



Jak instytucje publiczne mogą skorzystać z platformy CEF.AT?

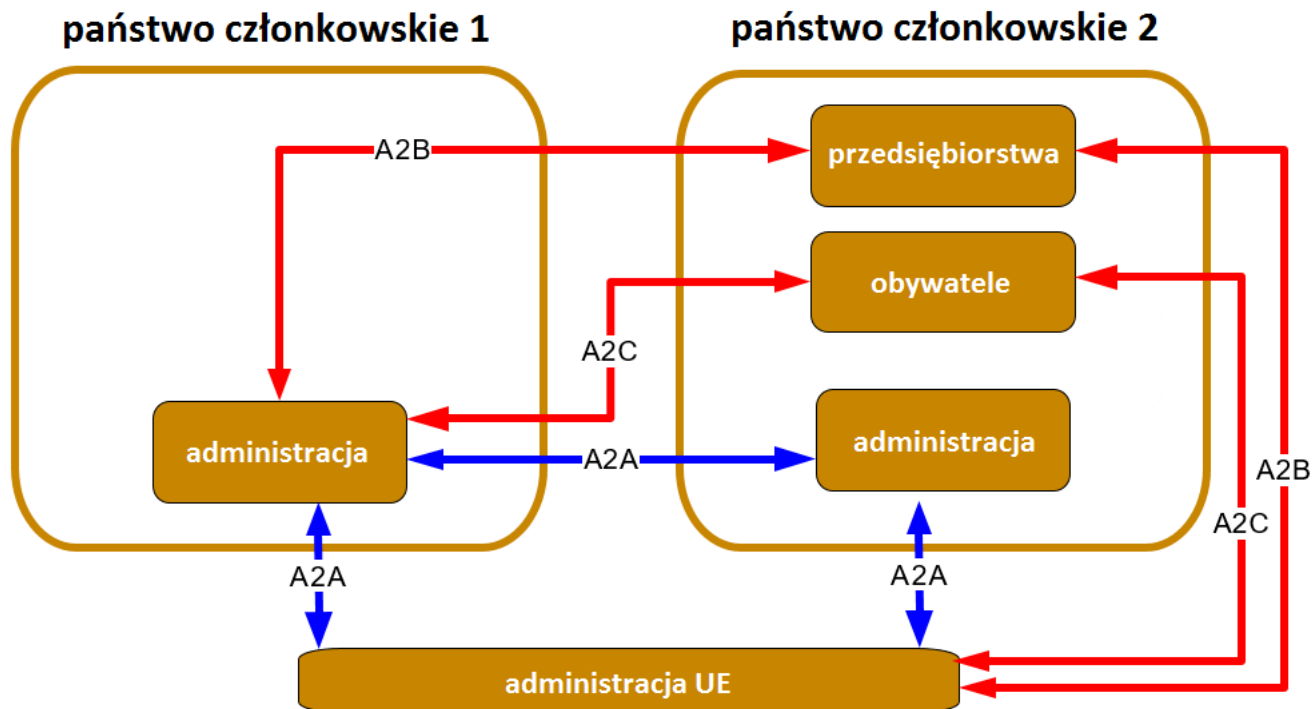
Szymon Klocek
Dyrekcja Generalna ds. Tłumaczeń Pisemnych
Komisja Europejska

Warszawa, 9 marca 2016 r.

Instytucje publiczne a tłumaczenie maszynowe

- *wielojęzyczność i tłumaczenie maszynowe*
- *obecnie: system MT@EC*
- *w przyszłości: platforma CEF.AT*
 - MT@EC = Machine Translation @ European Commission = system tłumaczenia maszynowego Komisji Europejskiej
 - CET.AT = Connecting Europe Facility – Automated Translation = instrument „Łącząc Europę” – tłumaczenie automatyczne

Interakcje pomiędzy podmiotami w państwach członkowskich i w UE



→ pierwszy rodzaj interakcji (A2B, A2C)

→ drugi rodzaj interakcji (A2A)

*Cel, do którego dążymy: czy nie byłoby wspaniałe, gdybym mógł **korzystać z usług publicznych w dowolnym państwie** członkowskim z **dowolnego miejsca** i **otrzymywać informacje w moim języku ojczystym?***



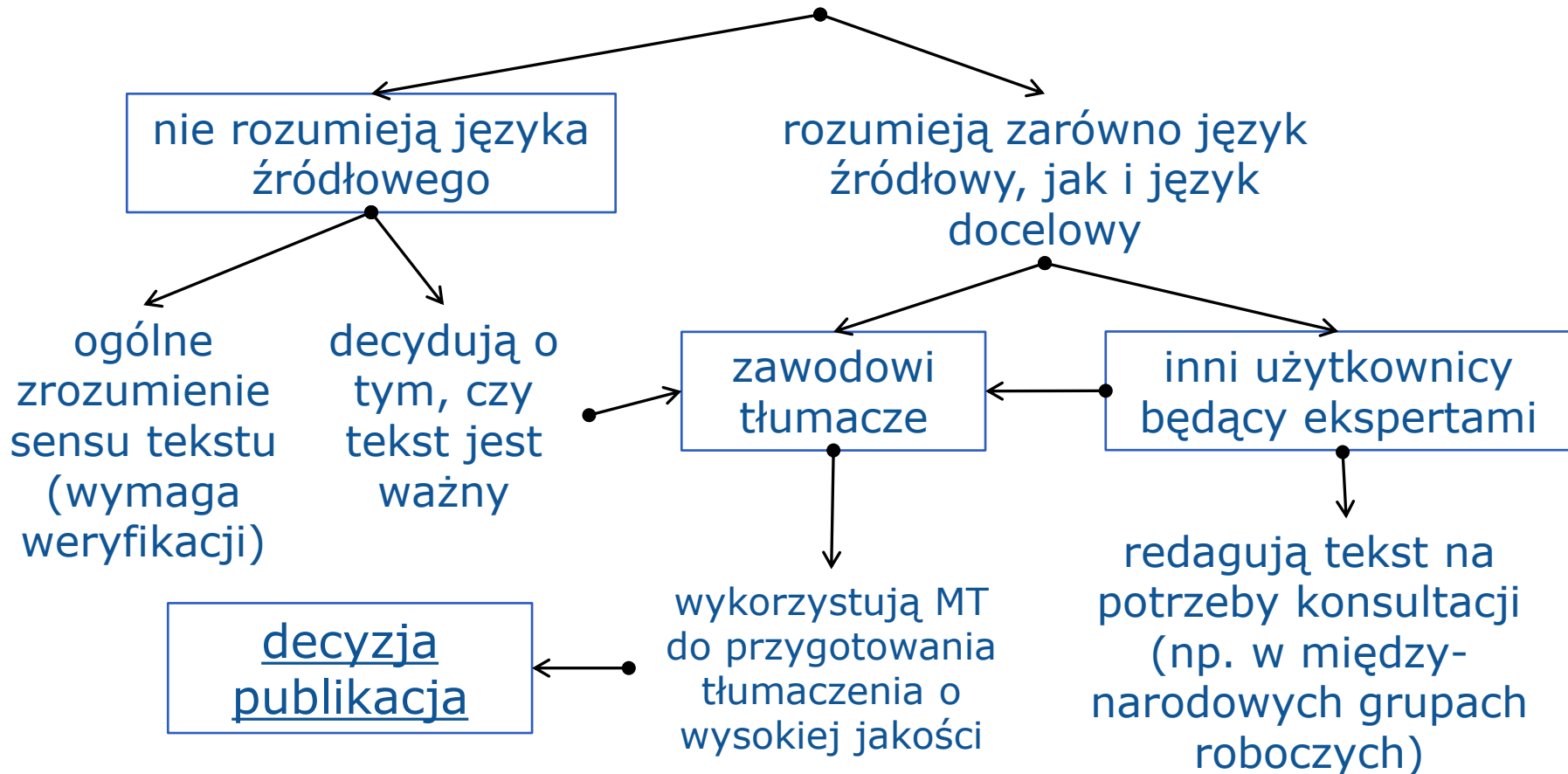
Dlaczego tłumaczenie maszynowe jest ważne?

MT jest jedynym dostępnym rozwiązaniem technicznym zapewniającym:

- szybki i tani **dostęp do informacji** w językach obcych
- zrozumienie **informacji** w językach obcych, których w przeciwnym wypadku nie dałoby się wykorzystać lub których przetłumaczenie wymagałoby czasu i znacznych wydatków
- korzystanie **ze stron internetowych w wielu językach**
- **wyszukiwanie i analizę** informacji dostępnych w różnych językach

Właśnie dlatego tłumaczenie maszynowe jest technologią **o kluczowym znaczeniu dla wielojęzycznej Europy.**

Użytkownicy tłumaczenia maszynowego



MT@EC jest już dostępne!

Uruchomienie systemu: 26 czerwca 2013 r.

Języki: wszystkie kombinacje w ramach 24 oficjalnych języków UE

Wykorzystywana technologia: statystyczne tłumaczenie maszynowe

wykorzystujące otwartoźródłowe oprogramowanie o nazwie „Moses”, współfinansowane przez programy ramowe UE w zakresie badań naukowych i innowacji

System MT@EC został zbudowany przez DGT w latach 2010-2013

i był współfinansowany z programu ISA (działanie 2.8)

http://ec.europa.eu/isa/actions/02-interoperability-architecture/2-8action_en.htm

Charakterystyka MT@EC

Odbiór tłumaczeń:

- ręczny (przez stronę internetową, przeznaczony dla ludzi)
- automatyczny (przez usługę sieciową, przeznaczony dla komputerów)

Cechy specjalne:

- **interfejs użytkownika** dostępny w 24 językach
- system **zachowuje formatowanie** dokumentów *[poza dokumentami .PDF]*
- specjalne **formaty docelowe** dla tłumaczy: .tmx i .xliff
- gotowe tłumaczenia można także otrzymywać **e-mailem**
- można przetłumaczyć **wiele dokumentów** na **wiele języków** jednocześnie
- informacja o **jakości tłumaczenia** w danej parze językowej (na podstawie metryk BLEU)

MT@EC a bezpieczeństwo

- Bezpieczne przechowywanie danych **w centrum danych UE**
- Dostęp poprzez **ECAS** (EC Authentication Service – system uwierzytelniania Komisji Europejskiej)
- Bezpieczne przesyłanie dokumentów:
 - w ramach **sTESTA***, bardzo bezpiecznej sieci prywatnej łączącej administracje publiczne w UE, oddzielonej od Internetu
 - przez **Internet** (z wykorzystaniem bezpiecznego połączenia przez protokół **https**)

** możecie Państwo sprawdzić czy Państwa organizacja ma dostęp do sieci TESTA na stronie:*

<https://portal.testa.eu/jetspeed/portal/homepage/about.psml>.



Tłumaczenie dokumentów

Tłumaczenie tekstu

Moje zlecenia tłumaczenia

Moje ustawienia

polski

Wyloguj

Wybierz co najmniej jeden oryginał:

Wybierz pliki...

LUB

Przecignij pliki tutaj.

Tłumacz z...

Przetłumacz na...

Format przetłumaczonego pliku:

Taki sam jak oryginał

TMX

XLIFF

Chcę otrzymać tłumaczenie e-mailem.

Usunąć po pobraniu.

Przetłumacz dokument

Wybrane pliki



Tłumaczenie dokumentów

Tłumaczenie tekstu

Moje zlecenia tłumaczenia

Moje ustawienia

polski ▼

Wyloguj

Tekst do przetłumaczenia:

Maksymalna długość: 4000 znaków.
Pozostało znaków: 4000 znaków.

Wyczyść

Tłumacz z...

Przetłumacz na...

- Chcę otrzymać tłumaczenie e-mailem.
 Usuń po pobraniu.

Przetłumaczony tekst:

Zaznacz wszystko

Przetłumacz tekst

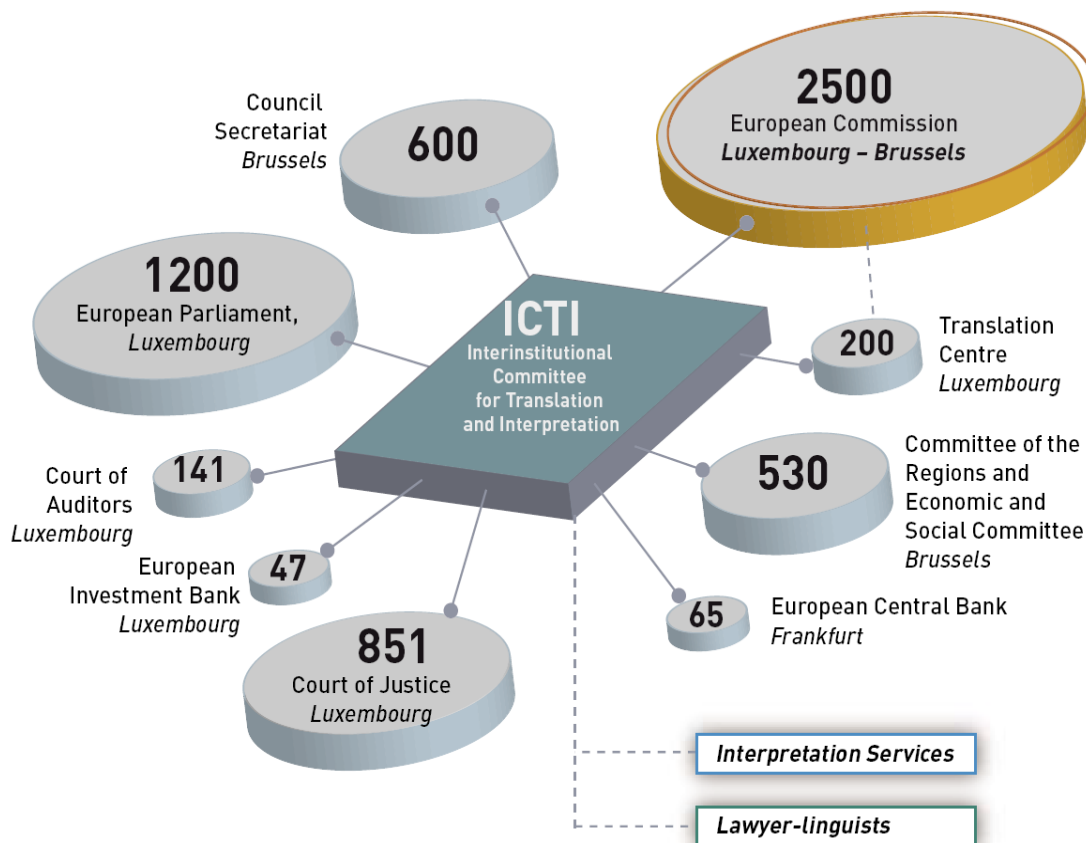
Kto może obecnie używać MT@EC?

→ *Instytucje i organy UE:*

- | | |
|---|--|
| <ul style="list-style-type: none">▪ Komisja▪ Parlament▪ Rada▪ Trybunał Sprawiedliwości▪ Trybunał Obrachunkowy | <ul style="list-style-type: none">▪ Europejski Komitet Ekonomiczno-Społeczny▪ Komitet Regionów▪ Europejski Bank Centralny▪ Europejski Bank Inwestycyjny |
|---|--|

→ *usługi internetowe finansowane lub wspierane przez UE*

→ *administracje publiczne państw członkowskich UE/EOG*



Euramis

940 milionów zdań
[na koniec 2015 r.]

2,6 miliona nowych
zdań co miesiąc

wszystkie 24 języki
[różna ilość danych]

„dokumenty UE”
[tematyka i język typowe dla UE]

Darmowy dostęp do MT@EC dla pracowników administracji publicznej

- korzystając ze służbowego adresu e-mail, można samodzielnie zarejestrować się i otrzymać indywidualną nazwę użytkownika i hasło w ramach **ECAS**
- *[proszę skorzystać z instrukcji zawartych na stronie <https://webgate.ec.europa.eu/cas/eim/external/register.cgi>]*
- następnie należy wysłać **e-mail** na adres DGT-MT@ec.europa.eu z prośbą o aktywację dostępu do usługi
- DGT poinformuje Państwa o zakończeniu procesu aktywacji e-mailem



Usługi internetowe **podłączone do MT@EC**

(stan: styczeń 2016 r.)

Usługa	Opis/adres URL/wykorzystanie MT@EC
IMI	system wymiany informacji na rynku wewnętrznym (Internal Market Information System) – sieć wewnętrzna (intranet) łącząca ponad 7000 organów administracji publicznych państw członkowskich UE http://ec.europa.eu/internal_market/imi-net/index_en.html → wymiana wiadomości tekstowych
SOLVIT	internetowa sieć umożliwiająca rozwiązywanie problemów dotyczących niewłaściwego stosowania prawa rynku wewnętrznego przez organy publiczne http://ec.europa.eu/solvit/ → wymiana wiadomości tekstowych
nLex	wspólny portal prawa krajowego http://eur-lex.europa.eu/n-lex/ → tłumaczenie wyników wyszukiwań
TED	Tenders Electronic Daily (elektroniczny dziennik o przetargach) to internetowa wersja Suplementu do Dziennika Urzędowego UE dotyczącego zamówień publicznych http://ted.europa.eu/ → informacje o zawartych umowach
CircaBC	Centrum zasobów dotyczących komunikacji i informacji dla organów administracji, przedsiębiorstw i obywateli (Communication and Information Resource Centre for Administrations, Businesses and Citizens) (platformy współpracy dla użytkowników) https://circabc.europa.eu/ → tłumaczenie dokumentów

Co dalej?

przekształcimy system MT@EC w platformę CEF.AT

Platforma **CEF.AT**:

- zostanie zbudowana w oparciu o **istniejący** system MT@EC – ale będzie oferowała znacznie więcej
- będzie kładła nacisk na **bezpieczeństwo, jakość i dostosowanie** MT do potrzeb ogólnoeuropejskich usług internetowych – ale będzie obejmowała także inne podmioty
- będzie służyła **promowaniu wielojęzyczności** – nie tylko poprzez tłumaczenie maszynowe

Jakie korzyści zaoferuje użytkownikom CEF.AT?

- *szybsze i bardziej bezpieczne tłumaczenia*
- *więcej ogólnoeuropejskich internetowych usług publicznych wykorzystujących tłumaczenie maszynowe i inne formy wielojęzyczności, dostosowane do potrzeb użytkowników*

oraz

- *lepszą ogólną jakość tłumaczenia*
- *wyspecjalizowane silniki tłumaczące (na początku oferujące tłumaczenia w dziedzinach najbardziej interesujących dla użytkowników)*
- ➔ *Potrzeba **więcej danych lepszej jakości** we wszystkich językach oraz dotyczących szerszej tematyki (nowych rodzajów i kategorii tekstów)*

Dostępne obecnie: silniki tłumaczące „ogólnego przeznaczenia” (unijne)

- *MT@EC* zbudowano w oparciu o „tłumaczenia unijne” (zawartość bazy danych pamięci tłumaczeniowych Euramis)
- Euramis zawiera dane we wszystkich 24 językach oficjalnych UE **ale** w różnych ilościach (od 97 milionów segmentów w języku angielskim i 58 milionów w języku francuskim do 2,8 miliona w językach chorwackim i irlandzkim – pozostałe języki mają od 22 do 38 milionów segmentów)
- Euramis (a zatem i *MT@EC*) zawiera dane (segmenty) dotyczące działalności UE i tematów, którymi zajmuje się UE, napisane w języku, jakim posługuje się UE, **ale** nie zawiera dużych ilości danych napisanych w języku potocznym
- *MT@EC* umożliwia tłumaczenie tekstów z wielu różnych dziedzin **ale** nie ma obecnie możliwości wybrania jednej z dziedzin.

Potrzeba także więcej danych (segmentów) zawierających fragmenty specjalistycznych tekstów (dane z „konkretnych” dziedzin)

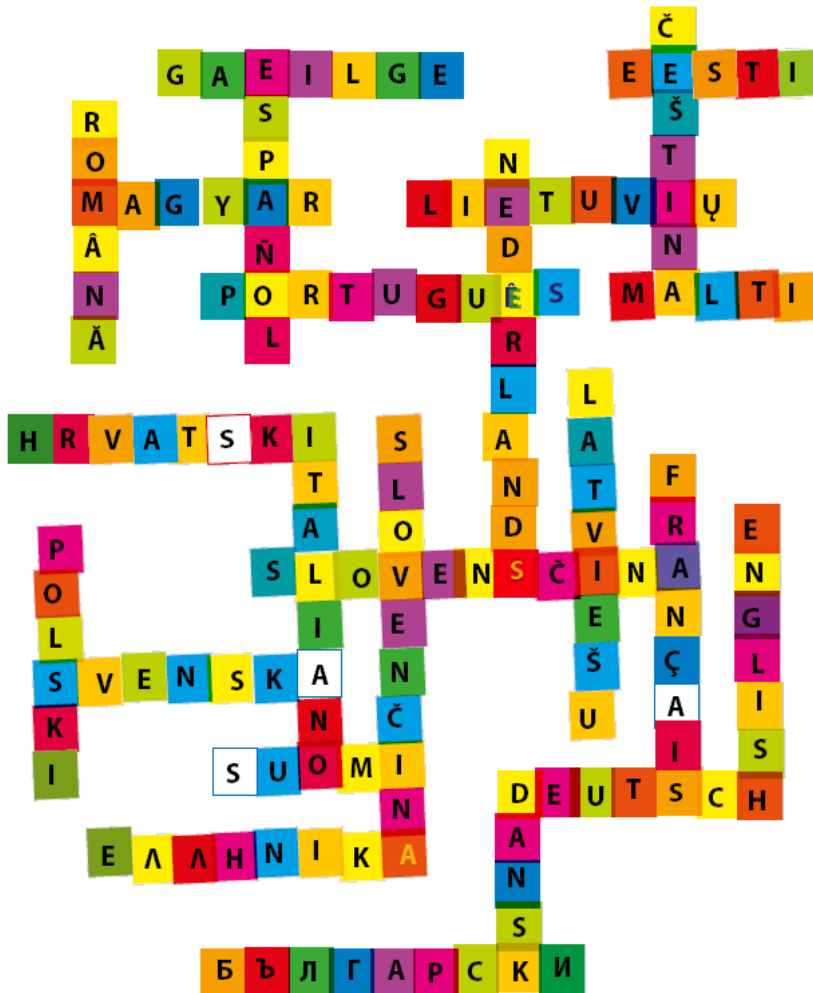
Usługi internetowe korzystające ze „zindywidualizowanych” silników tłumaczących które będą podłączone do platformy CEF.AT (a na razie do MT@EC)

Usługa	Opis/adres URL/wykorzystanie MT@EC
Europeana	Europejska biblioteka cyfrowa, oferująca scentralizowany dostęp do zdigitalizowanych treści w wielu językach, wchodzących w skład europejskiego dziedzictwa kulturowego.
ODP*	Ogónoeuropejski Portal Otwartych Danych (O pen D ata P ortal), umożliwiający dostęp do otwartych danych przechowywanych przez UE i państwa członkowskie.
EESSI	System elektronicznej wymiany informacji dotyczących zabezpieczenia społecznego (E lectronic E xchange of S ocial S ecurity I nformation), łączący 32 państwa EU i EFTA.
ODR*	Internetowe rozstrzyganie sporów (O nline D ispute R esolution), platforma umożliwiająca rozstrzyganie sporów dotyczących umów między konsumentami i sprzedawcami, łącząca wszystkie krajowe podmioty zajmujące się pozasądowym rozstrzyganiem sporów.
e-justice*	Portal umożliwiający scentralizowany dostęp do aktów prawnych, umożliwiający współpracę sądową w ramach UE.
SaferInternet	Usługi umożliwiające zwiększenie bezpieczeństwa dzieci w Internecie.

* usługi już podłączone do systemu MT@EC (oferującego „standardowe” tłumaczenia)



European
Commission



Pytania na temat MT@EC
proszę kierować na adres:
dgt-mt@ec.europa.eu