Criptografia

Engenharia Biomédica

J. Bacelar Almeida
jba@di.uminho.pt

Departamento de Informática
Universidade do Minho

2008/2009

O que é a Criptografia?

- Historicamente, a **Criptografia** consiste em técnicas que procuram tornar possível a comunicação secreta entre dois agentes, sobre um canal aberto.
- Por extensão, a criptografia hoje procura dar resposta a numerosas propriedades do processo de comunicativo para além da *confidencialidade* (segredo), como sejam a *integridade, autenticidade, anonimato, etc.*. Iremos designar essas propriedades genericamente por **propriedades de segurança**.
- Em contrapartida, a **Cripto-análise** tenta gorar os objectivos da Criptografia, isto é, *quebrar* a segurança da comunicação.
- Conjuntamente, a Criptografia e a Cripto-análise formam uma área a que podemos chamar **Criptologia**; uma área com profundas raízes na Matemática e nas Ciências da Computação.

Criptografia Moderna

- A Criptografia existe desde a antiguidade, normalmente associada a actividades militares e diplomáticas.
- A segurança dependia, em grande parte, do secretismo que rodeava as técnicas utilizadas (o que, historicamente, se revelou "catastrófico").
- Esta tendência fez-se notar ainda no Século XX, durante as 1ª e 2ª Guerras Mundiais e prolongou-se durante as primeiras décadas da Guerra Fria.
- Só no princípio dos anos 70 surgiu como área de investigação académica de conhecimento generalizado.
- Hoje é reconhecida a importância de eliminar o obscurantismo como factor na segurança dos sistemas criptográficos.

História da Criptografia Moderna

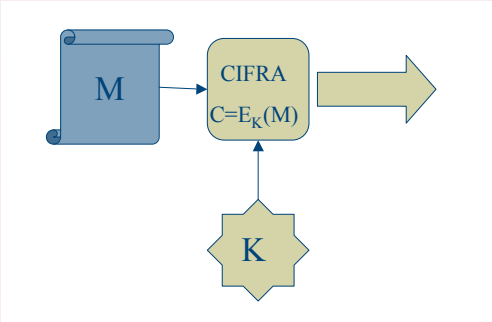
- **1948-1949** – *Claude Shannon* desenvolve a *Teoria da Informação* e enquadra o estudo da Criptografia nessa teoria.
- **1970-1977** – Desenvolvimento e estandardização do Data Encryption Standard (DES).
- **1976** – Primeiro paper de Diffie e Hellmann definindo os princípios da criptografia de chave pública.
- **1978** – Rivest, Shamir e Adleman descobrem a primeira cifra assimétrica: o RSA.
- **1985** – Descoberta da cifra assimétrica El Gamal.
- **1995** – Estandardização do Digital Signature Algorithm.
- **2001** – Escolha do substituto do DES: Advanced Encryption Standard (AES).

Princípio de Kerckhoff

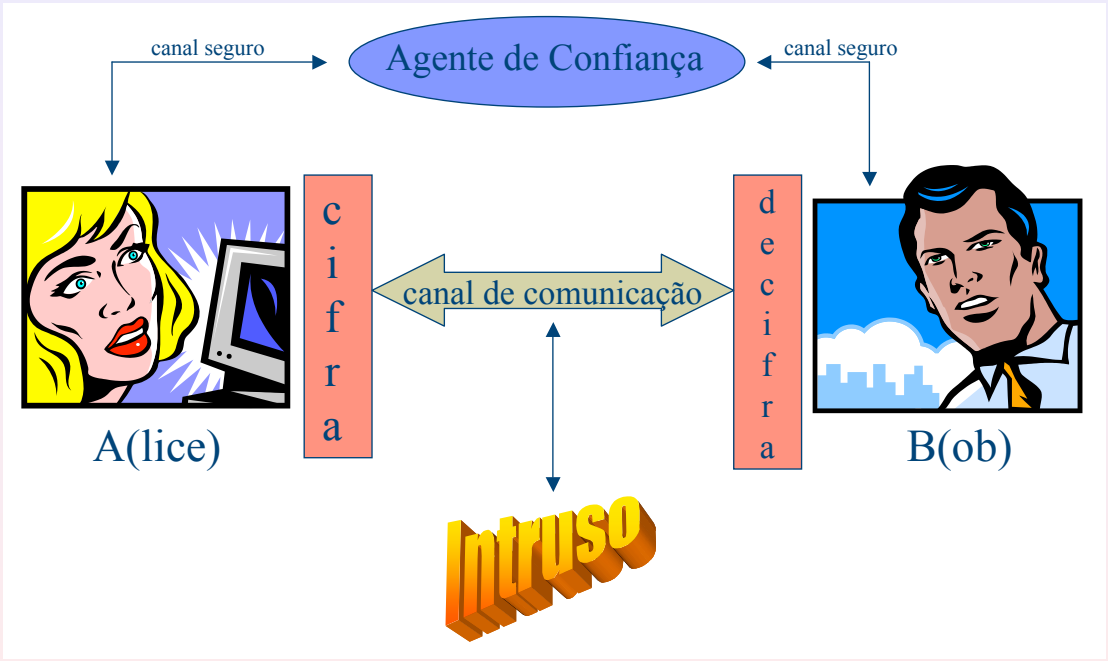
Para avaliar a segurança de uma técnica criptográfica devemos assumir que esta é do conhecimento de eventuais inimigos.

Corolário: a segurança da cifra é assegurada por um parâmetro explícito — a **chave**.

$$C = \mathrm{enc}_K(M)$$



Modelo



Terminologia

- **texto limpo:** mensagem a transmitir.
- **cifra:** operação que transforma o texto limpo numa mensagem "com significado obscurecido" — o **criptograma**.
- **chave:** parâmetro de segurança da operação de cifra
- **sistema criptográfico:** especificação das operações de "inicialização"; "cifra" e "decifragem".
- **ataque:** comprometimento dos objectivos da técnica criptográfica (e.g. obtenção do texto limpo sem conhecimento da chave; descobrir a chave utilizada; etc.).
- **intruso/adversário/inimigo/*spy*:** entidade que personifica quem pretende comprometer os objectivos da técnica criptográfica.

Segurança de Cifras

Dependendo das capacidades computacionais atribuídas ao adversário, classificam-se as noções de segurança das cifras como:

- **Segurança Absoluta** – quando a segurança da cifra é estabelecida perante um adversário sem limitações computacionais.
- **Segurança Computacional** – quando se considera que o adversário dispõe de limitações do *poder computacional* "realistas" (tempo de processamento, capacidade de memória, etc.).

Exemplo: One Time Pad

- Cifra demonstrada *incondicionalmente segura* por *Claude Shannon* (1949).
- Utiliza uma chave aleatória com o mesmo tamanho da mensagem a transmitir.
- Operações de cifra/decifragem são simplesmente o *xor* com a chave.

$$C_i = T_i \oplus K_i \qquad M_i = C_i \oplus K_i$$

- Chaves só podem ser utilizadas numa única operação de cifra.
- Segurança da cifra resulta do facto de o conhecimento do criptograma não resultar na *diminuição de incerteza* relativa ao conhecimento do texto limpo.
- Os problemas inerentes à geração e distribuição da chave tornam a cifra *inviável*.

Ataques

Vamos distinguir dois tipos de ataques dependendo das faculdades que se atribuem ao adversário:

- **Passivo:** atribui-se ao adversário unicamente a capacidade de escutar o canal de comunicação (i.e. de observar todo o tráfego que circula do canal).
- **Activo:** atribui-se adicionalmente para manipular a informação que circula no canal de comunicação (alterar/bloquear/injectar mensagens).

Ataque por *Força Bruta*

- Adversário percorre todo o *espaço de chaves* na expectativa de encontrar o texto limpo original.
- Pressupõe que:
  - existe suficiente redundância no texto original;
  - espaço de chaves é muito inferior ao espaço de mensagens.
- No entanto, estes condicionalismos são habitualmente cumpridos pelas aplicações correntes de cifras.
- É assim normalmente tido como *um ataque que é sempre passível de ser aplicado a uma cifra*.
- Mas cuja *viabilidade* se encontra condicionada pelo tempo que demora percorrer todo o espaço de chaves!!!
- Pode, portanto, ser ultrapassado adoptando “tamanhos razoáveis” para as chaves.

...sobre números grandes...

O tamanho do *espaço de chaves* é exponencial em relação ao tamanho da chave.

Tam. Chave	**Tempo** $(1\mu sec/test)$	**Tempo** $(1\mu sec/10^6 test)$
32 bit	35.8 min.	2.15 msec.
40 bit	6.4 dias	550 msec
56 bit	1140 anos	10 horas
64 bit	500000 anos	107 dias
128 bit	$5 * 10^{24}$ anos	$5 * 10^{18}$ anos

Exemplo de um Ataque

- Considere-se uma cifra *por substituição mono-alfabética*.

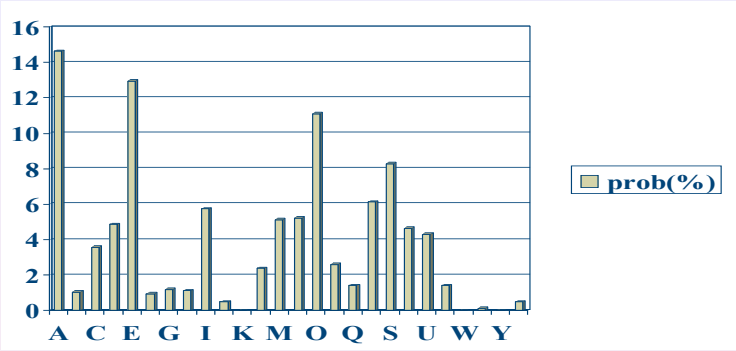
A	B	.....	Z
?	?	.....	?

- O número de possíveis chaves é então de $26! \approx 17.5 * 10^{24}$.
- Interceptou-se o seguinte criptograma:

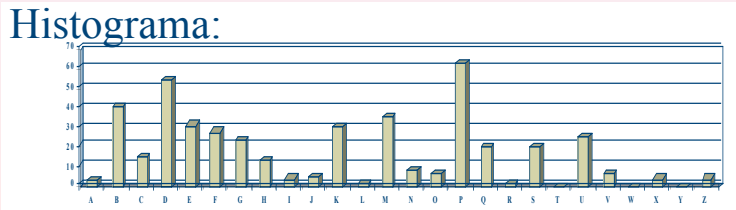
```
FPGFBNBVPKFBDMSBEMDMGUCDKDGUGDMUSPMMDBEFLEFEQDCPPGIDEXDCBKPMDHKPMPFQBUGPSUGHKEGPF
QBMPXPKSESEBSURBHKBHBMEQBFUFSDSBGHKPPFCECPHDKQPDHDKQPFDBADVEDFDCDDCEZPKLDZEDGMPPM
NDKPMDGDVPEMPDNUPFQDVDMPCPZGEFUQBMCPUGMEOPFSEBHPFBMBFBDUNPCDPXSEQDSDBCBBUQBFBCPMU
KNEKDGUGDMHPKMBFDNPFMCPBKENPGBIMSUKDSBGJUPGPFQKPQEVPSBFSEOEDIUOBMPGKPMQDUKDFQPMPX
SPFQKESBMPMMPMQKDZEDGUGDHPKNUFQDQPKKEVPOBJUPPJUPVDEMUSPCPKKPMHBFCEOAPMJUPFDBMDIED
PPOPMBOADKDGHDKDBHKDQBSBGEFJUEPQDSDB
```

- Sabe-se que a mensagem transmitida é um texto em Português.
- ?Como proceder à cripto-análise desta crifra?

- Explora-se o facto de, em Português, existirem diferentes probabilidades de ocorrência de letras nos textos.



- Por outro lado, também são distintas as probabilidades com que estas se agrupam (e.g. “as”, “os”, “es”, “que”, “nao”, ...)
- Assim, realizando a análise de frequências no criptograma obtemos:



- Podemos prosseguir por “palpites”: as letras *P*, *D* e *B* deverão corresponder ao *A*, *E* e *O*. Por outro lado, a existência de várias ocorrências dos pares *PM*, *PF*, *MP*, *JUP*, ... sugerem-nos a seguinte *decifragem parcial*:

```
ME-MO-O-E-MOAS-O-SAS-U-A-A-U-ASU-ESSAO-M--M--A-EE--A--A-O-ESA--ESEM-OU-E-U--
---EM-OSE-E-----O-U-O--O-OS--OMUM-A-O---EEM---E-A--EA-A--EMAO-A--AMA-AA---E-
-A--A-SEES-A-ESA-A-E-SEA-UEM-A-ASE-E---MU-OS-EU-S--EM--O-EMOSOMOAU-E-AE----
A-AO-OOU-OMO-ESU----A-U-AS-E-SOMA-EMS-EO---E-O-S-U-A-O-QUE-EM--E---E-OM----
A-U-OSE--ES-AU-AM-ESE--EM----OSESSES--A--A-U-A-E--UM-A-E----E-OQUEEQUE-A-SU-
E-E--ES-OM----ESQUEMAOSA--AEE-ESO--A-A--A-AO--A-O-O--MQU-E-A-AO
```

- ...que não parece fazer muito sentido!!! :-(

- Fazendo *backtracking* e tentando outra alternativa, obtinhamos:

```
NE-NO-O-ERNOAS-O-SAS-U-ARA-U-ASU-ESSAO-N--N-TA-EE--A--A-ORESA-RESENTOU-E-U--R--
ENTOSE-ER----O-U-O-RO-OS-TONUN-A-O--REEN---E-ARTEA-ARTENAO-A--ANA-AA---ER-A--A-
SEES-ARESA-A-E-SEA-UENTA-ASE-E---NUTOS-EU-S--EN--O-ENOSONOAU-E-AE---TA-AO-
OOUTONO-ESUR--RA-U-AS-ERSONA-ENS-EOR--E-O-S-URA-O-QUE-ENTRET--E-ON----A-U-OSE-
RESTAURANTESE--ENTR--OSESSESTRA--A-U-A-ER-UNTATERR--E-OQUEEQUE-A-SU-E-ERRES-ON---
-ESQUENAOSA--AEE-ESO--ARA--ARAO-RATO-O--NQU-ETA-AO
```

- Que finalmente nos conduz a:

```
NEMNOGOVERNOASCOISASMUDARAMUMASUCESSAOINFINITADEEMBAIXADORESAPRESENTOUMECUMPRIMEN
TOSEXERCICIOCUJOPROPOSITONUNCACOMPREENDIDEPARTEAPARTENAOHAVIANADAADIZERFAZIAMSEES
GARESAMAVEISEAGUENTAVASEDEZMINUTOSDEUMSILENCIOPENOSONOAUGEDAEXCITACAODOOUTONODESU
RGIRAMUMASPERSONAGENSDEORIGEMOBSCURACOMQUEMENTRETIVECONCILIABULOSEMRESTAURANTESEX
CENTRICOSESSESTRAZIAMUMAPERGUNTATERRIVELOQUEEQUEVAISUCEDERRESPONDILHESQUENAOSABIA
EELESOLHARAMPARAOPRATOCOMINQUIETACAO
```

Em Resumo...

- Na cripto-análise, explora-se *toda a informação* disponível, como sejam:
  - a natureza na mensagem transmitida;
  - informação parcial dessa mensagem;
  - histórico sobre a utilização da cifra (e.g. existência de mensagens cifradas com a mesma cifra/chave);
  - possíveis *vícios de utilização* da cifra (e.g. deficiências na escolha das chaves, etc.).
- Uma técnica de cripto-análise é tanto mais efectiva quanto se consiga afastar significativamente de um ataque por força bruta.

Classificação de Ataques a Cifras

Num ataque a uma cifra, o adversário é colocado perante o *desafio* de descobrir o texto limpo associado a um criptograma[1]. Dependendo do conhecimento adicional atribuído ao adversário, classificamos o ataque como:

- **criptograma conhecido** — o adversário só conhece o criptograma sobre o qual é desafiado.
- **texto limpo conhecido** — adicionalmente, o adversário conhece um determinado número de pares "texto-limpo/criptograma" (que não incluem o criptograma de desafio).

[1] Por vezes esse desafio é colocado sobre a forma de teste de *indistinguibilidade*: o adversário sabe que o criptograma resulta da cifra de uma de duas mensagens diferentes escolhidas por ele, só precisando assim descobrir qual das duas foi cifrada.

- **texto limpo escolhido** — o adversário pode escolher quais os textos limpos para os quais conhece os respectivos criptogramas (i.e. tem acesso à operação de cifra). Diz-se ainda que é *adaptativo* quando essa escolha pode ser condicionada pelo desafio (caso contrário, essa escolha é realizada antes da recepção do desafio).
- **criptograma escolhido** — o adversário pode escolher criptogramas para os quais pretende saber os textos limpos associados (desde que não seja o próprio desafio). Também aqui se distingue a versão *adaptativa* quando essa escolha depende do desafio.

Propriedades de Segurança

A criptografia é hoje utilizada para fornecer garantias referentes a um vasto leque de *propiedades de segurança*:

- **Confidencialidade:** garantir que o conteudo da mensagem só é do conhecimento dos intervenientes legítimos.
- **Integridade:** garantir que o receptor não “aceita” mensagens que tenham sido manipuladas.
- **Autenticidade:** assegurar a “origem” da mensagem.
- **Não repúdio:** demonstrar a “origem” da mensagem.
- **Anonimato:** não fornecer qualquer informação sobre a origem da mensagem.
- **Indentificação:** assegurar a “identidade” do interveniente na comunicação.
- **...**

Serviços e Protocolos Criptográficos

- Estamos normalmente interessado numa combinação de propriedades (e.g. num *canal seguro* entre duas partes pretende-se garantir a confidencialidade, autenticidade e integridade).
- Por outro lado, algumas das propriedades referidas não resultam directamente de uma técnica criptográfica específica, mas antes de uma conjugação de técnicas.
- Esta combinação de técnicas resultam tipicamente no que se designa por **protocolos criptográficos** — aí especificam-se as trocas de mensagens (e as técnicas criptográficas utilizadas) para se atingirem os fins pretredidos.
- A segurança de protocolos criptográficos (i.e. se eles realmente cumprem os requisitos para que foram desenvolvidos) não depende unicamente da segurança das técnicas que os suportam.

Criptografia e Segurança

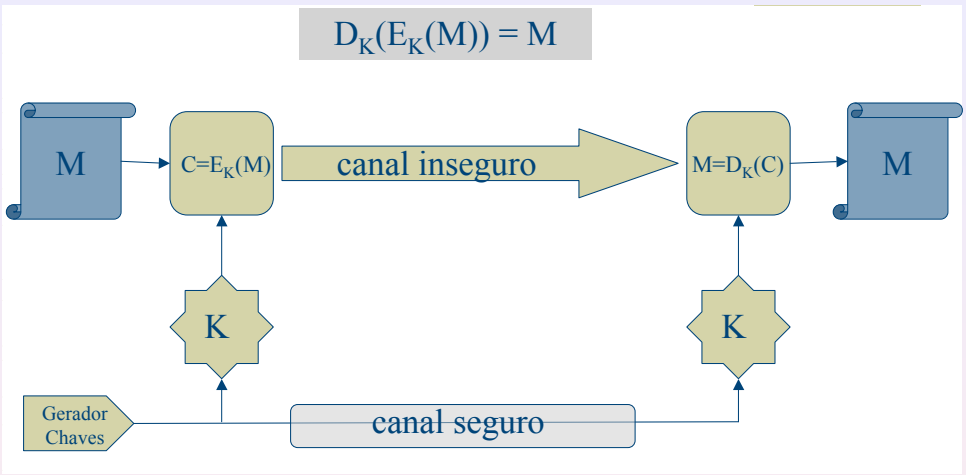
A segurança das técnicas criptográficas constituem um ingrediente fundamental e necessário na segurança de sistemas informáticos, mas não suficiente.

- Podemos distinguir (pelo menos) os seguintes níveis no estabelecimento da segurança de um sistema informático:
  - Técnica criptográfica;
  - Protocolo;
  - Implementação;
  - Utilização.
- Uma brecha de segurança em qualquer um destes níveis compromete a segurança de todo o sistema.

Breve Catálogo de Técnicas Criptográficas (e afins)

- Cifras.
- Assinaturas Digitais.
- Funções de *hash*
- *Message Authentication Codes*.
- ...

Cifras Simétricas



- A mesma chave é utilizada na operação de *cifra*/*decifragem*.
- Pressupõe por isso que, numa fase prévia à comunicação, se procedeu ao **acordo de chaves**.
- ...operação que “tipicamente” envolve a utilização de **canais seguros**.
- Exemplos: RC4; DES; IDEA; AES.

Geração (e manipulação) de chaves

- Um factor determinante para a segurança das técnicas criptográficas é a *qualidade das chaves* utilizadas.
- A sua principal característica é a *aleatoriedade* e *imprevisibilidade*.
- O tamanho (número de *bits*) depende da técnica concreta. Para cifras simétricas são comuns chaves de 40 a 256 bits.
- Na perspectiva da programação, é apropriado olhar para as chaves como *tipos opacos*.
- ...devendo ser prestada "toda a atenção" à sua manipulação (armazenamento; âmbito de vida na execução do programa; etc.)
- Os requisitos impostos impedem a utilização de *PINs* ou de *palavras/frases passe* directamente como chave de uma cifra moderna.
- ...mas existem métodos para derivar chaves a partir dessas *passwords* (*key derivation functions*).

Distribuição/Acordo de Chaves

- A pré-distribuição das chaves constitui a maior dificuldade na utilização das cifras simétricas, já que o estabelecimento de *canais seguros* é oneroso.
- Note que pode ser vista alguma *circularidade* neste domínio: *a criptografia pode ser utilizada para construir canais seguros, mas ela própria depende da existência de canais seguros*.
- Os *Protocolos de Distribuição de Chaves* fazem uso de uma "rede mínima de confiança" para distribuir as chaves entre os intervenientes.
- A criptografia assimétrica abriu novas perspectivas: o *Acordo de Chaves* — em vez de se gerar e distribuir a chave, define-se uma forma de cada interveniente "derivar" uma chave apropriada (sem que se ninguém mais possa ser capaz de derivar essa chave...).

Cifras Assimétricas

- Faz uso de chaves de *cifra/decifragem* "distintas" — *o conhecimento de uma não revela informação sobre a outra*.
- Só a chave para decifrar necessita ser *secreta* (chave privada).
- Permite tornear problema da distribuição de chaves — a chave para cifrar pode ser comunicada sem requisitos de confidencialidade.
- ...mas permanecem algumas dificuldades (autenticidade da origem...).
- Exemplos: RSA; El-Gamal.

Funções de Hash

- As funções de *Hash* criptográficas (ou *message digest*; *fingerprint*; etc.) permitem produzir um "resumo" de tamanho fixo a partir de uma mensagem de comprimento arbitrário.
- ...de tal forma que "não é viável" encontrar outra mensagem que disponha do mesmo resumo (*pre-image resistant*).
- Tratam-se, por isso, de **funções *one-way*** (não invertíveis).
- Exemplos: MD5; SHA-1.
- Trata-se de um exemplo de uma técnica fundamental em criptografia que, por si só, não dá resposta directa a nenhuma propriedade de segurança — *o seu poder resulta da combinação com outras técnicas*.

Message Authentication Codes (MACs)

- Um *MAC* pode ser entendido como uma função de hash cujo resultado depende, para além da mensagem, de um segredo (chave secreta).
- Garante assim a *integridade* de uma mensagem:
  - Quem envia a mensagem gera o MAC respectivo que envia junto;
  - O Receptor por sua vez gera o MAC da mensagem recebida e compara-o com o recebido.
  - Se alguém alterar a mensagem não poderá recalcular o MAC (não dispõe da chave).
  - Assim o receptor não aceitará a mensagem manipulada (porque não verifica o MAC).
- Exemplos: HMAC-MD5; HMAC-SHA1.

Assinaturas Digitais

- As **assinaturas digitais** permitem associar uma mensagem a uma determinada “entidade”.
- Cumprem assim um papel análogo ao das *assinaturas correntes* que associam documentos a pessoas.
- Ao nível das propriedades de segurança, estamos interessados em garantir:
  - **autenticidade:** o destinatário deverá confiar na identidade do signatário.
  - **integridade:** que o documento objecto da assinatura não é manipulado.
  - **não repudiável:** o signatário não poderá negar, posteriormente, que realmente assinou o documento.
- Exemplos: RSA; DSA.
- Pode ser entendido como o “contributo mais significativo” da criptografia assimétrica.

openSSL

- Autentico **canivete Suiço** para quem trabalha em criptografia.
- Originalmente concebido como uma biblioteca que implementa o protocolo *SSL*.
- ...mas disponibiliza uma *shell* que dá acesso à funcionalidade implementada (cifras, funções de hash, assinaturas, ...)
- Disponível para a generalidade das plataformas (Unix, MacOS, Windows).
- Apontadores:
  - `http://www.openssl.org`
  - `http://www.modssl.org`