INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

UNIVERSIDADE DE ÉVORA
JOCLAD2016
31 MAR a 2 ABR 2016
XXIII JORNADAS DE CLASSIFICAÇÃO E ANÁLISE DE DADOS

Transmissão Automática de Dados para o INE

JOCLAD 2016

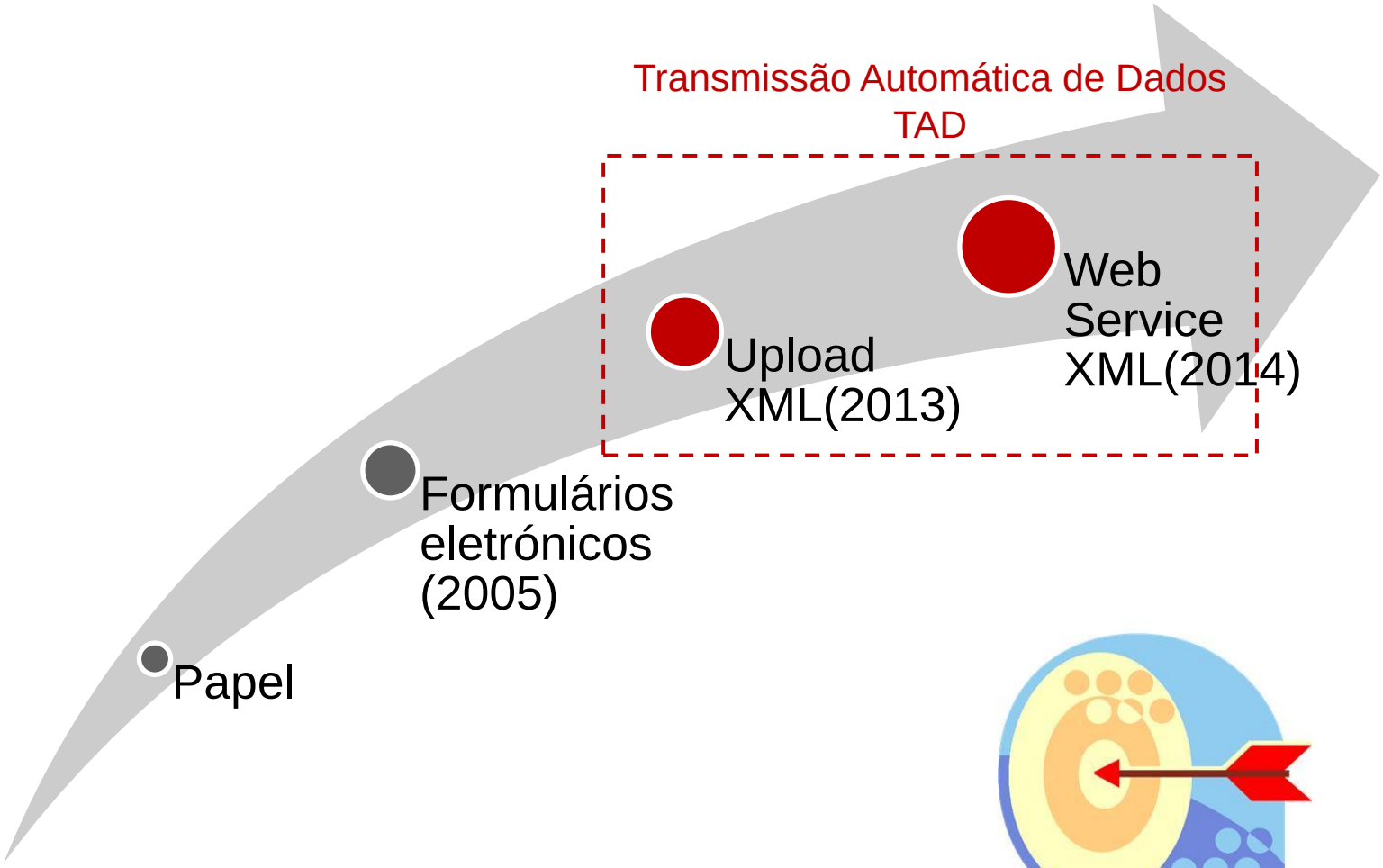
Luísa Pereira

Évora – 1 de abril de 2016

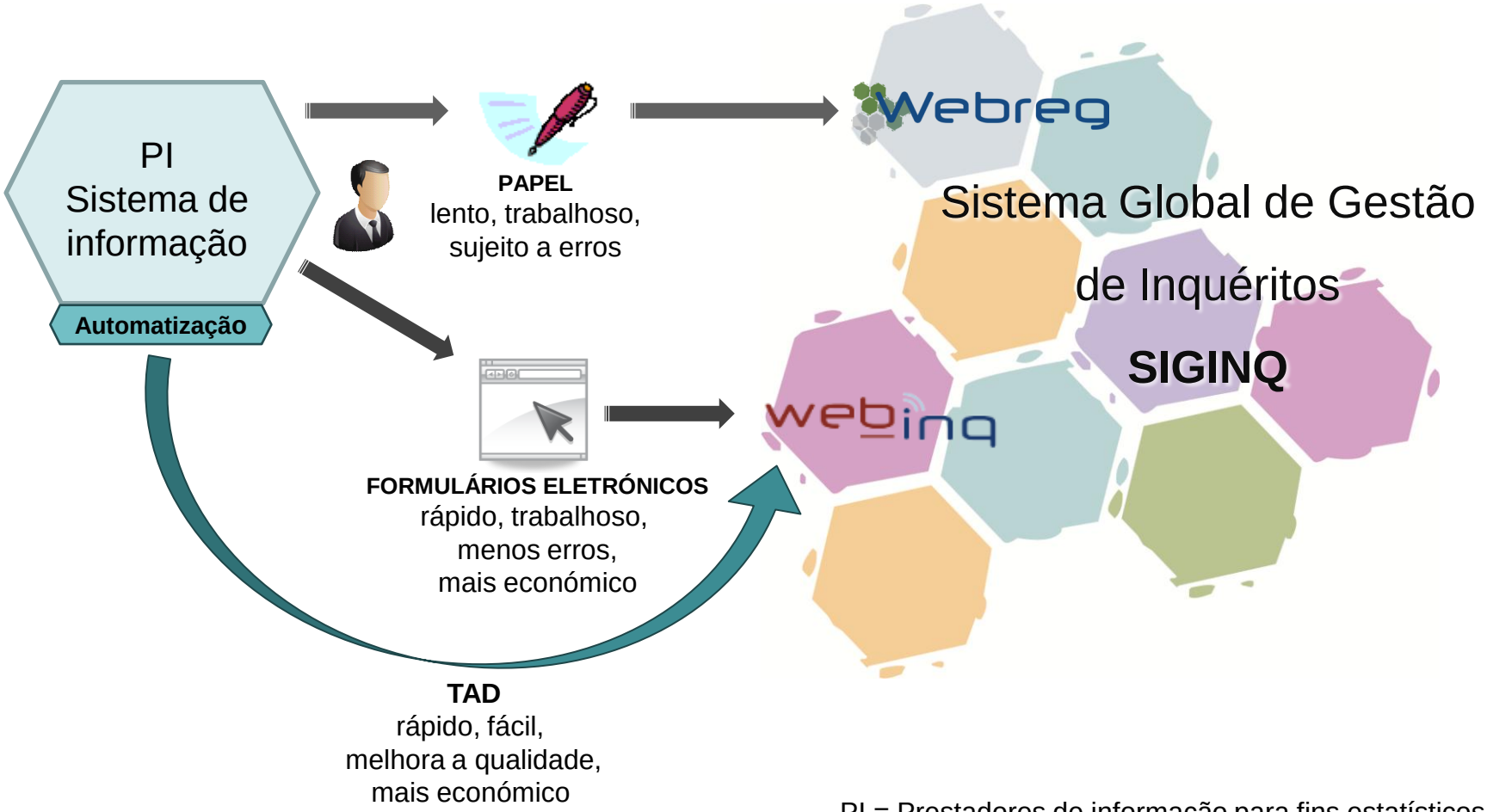
Transmissão Automática de Dados
Índice

- Visão geral - Empresas
- Responsabilidade partilhada
- Escolha dos candidatos ideais
- Razões e expectativas
- Questões previsíveis
- Porquê XML/XSD?
- Etapas
  - Etapa I
  - Etapa II
- Primeira escolha
- Algumas contagens
- Opinião das empresas

Transmissão Automática de Dados
Visão geral – Empresas



Transmissão Automática de Dados
Visão geral – Empresas



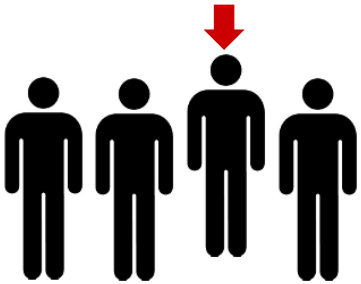
PI = Prestadores de informação para fins estatísticos

Transmissão Automática de Dados
Responsabilidade partilhada

- A TAD não é para todos
  - Não substitui o Papel nem os Formulários eletrónicos
  - Continuarão a ser a melhor opção para alguns PI
- Pensado para o público certo
  - Focalizado nos PI que mais beneficiem
    - Solução simples e com boa performance
    - Curva pequena de aprendizagem
    - Baixo custo de desenvolvimento
- A chave é a comunicação e a diplomacia
  - A cooperação dos PI é vital
  - Garantir que os PI percebem os seus benefícios
  - Demonstrar total disponibilidade do INE
  - Sem os PI, a TAD não será usada, logo será inútil

PI = Prestadores de informação para fins estatísticos

Transmissão Automática de Dados
Escolha dos candidatos ideais



- Recomendada para inquéritos
  - Sem alterações frequentes
    - Questionário
    - Amostra
  - Baixa periodicidade (semanal, mensal, ...)
  - Em que os PI usam soluções informacionais/informáticas
    - *Software houses* ficarão satisfeitas por poderem proporcionar novas funcionalidades
    - Uma única implementação numa ferramenta de uma *software house* permite a sua adoção em vários PI
- Recomendado para PI (Prestadores de informação)
  - Com sistemas de informação/informáticos
  - Dispostos a investir na redução do custo da resposta a inquéritos
  - Com grandes volumes de dados
  - Com carga estatística alta

Transmissão Automática de Dados
Razões e expectativas

- Reduzir a carga atual de recolha
  - A entrega de dados (submissão) deve exigir menos esforço e ser mais pontual
- Reduzir erros na entrada de dados
  - As respostas têm maior garantia de qualidade e são incorporadas automaticamente
- Reduzir o trabalho da equipa de recolha de dados
  - Eliminado o trabalho de introdução de dados
  - Mais disponibilidade para a análise de dados
- Os erros são expectáveis (e corrigidos) durante o período de teste
  - A automatização viabiliza a repetição com qualidade
  - Previne erros
    - Tipo de dados
    - Listas de códigos inválidos
    - Más interpretações

Transmissão Automática de Dados
Razões e expectativas

- Reduzir os custos operacionais globais
  - Preparado o sistema, o custo operacional é menor para o prestador e para o recetor (INE) de informação
- Adotar métodos reutilizáveis
  - Desenho genérico
  - Integrado no sistema do INE (SIGINQ)
  - Permite o uso futuro em inquéritos de outros domínios
  - Menor custo para inquéritos de outros domínios
- Minimizar o impacto na arquitetura atual
  - Minimizado o impacto no complexo sistema de recolha existente
  - Foram escolhidos cautelosamente pontos de integração

Transmissão Automática de Dados
Questões previsíveis

- Automatização pode levar os PI a descurar a avaliação dos dados
  - Diminuição da qualidade
- Questões abertas, tais como justificações, podem ser esquecidas
  - Resposta não existe nos sistemas de informação dos PI
- Dados são válidos perante regras automáticas
  - Não são analisados manualmente
  - Podem existir erros não detetados inicialmente e difíceis de encontrar
- Novas entregas de dados são sempre totais
  - Podem conduzir ao desperdício de recursos de hardware ou até a problemas de coerência

Transmissão Automática de Dados
Porquê XML/XSD?

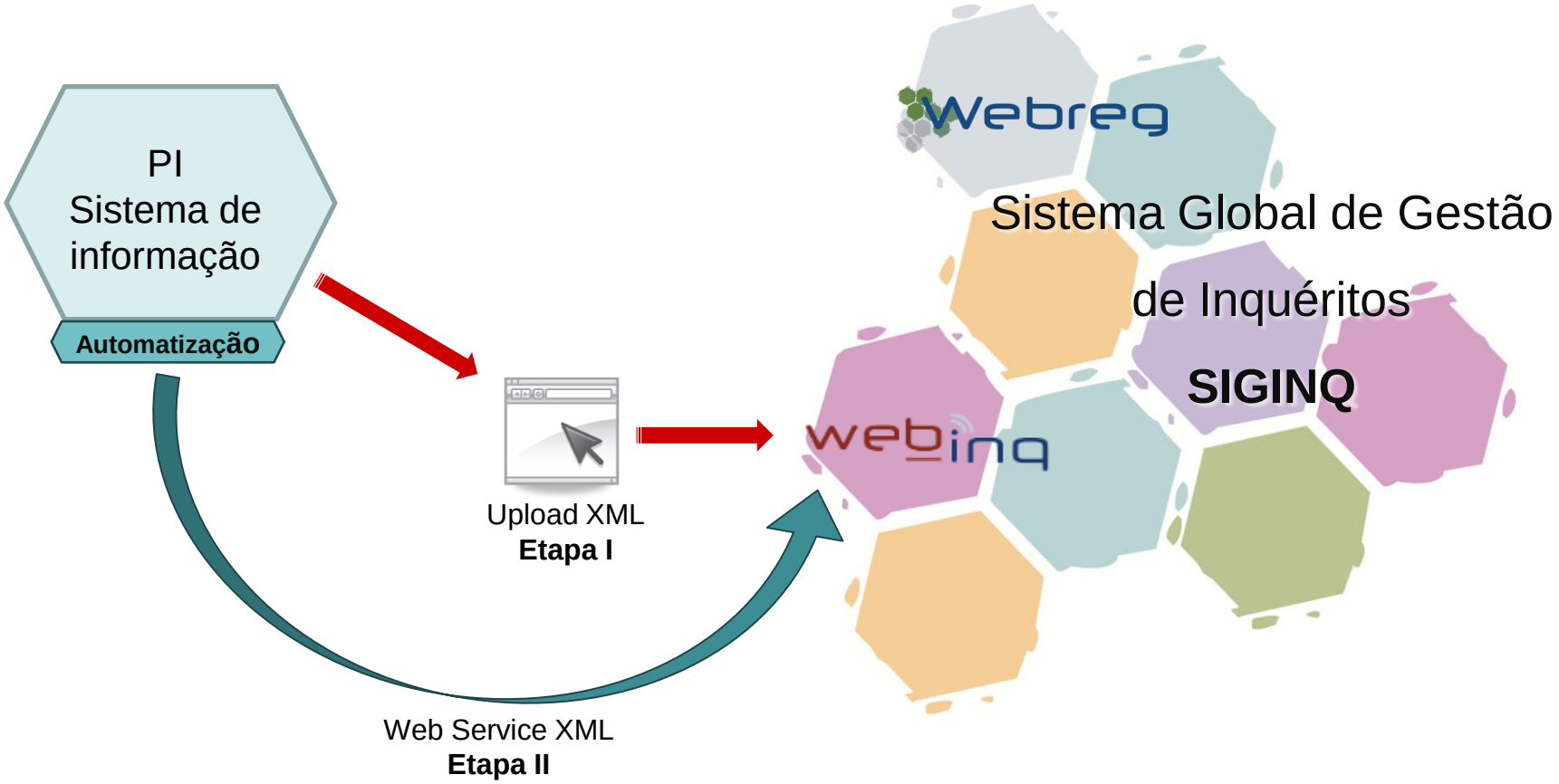
- Independência de plataforma
  - Ficheiro de texto XML (*Extensible Markup Language*)
  - Futuro garantido
- Fácil
  - Ficheiros de texto XML são fáceis de produzir e de processar
- Complexidade ilimitada
  - Estrutura do ficheiro suporta complexidade ilimitada
  - Definição de forma a suportar
    - Inquéritos múltiplos
    - Ocorrências múltiplas
    - Unidade estatísticas múltiplas
- Legível
  - Aplicações
  - Pessoas

Transmissão Automática de Dados
Porquê XML/XSD?

- XML simples – Fácil adoção
  - Formato XML simples reduz custo de implementação dos PI
  - Esquema similar à arquitetura interna, simplifica a implementação e integração

- XSD público – Validação antes da entrega
  - Disponível um ficheiro XSD (XML Schema Definition) que permite validar o formato dos dados
  - Permite uma validação inicial, fácil e básica, antes da entrega (submissão) da resposta

Transmissão Automática de Dados
Duas etapas



PI = Prestadores de informação para fins estatísticos

Transmissão Automática de Dados
Etapa I – Upload XML

- Desenvolvimento inicial (INE)
  - Conceber e desenvolver a importação de ficheiros XML de resposta a inquéritos
    - Guardar dados do ficheiro no repositório do INE
    - Dar feedback: reportar erros detetados ou uma mensagem de sucesso
- Teste e afinação (Prestadores de informação)
  - Disponibilizado às *software houses* um **website privado** (GUID de acesso)
    - Permite a submissão de respostas XML de teste para afinação da sua formatação e integridade interna
    - Antecipa e permite avaliar eventuais dificuldades
    - Documenta todo o processo: Especificações e perguntas frequentes
    - Suporte técnico
- Em uso (Prestadores de informação)
  - Na área privada e segura do **WebInq** (Inquéritos do INE na WEB), está disponível uma funcionalidade de upload

Transmissão Automática de Dados
Etapa I – Upload XML – swhupload.ine.pt

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

TAD
Transferência Automática de Dados

TAD - Transmissão automática de dados para o INE

O Webinq é um serviço disponível na Internet orientado para a recolha de informação por via eletrónica, promovendo a diminuição do esforço exigido às organizações para resposta aos questionários oficiais do INE, visando com esta facilidade um melhor relacionamento com as/os respondentes, ao criar processos que reduzam e agilizem o seu trabalho.

Em 2013, o INE introduziu neste serviço um novo método de recolha, TAD - Transmissão Automática de Dados, apenas para um subconjunto dos inquéritos disponíveis.

Atualmente, os seguintes Inquéritos aceitam este novo método de recolha:

- ICT
- IHSP
- IPCAMP
- IPCOL
- IPHH
- IREPEXP
- IREPIMP
- ITRP
- SIOUQ1
- SIOUQ2
- SIOUQ3
- SIOUQ4
- SIOUQ6
- UCDR

O TAD é disponibilizado às organizações de duas formas:

- Upload de ficheiros XML;
- Webservice para envio de ficheiros XML.

Nesta página, após introdução do GUID que identifica univocamente cada organização em ambiente de desenvolvimento e de produção, poderá ter acesso a:

- Upload de XML em ambiente de desenvolvimento – Com o objetivo de afinar a construção dos ficheiros XML;
- Documentação genérica da TAD;
- Documentação específica de cada inquérito que permite a TAD;
- Colocação de dúvidas.

Entrada

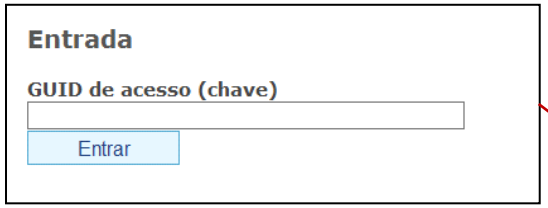
GUID de acesso (chave)

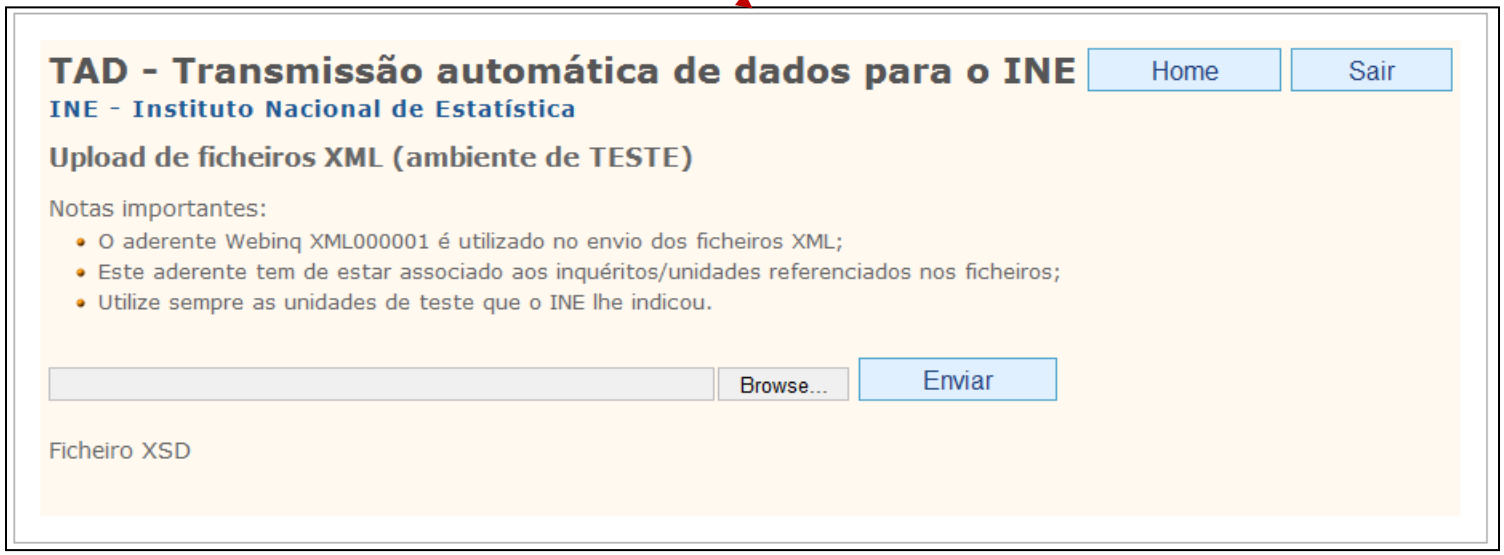
Entrar

Como aderir a este serviço?
Cada organização, a qual poderá ser uma empresa de software ou a própria empresa que responde ao INE deverá manifestar a sua intenção de participar neste novo método de recolha solicitando ao INE a criação de um GUID para acesso.

Este contacto deverá ser efetuado por mail para o endereço Webinq@ine.pt.

Transmissão Automática de Dados
Etapa I – Upload XML – swhupload.ine.pt





Transmissão Automática de Dados
Etapa I – Estrutura XML

```
<ine xmlns="http://ine.pt/public/xsd/forms" xmlns:xsi="http:...xsd">
  <header chave="INEFADM1-2783-4585-9000-2ED038CB5BC8">
    <geracao>
      <data> 2014-03-28T11:41:15 </data>
      <software ver="1.0"> SOFTWARE HOUSE </software>
    </geracao>
    <qual> REAL </qual>
  </header>
  <respostas>
    ▶ <resposta>
  </respostas>
</ine>
```

1º Nível

```
<resposta>
  <referencia>
    <inq> IPHH </inq>
    <estab> 0000000000002 </estab>
    <mensal>
      <ano> 2014 </ano>
      <sub-periodo> 10 </sub-periodo>
    </mensal>
  </referencia>
  ▶ <dados tipo="normal" ult_actualizacao="2014-08-28T11:4
</resposta>
```

2º Nível

Transmissão Automática de Dados
Etapa I – Estrutura XML

```
<resposta>
  <referencia>
    <inq> IPHH </inq>
    <estab> 0000000000002 </estab>
    <mensal>
      <ano> 2014 </ano>
      <sub-periodo> 10 </sub-periodo>
    </mensal>
  </referencia>
  <dados tipo="normal" ult_actualizaca
</resposta>
```

2º Nível

```
<dados tipo="normal" ult_actualizacao="2014-08-28T11:41:15">
  <v nome="BC101"> 55112 </v>
  <v nome="BC105"> 20 </v>
  <v nome="VPP_1"> 5 </v>
  <v nome="VPP_2"> 5 </v>
  <v nome="VPP_3"> 5 </v>
  <v nome="pais__LU____V1"> 6 </v>
  <v nome="pais__LU____V2"> 6 </v>
  <v nome="pais__LU____V3"> 6 </v>
  <v nome="pais__LU____pais_dsg">
    <![CDATA[ Luxemburgo ]]>
  </v>
</dados>
```

3º Nível

Transmissão Automática de Dados
Etapa I – Correspondência

Início | Quadro 1 | Quadro 2 | Notas

Permanência de hóspedes (durante o mês de referência)

País de residência habitual		Nº de Hóspedes que entraram no mês (1)	Nº de Hóspedes que dormiram no mês a) (2)	Nº de Dormidas (Noites) (3)
Portugueses residentes em Portugal	VPP	5	5	5
Estrangeiros residentes em Portugal	VEP			

Nota: Na linha anterior devem registar-se os Hóspedes de nacionalidade estrangeira que residem em Portugal. Os Hóspedes, que residem no estrangeiro, sejam portugueses ou estrangeiros, devem ser registados nas linhas seguintes, por país de residência habitual.

Residentes no Estrangeiro, por países de residência	Nº de Hóspedes que entraram no mês (V1)	Nº de Hóspedes que dormiram no mês a) (V2)	Nº de Dormidas (Noites) (V3)	Apagar
DE Alemanha	12	12	12	☐
LU Luxemburgo	6	6	6	☐
Sub-Total	18	18	18	

Outro país: Adicionar

```
<v nome="V162">200</v>
<v nome="V163">200</v>
<v nome="VPP_1">5</v>
<v nome="VPP_2">5</v>
<v nome="VPP_3">5</v>
<v nome="num_dias">31</v>
<v nome="pais__DE____V1">12</v>
<v nome="pais__DE____V2">12</v>
<v nome="pais__DE____V3">12</v>
<v nome="pais__DE____pais_dsg"><![CDATA[Alemanha]]></v>
<v nome="pais__LU____V1">6</v>
<v nome="pais__LU____V2">6</v>
<v nome="pais__LU____V3">6</v>
<v nome="pais__LU____pais_dsg"><![CDATA[Luxemburgo]]></v>
</dados>
```

Transmissão Automática de Dados
Etapa II – Web Service

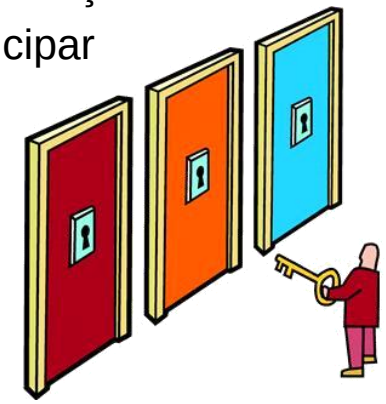
- Evolução
  - Sistemas externos interagem diretamente com o sistema do INE
- Web Service
  - Complementa as funcionalidades disponibilizadas pelo upload
  - Disponibiliza uma interface aplicacional (API) exposta como Web Service
- Processo simples e seguro
  - Comunicação em canal seguro (dados encriptados via certificado SSL)
  - Credenciais do aderente ao WebInq (código/password)
  - Chamado diretamente pelos sistemas externos
  - Executado em modo assíncrono
  - Resultado do processamento disponível no Webinq

Transmissão Automática de Dados
Primeira escolha

- Estatísticas do Turismo

- Inquérito à permanência de campistas nos parques de campismo
  - Ocorrências mensais
  - Questionário estável
  - Amostra estável (recenseamento)
  - Existem sistemas de informação/informáticos externos genéricos
  - Questionário solicita dados já presentes nos sistemas de informação dos PI
  - A equipa de recolha de dados demonstrou interesse em participar
  - O Conselho Diretivo do INE concordou

IPCAMP



Transmissão Automática de Dados
Algumas contagens

- TAD (2015)
  - Disponível para 14 inquéritos
  - 23,15 % dos dados recolhidos (22.282 respostas em 96.234 possíveis)
- TAD (2016)
  - Planeados mais 3 inquéritos
  - Previsto reforçar contactos com os PI e *software houses*
- WebInq (2015)
  - 95,5 % dos dados recolhidos (606.353 respostas em 634.929 possíveis)
  - Restantes 4.5 % recolhidos por outros métodos de recolha (papel, fax e e-mail)

Transmissão Automática de Dados
Opinião das empresas

- A “ORBITUR” referiu o enorme benefício de que passou a usufruir pela implementação da TAD no IPCAMP

- Recolha e reporte de informação estatística ao INE

126 h/ano para 8 h/ano

Workshop, promovido pelo INE sobre o tema
“A Produção Estatística Oficial e as Empresas” Outubro de 2013

Nota: ORBITUR - 21 Estabelecimentos

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

Transmissão Automática de Dados para o INE

JOCLAD 2016

luisa.pereira@ine.pt

Évora – 1 de abril de 2016

Obrigada!