



## **IO & Big Data**

Uma série de *workshops* sobre a Investigação Operacional (IO) em ambiente *Big Data* (BD) foram desenvolvidos em colaboração entre a Associação Portuguesa de Investigação Operacional (APDIO), a IBM Portugal e o Instituto Politécnico de Portalegre (IPP), tendo decorrido na Escola Superior de Tecnologia e Gestão (ESTG/IPP), Portalegre, em setembro e outubro de 2014.

Esta colaboração com a APDIO e a IBM Portugal vem também dar continuidade a programas internacionais no âmbito da IO que têm sido desenvolvidos na ESTG/IPP, e.g., [http://tiny.cc/sc\\_sem2014/](http://tiny.cc/sc_sem2014/) , tendo em comum alguns dos elementos organizadores e um horizonte de colaboração já com alguns anos. Utilizando-se a habitual abordagem dum programa internacional, o plano de atividades prossegue até à segunda metade de 2015.

A seleção da temática BD baseou-se no papel importante que estas tecnologias podem ter na designada “especialização inteligente”, na agilidade dos agentes do setor público, bem como no impacto potencial em redes sociais, *smart cities*, tecnologias de informação, plataformas móveis para a saúde e cuidados médicos, redes de comunicação, energia e transporte, entre outros campos em desenvolvimento rápido.

Embora o registo e a presença nas diferentes sessões tivessem decorrido de forma autónoma, o acompanhamento integral das sessões foi recomendada a todos os participantes, em especial aos interessados em desenvolvimentos práticos e aplicações BD. O “Preço Reduzido” manteve-se durante todo o horizonte temporal para inscrições, as dificuldades com as inscrições foram sendo ultrapassadas em tempo útil, os elementos do IPP e os sócios da APDIO usufruíram de condições especiais, tendo os elementos externos a possibilidade de se registarem simultaneamente como sócios da APDIO.

Atendendo à pausa estival, em especial com o encerramento das instituições de ensino superior durante parte do mês de agosto, as atualizações ocorreram preferencialmente através de meios eletrónicos e das redes sociais (Facebook-FB).



## **IO & Big Data**

A oferta de alojamento foi oportunamente disseminada, não só a disponibilidade das residências do IPP, mas também através dos sítios eletrónicos dos municípios da área de Portalegre. De igual modo se procedeu para as atividades culturais promovidas na área.

A jornada inicial (09-setembro) reuniu vários oradores em painéis temáticos, permitindo uma visão geral da situação presente e dos desenvolvimentos futuros em cada tema: Saúde, Energia/Indústria, Logística/TIC. Estes temas foram selecionados atendendo à sua atualidade e relevância para aplicações BD, pretendendo abordar-se em cada tema um triplete de vertentes: a da IO, a empresarial, a de inovação/internacionalização. Deste modo, consolida-se a componente geral de formação, permitindo o adequado e consequente desenvolvimento das componentes específicas.

Subsequentemente, as jornadas laboratoriais de 10-setembro e 08/09-outubro dirigiram-se a ferramentas específicas em BD, respetivamente, ferramentas para “Hadoop” e “Big Data” da IBM, visando aplicações laboratoriais e casos de estudo.

### **IO&BD-2014 (I) Jornadas “IO & Big Data” e IBM HADOOP Portalegre, 9/10 de setembro-2014**

As jornadas abriram a 09-setembro com a presença de Luís Batista (Vice-Diretor da ESTG/IPP) representando tanto a ESTG como o IPP, Ana Póvoa (IST/UL) representando a APDIO, Cristina Semião (*Healthcare & Public Sector*, IBM Portugal ) em representação da IBM Portugal, João Miranda (Comissão Organizadora, ESTG/IPP), bem como dos ilustres moderadores e oradores, participantes, e ainda alguns Colegas da ESTG e de outras escolas do IPP. Foram brevemente lembradas as necessidades gerais e específicas de formação, nomeadamente as necessidades decorrentes da juventude da temática BD, da modernização de equipamentos e ferramentas tecnológicas, da continuidade da boa dinâmica do “triângulo do conhecimento” que envolve investigação, inovação e educação.



## IO & Big Data

O escalonamento dos painéis temáticos foi ligeiramente ajustado, pois por motivos de força maior não foi possível concretizar o painel em **Energia/Indústria**. Agradecendo a presença da Moderadora convidada, Anabela Oliveira ( [ESTG/IPP](#) ), a possibilidade de re-alendarizar este painel ficou em ponderação.

### a) Painel em Saúde

O painel em Saúde teve a moderação de Adriano Pedro ( [ESS/IPP](#) ), o qual introduziu este tema partilhando uma visão alargada, completa, bem como apresentou detalhadamente os oradores nas três vertentes: IO, empresarial, e inovação/internacionalização. As contribuições dos oradores seguiram por ordem alfabética, havendo ainda disponibilidade para esclarecer alguns pontos apresentados pelos participantes. Descrevem-se em seguida os pontos principais que os oradores focaram:

- Cristina Semião (Healthcare & Public Sector, [IBM Portugal](#) ) abordou a sustentabilidade do sistema de saúde em face da evolução demográfica; os programas de intervenção, a situação económica e a gestão da saúde em termos do estilo de vida, situações não clínicas, e tratamento da doença; contextualizou as tecnologias BD, a “medicina com base na evidência”, e a necessidade dos sistemas de informação em Saúde; indicou ainda alguns desenvolvimentos no âmbito do Ministério da Saúde, da Direção Geral de Saúde, e do potencial papel do Watson no apoio à decisão médica.
- [Maria Eugénia Captivo](#) (FCUL) introduziu alguns problemas de IO que têm vindo a ser analisados no setor da Saúde, distinguindo aí a abordagem em BD; enquadrou os contributos do *working group* em ORAHS-*Operational Research Applied to Health Services*, bem como indicou alguns desenvolvimentos de IO em termos de serviços de saúde; focou o planeamento de cirurgias, os modelos e técnicas de resolução, a redução de custos, os atributos dos profissionais de saúde, bem como o volume e a velocidade de informação a tratar; como desafios principais em BD referiu as fontes de informação, a respetiva análise de dados e as evidências a retirar, relevando a qualidade de resposta ao doente como o ponto crítico.



## IO & Big Data

- [Nuno Garcia](#) (UBI) revisitou brevemente alguns dos pontos apresentados anteriormente, adicionando pontos históricos aos fundamentos do quotidiano presente; introduzindo a Internet do futuro, explanou a extração de valor da informação e do volume de dados em BD, especificando o desafio atual do “open-data”, da transparência e da anonimidade em face da validade e privacidade dos dados; além das dificuldades em tratar grandes volumes de dados, da computação ubíqua no estudo de moléculas e em aplicações farmacêuticas, também posicionou as ferramentas BD em termos da medicina convencional, dos cuidados personalizados, e do enquadramento para a disponibilização dos dados pessoais; como desenvolvimentos dum longo futuro em IO e BD, encadeou a recolha e estruturação dos dados pessoais, a integração dos sistemas de informação, o impacto da BD e das aplicações móveis nos sistemas de saúde.

A discussão desta temática foi alargada aos participantes, numa troca de impressões bem interativa sobre saúde que focou principalmente: a estruturação dos dados; *i*) o custeio e a definição de políticas de saúde; *ii*) as ferramentas de IO mais importantes; *iii*) alguns casos típicos, situações específicas.

### **b) Painel em Logística/TIC**

O painel em Logística/TIC teve a moderação de Valentim Realinho ( [ESTG/IPP](#) ), o qual procedeu a breve introdução da temática e à apresentação dos três oradores, nas vertentes referidas da IO, empresarial, e inovação/internacionalização. As contribuições dos oradores foram acompanhadas com a exposição breve de alguns diapositivos, havendo ainda um período para discussão com a audiência. Os principais pontos que foram explorados pelos oradores são indicados em seguida:



## IO & Big Data

- [Ana Paula Póvoa](#) (IST/UL) apresentou uma perspectiva geral das cadeias logísticas, referindo boas práticas em diversos setores, no setor público, na saúde, na distribuição, ou em produção industrial; destacou a eficiência operacional e exemplificou-a, bem como relevou as mais-valias da experiência do consumidor e do valor para o cliente; focou a gestão de risco em cadeias logísticas, os novos modelos de negócio; identificou fatores de sucesso para a IO em BD, como o conhecimento científico, o desenvolvimento de técnicas e métodos, o tratamento e a *clusterização* de dados.
- José Eduardo Fonseca (*Strategic Outsourcing - Service Delivery & Chief Technologist*, [IBM Portugal](#) ) introduziu as principais inovações na linha de desenvolvimento dos serviços tecnológicos, identificando os principais passos na evolução para a era dos sistemas cognitivos; descreveu as principais eras de tecnologia até à confluência a nível global das tecnologias Mobile, Cloud, Big Data/Analytics; caracterizou a geração deste milénio, as mega-tendências, a atividade massiva dos meios multimédia e a produção de dados estruturados e não-estruturados, bem como complementou as oportunidades emergentes com os novos sistemas integrados de informação (*engaging, recording, insight*); exemplificou com casos do sistema financeiro, relevando os atributos Volume/Velocidade/Variabilidade/Veracidade dos dados (“4Vs”); ainda apontou desenvolvimentos pertinentes da segurança na *cloud*, a importância da dinamização ou promoção tecnológica, ou a implementação em curso de sistemas cognitivos.
- Luís Vaz (*Business Intelligence*, [Grupo Jerónimo Martins](#) ) apresentou uma abordagem pessoal ao BD e suas ferramentas, confrontando tal com uma abordagem empresarial que poderia ser conservadora; contextualizou algumas noções básicas de informação, palestrou sobre algumas *buzzwords*; confrontou a necessidade e paradigma de BD, direccionando tal para a competição analítica em termos empresariais; provocando a audiência com a sua abordagem dicotómica, considerou o atributo Valor como complementar aos “4Vs”, discutiu algumas dúvidas sobre o BD como tecnologia disruptiva; como principais desafios, apontou o balanço custo-benefício e caracterizou alguns dos seus atributos, a extração de valor para a empresa, a personalização de canais e esquemas de lealdade dirigidos ao cliente.



A discussão decorreu também com boa dinâmica, com os participantes e oradores interagindo positivamente, e sendo esmiuçados: *i)* os factores de sucesso em distribuição; *ii)* a avaliação do retorno de investimento, ROI; *iii)* o balanço entre produtividade e inovação, considerando-se a eficiência operacional; *iv)* a apreciação do valor extraído dos dados; *v)* os principais desafios à evolução da BD, seja em termos de recursos humanos nas empresas, de ferramentas tecnológicas na academia, ou da liderança por inovação como elemento disruptivo; *vi)* a definição de BD, o tipo de conhecimento e os processos respeitantes; *vii)* o BD no setor industrial e no setor da distribuição; *viii)* possíveis sinergias entre dados empresariais, tecnologias inovadoras e desenvolvimentos académicos.

Tendo sido clarificados, em momento prévio ao início destas jornadas, alguns pontos relacionados com a disseminação dos materiais e com a recolha de imagens, os diapositivos das apresentações e materiais complementares foram temporariamente disponibilizados em [http://tiny.cc/IOBD\\_Sep14](http://tiny.cc/IOBD_Sep14).

### c) IBM HADOOP

A jornada de 10-setembro dirigida a **ferramentas IBM HADOOP** decorreu com a supervisão e monitorização de Paul Reeves (IBM France), sendo a língua de trabalho o inglês. A sessão iniciou-se com a apresentação do formador e de todos os participantes, das expectativas e necessidades de cada um: a audiência era constituída quase na totalidade por docentes e investigadores, cada qual com a sua especialidade. Note-se que a audiência típica destas ações de formação é habitualmente constituída por estudantes avançados e das áreas de computação e informática.

Por conseguinte, o formador doseou os temas de enquadramento e motivação com as componentes técnicas e laboratoriais, de forma a promover um melhor acompanhamento dos diferentes tópicos: noções introdutórias; *i)* *Hadoop* e sua arquitetura; *ii)* *MapReduce*; *iii)* *Pig*, *Hive* and *Jaql*; *iv)* *Flume*; *v)* *Big Insights*.



## **IO & Big Data**

Note-se que as ferramentas BD foram inicialmente desenvolvidas por encomenda do governo norte-americano na sequência dos eventos de Setembro-2001, no sentido de permitir o tratamento automático de um elevado volume de dados, sendo estes dados de diversa natureza (numérica, texto, audio, imagem,...), com estruturas variadas; analogamente ao que ocorreu com a Programação Linear nos anos cinquenta do século passado, a disponibilização de tais ferramentas no domínio público e empresarial parece prometedora.

Mais informação sobre esta sessão de 10 de Setembro de 2014, sobre a agenda, as inscrições e os pré-requisitos, foram disponibilizadas oportunamente através de meios electrónicos e na rede social FB, estando os materiais provisoriamente disponíveis em:

<https://www.dropbox.com/sh/z827qjr04wedctt/AAAnXuM03o9jYEpqWe47rJb3a?dl=0>.

### **IO&BD-2014 (II)**

#### ***Workshops IBM Big Data***

#### **Portalegre, 8/9 de outubro-2014**

A segunda série de workshops, visando a formação específica em aplicações laboratoriais “Hadoop” e “Big Data” com ferramentas IBM, decorreu nos dias 8 e 9 de outubro de 2014. Esta acção foi frequentada maioritariamente por docentes e investigadores do IPP, estando também presentes alguns profissionais e estudantes interessados em desenvolvimentos práticos e aplicações *Big Data*, num total de mais de duas dezenas de participantes.

Atendendo ao carácter laboratorial (“hands-on”) desta série, foram oportunamente recomendados conhecimentos básicos de Linux como pré-requisitos, tendo-se também convidado os participantes a trazer algum equipamento (*laptop*, disco externo) dado o número de máquinas disponíveis ser limitado.

O docente Jorge Machado (ESTG/IPP) preparou e apresentou uma sessão de normalização visando competências básicas em Linux. Esta sessão com a duração de 4 (quatro) horas decorreu na manhã de 6 de outubro, sendo frequentada pela quase totalidade dos elementos da ESTG que participaram nestas atividades de formação, bem como por alguns funcionários de diversos centros de informática do IPP.



## **IO & Big Data**

As sessões de 8/9 de outubro decorreram em inglês, com a apresentação e monitorização de Paul Reeves (IBM France), sendo também possível a clarificação de alguns tópicos particulares dos participantes. Os participantes que se registaram como membros ou parceiros de grupos IBM (*IBM Partner World*, *IBM Academic Initiative*) também usufruíram de condições especiais de inscrição.

Mais informações sobre estas sessões foram disponibilizadas através de meios electrónicos e na rede social FB, estando os materiais provisoriamente disponíveis em:

[https://www.dropbox.com/sh/twglprk9mpglkb8/AAC4UD\\_TcC9rat-sxgDOJvoVa?dl=0](https://www.dropbox.com/sh/twglprk9mpglkb8/AAC4UD_TcC9rat-sxgDOJvoVa?dl=0) .

### **IO2015: “IO & Big Data”**

Considerando o tema “IO& Big Data”, está também calendarizado o **IO2015-XVII Congresso da APDIO** para 07-09 de Setembro-2015, na ESTG/IPP, sendo este o ponto de encontro privilegiado para a promoção da IO e tópicos associados, através da troca de ideias entre docentes, investigadores, estudantes, e profissionais com diferentes formações de base.

O início de inscrições para o **IO2015** é esperado em breve, estando informação adicional disponível em <http://tiny.cc/iobd-1415> e agradecendo-se consulta regular das atualizações.

#### **Inscrição e Contactos :**

Instituto Politécnico de Portalegre  
Gabinete de Relações Públicas e Cooperação  
Susana Dias  
Praça do Município  
Apartado 84  
7301 – 901 Portalegre  
Telefone – 245 301 500 / 245 301 545  
Fax – 245 330 353  
[formacao@ippportalegre.pt](mailto:formacao@ippportalegre.pt)



**Ligações frequentemente utilizadas:**

- Sítio “IO&BD-2014” <http://tinyurl.com/iobd2014>
- Sítio “IO&BD-1415” (FB) <http://tiny.cc/iobd-1415>
- 09-Set-2014 (FB) <https://www.facebook.com/events/1424186301193954/?ref=22&source=1>  
(materiais) [http://tiny.cc/IOBD\\_Sep14](http://tiny.cc/IOBD_Sep14)
- 10-Set-2014 (FB) <https://www.facebook.com/events/1478745862367749/>  
(materiais) <https://www.dropbox.com/sh/z827qjr04wedctt/AAAnXuM03o9jYEpgWe47rJb3a?dl=0>
- 08/09-Outubro-2014 (FB) <https://www.facebook.com/events/725221090892092/>  
(materiais) [https://www.dropbox.com/sh/twglprk9mpglkb8/AAC4UD\\_TcC9rat-sxgDOJvoVa?dl=0](https://www.dropbox.com/sh/twglprk9mpglkb8/AAC4UD_TcC9rat-sxgDOJvoVa?dl=0)
- Alojamento (FB) <https://www.facebook.com/events/282254538649275/?ref=5>
- Atividades culturais (FB) <https://www.facebook.com/events/335324536634397/?ref=5>

\*