GridFS que é responsável por armazenar arquivos de grandes dimensões.

5. Experimento

Durante a Copa do Mundo de 2014 no Brasil, a seleção alemã se utilizou das tecnologias do Big Data para aprimorar seus treinos e análises de adversários, com a ajuda da também alemã SAP, empresa de desenvolvimento de soluções em TI. Foi criada uma ferramenta que conseguia analisar inúmeros dados de treinamentos e jogos, o que ajudou a melhorar o desempenho do time nos jogos.

A solução que foi chamada de Match Insights foi desenvolvida pela empresa de TI em parceria com a confederação alemã de futebol. O aplicativo, que roda na plataforma HANA da própria SAP possibilita que o técnico e os preparadores processem todo o volume de informações que é gerado em uma partida ou mesmo na preparação.

HANA é a plataforma de computação IN-Memory da SAP e significa High Performance Analytic Appliance (ferramenta analítica de alto desempenho) e foi desenvolvida para aplicativos analíticos que podem ser implementados em alta velocidade e também para muitos dados, se tornando perfeita para o Big Data.

Os resultados conseguidos foram excelentes, já que além de vencer a copa do mundo, o time alemão conseguiu uma superioridade estatística sobre todos os adversários. Isso se deve também a facilidade que a interface do aplicativo, visto que até mesmo os jogadores tinham acesso aos dados e podiam analisá-los e entendê-los por conta própria.

Além do lado esportivo o aplicativo Match Insights da SAP também foi útil na parte mais social porque ele analisava as redes sociais para avaliar o humor e a participação de torcedores. Os dados que eram adquiridos dessas análises em cada partida eram exibidos em vídeos compartilhados pela empresa de TI, mostrando da porcentagem de aprovação do time pelos fãs até os jogadores mais populares no decorrer da partida.

Esses tipos de ferramentas estão se tornando cada vez mais comuns no meio esportivo, outros esportes além do futebol como basquete, vôlei, futebol americano entre outros, também fazem uso de soluções big data, pois é notável a diferença de desempenho de equipes que tem um acesso mais privilegiado aos dados daquelas que não tem, se tornando essencial para aqueles que querem alto rendimento e disputar competições em excelente nível..

6. Considerações Finais

O conceito de Big Data está relacionado ao grande crescimento do volume dados que é produzido

diariamente. Esse aumento se deve principalmente pelo aumento no uso de aplicações na web, que é uma grande geradora de dados em alta velocidade, porém ainda existem dúvidas de como tornar esse conceito tangível criando soluções de negócio através dessa massa de dados, e para isso serão necessárias o uso de outras tecnologias, importantes, nesse processo.

O processo de descoberta de conhecimento em banco de dados é realizado, nessa massa de dados, para a extração de conhecimento, tendo como principal etapa a mineração de dados ou data mining. Nessa etapa são pesquisados os dados em busca de erros e relacionamentos, encontrando falhas ou tendências que antes não eram vistas.

Essa quantidade de dados na escala do Big Data criou uma necessidade que os bancos de dados relacionais não podiam suportar, como alternativa, grandes empresas aderiram aos bancos de dados não relacionais, conhecidos também como NoSQL. Esses bancos possuem características diferentes ao dos bancos tradicionais, que o fazem uma melhor escolha quando se trata de big data.

O artigo tem como finalidade ser um guia básico de conhecimento sobre o conceito de big data, mostrando também, as principais características do processo de data mining e do sistema de banco não relacional, além de mostrar um experimento em que foram utilizados esses conceitos e de que forma os dados foram transformados em solução.

REFERÊNCIAS BIBLIOGRÁFICAS

BARASUOL, Érion Ricardo. **MongoDB uma base de dados orientada a documentos que utiliza orientação a objetos**. 2012. 101 f. Trabalho de Conclusão de Curso – Instituto Municipal de Ensino Superior de Assis, 2012.

BARTH, J. Fabrício. **Uma Introdução a Mineração de Informações na era do Big Data**. VAGAS Tecnologia e Faculdades BandTec.

BARTH, J. Fabrício. **Uma Introdução á Mineração de Informações na era do Big Data**. 2012. 75. Tipo de trabalho (Titulação) - VAGASTecnologia e Faculdades BandTec.

BIRGNOLI, T. Juliano, JUNIOR, S. Egon, MIGUEZ, B. Viviane, SANTOS, Neri, SPANHOL, Fernando. **A Intervenção Humana na Qualificação de Processos de Data Mining: Estudo de Caso em uma Base de Dados Hipotética**. Universidade Federal de Santa Catarina.

CAPETTA, M. Leonardo. **Consulta a banco de dados em linguagem natural**. Omnia Exatas, v.4, n.1, p.72-80, 2011.

CARVALHO, R.,JEFFERSON. **Conceitos e Técnicas sobre Data Mining**. Disponível em: <<http://www.devmedia.com.br/conceitos-e-tecnicas-sobre-data-mining/19342>>. Acesso em: 05 de Junho de 2016.

CHODOROW, Kristina; Michael DIROLF, Michael.. **MongoDB The Definitive Guide, First Edition**. Gravenstein Highway North: O'Reilly, 2010.

DOURADO, Joana. **Semantix – Treinamentos**. Disponível em: <<http://www.semantix.com.br>>. Acesso em: 05 dezembro. 2013.

FAYYAD, U. M., PIATETSKY-SHAPIRO, G., and SMYTH, P. (1996). **Advances in knowledge discovery and data mining. chapter From Data Mining to Knowledge Discovery: An Overview**, pages 1–34. American Association for Artificial Intelligence, Menlo Park, CA, USA.

GOLDMAN, A., KON, F., JUNIOR, F. P., POLATO, I, PEREIRA, R. F.. **Apache Hadoop: conceitos teóricos e práticos, evolução e novas possibilidades**.

GOMES, M. Heitor., HAUTH, G. Luiz., CARVALHO, R. Deborah. **Mineração de dados temporal: Descoberta de Regras de Causa e Efeito**. Faculdade de Ciências Exatas e Tecnologia (FACET) – Universidade Tuiuti do Paraná – Curitiba – PR – Brasil.

HAN, J., KAMBER, M. and PEI, J. (2011). Data Mining Concepts and Techniques. **The Morgan Kaufmann Series in Data Management Systems. Elsevier Science, 3ª edition**. São Francisco. USA

KIMBALL, R. **The Data Warehouse Toolkit: Guia Completo para modelagem dimensional**. Tradução da segunda edição. Rio de Janeiro: Campus Ltda, 2002.

KOHONEN, Teuvo. **An Introduction to Neural Computing**. Finland: Helsinki University of Technology, pp. 3-16, 1988.

KOHONEN, Teuvo. **Self-Organization and Associative Memory**. 2ª Edição. USA: Spring-Verlag, 1989.

KOHONEN, Teuvo. **The self-organizing map**. Proceedings of the Institute of Electrical and Electronics Engineers, vol.78, pp. 1464-1480, 1990.

LIPPMANN, Richard. **An Introduction of Computing with Neural Nets**. IEEE Computer Society, v.3, nº 4, pp. 4-22, abr.1987.

MAYER-SCHÖNBERGER, Viktor., CUKIER, Kenneth. **BIG DATA, Como Extrair Volume, Variedade, Velocidade e Valor da Avalanche Cotidiana**. Editora Elsevier Ltda, 2013.

MEDEIROS, HIGOR. **Introdução ao MongoDB**. Disponível em: <<http://www.devmedia.com.br/introducao-ao-mongodb/30792>>. Acesso em: 11 de Junho de 2016.

NASCIMENTO, O., J., Rafael. **Mineração e Análise de Dados em SQL**. Disponível em: <<http://www.devmedia.com.br/mineracao-e-analise-de-dados-em-sql/29337>>. Acesso em: 04 de Agosto de 2014.

SANTOS, Ronaldo Canofre M.; BORGES, Eduardo N.; MACHADO, Karina dos Santos. **Uma avaliação de algoritmos para mineração de dados disponíveis na web**. XII Escola Regional de informática de banco de dados. ISSN 2177 – 4226 – SBC. 13 a 15 de Abril, Londrina, PR, 2016.

SBPJor – Associação Brasileira de Pesquisadores em Jornalismo, IX, 2011, Rio de Janeiro. **Jornalismo Computacional em função da Era do Big Data**: 2011.12.

STEPPAT, NICO. Bancos de dados não relacionais e o movimento NoSQL. Disponível em: <<http://blog.caelum.com.br/bancos-de-dados-nao-relacionais-e-o-movimento-nosql/>>. Acesso em: 20 de Novembro de 2016

TAN, P., STEINBACH, M., and KUMAR, V. (2006). **Introduction to Data Mining**. Pearson international Edition. Pearson Addison Wesley.

TAURION, Cezar. **Coletânea de posts publicados no Blog developerWorks em 2012 developerWorks Brasil**. Disponível em: <<http://www.ibm.com/developerworks/blogs/page/ctaurion>>.

TIWARI, Shashank. **Professional NoSQL**. Indianapolis: John Wiley & Sons, 2011.